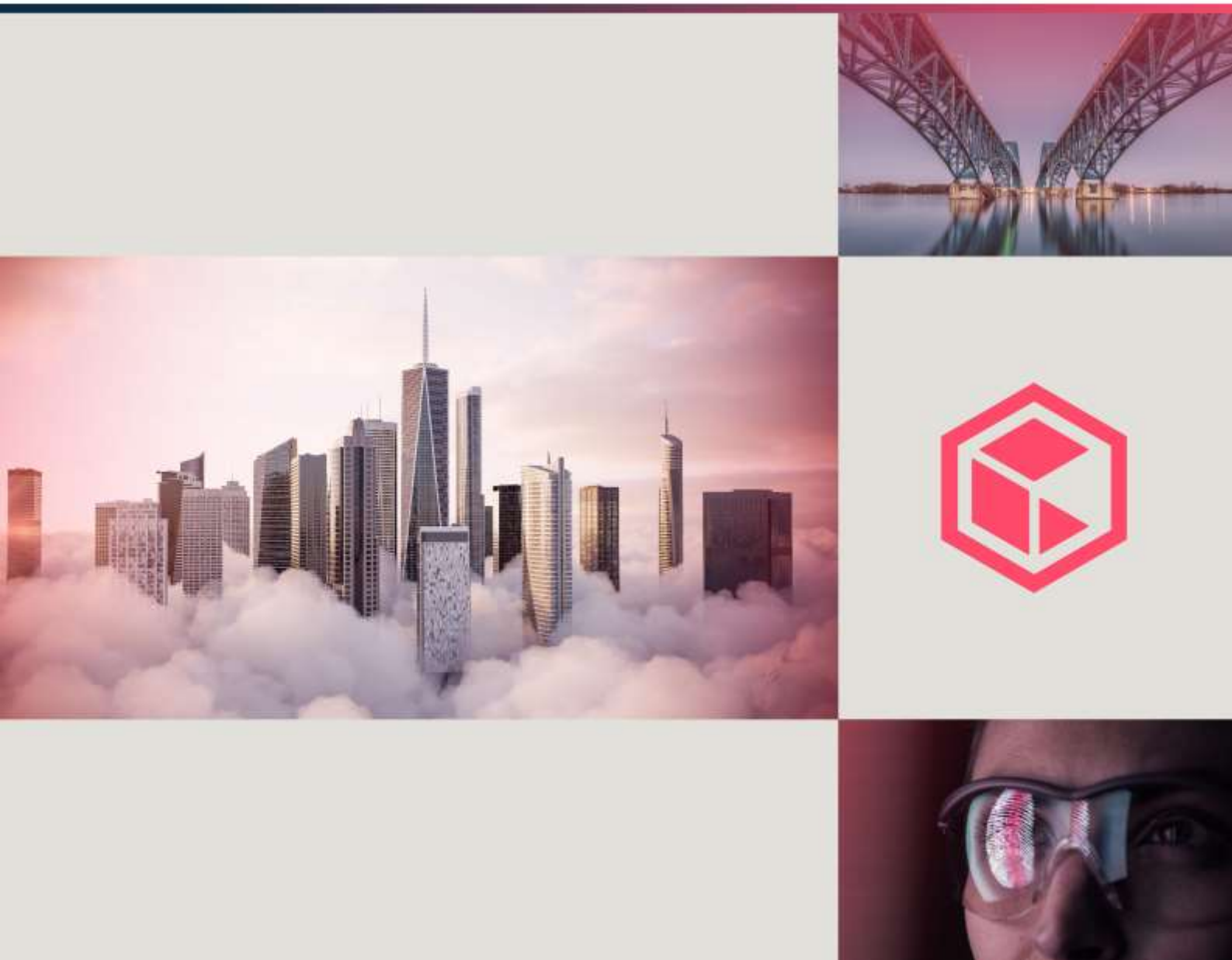# AWS Cloud Architecture Guide

Commvault Platform Release 2023E

August 2023

INTRODUCTION

As the modernization of business applications accelerates, the adoption of Amazon Web Services (AWS) can help you transform your IT solutions, operations, and culture to embrace the speed and flexibility of the Cloud. As a leader in the **2022 Gartner Cloud Infrastructure and Platform Services Magic Quadrant**, AWS is uniquely positioned to meet your Cloud business needs.

Commvault® software is designed to secure your AWS workloads, defend against cyber-security threats, and recover from unauthorized data tampering or resource failure. Commvault software is available from the AWS Marketplace, allowing you to start **trading fixed expense for variable expense**, and **benefitting from massive economies of scale** today.
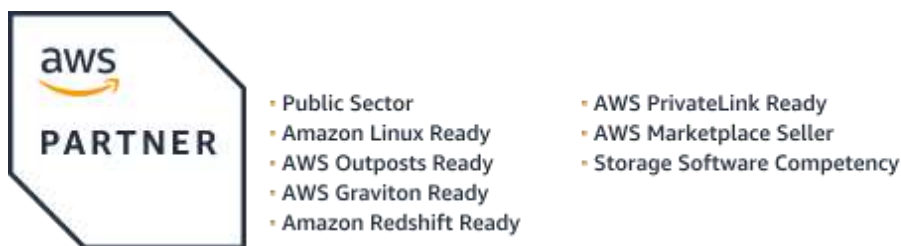
Commvault software is designed to help you **stop guessing (and paying) for capacity** by deploying minimal data management resources and scaling up/down based on observed demand. Commvault is enhanced to use AWS Graviton for the best price-performance of any general compute in AWS.

Commvault software with AWS helps you **increase speed and agility** by taking on the undifferentiated heavy lifting of data protection of Cloud services. Commvault will automatically discover workloads to protect, deploy temporary resources to perform protection, and power-up/power-down compute resources as required. This intelligent automation helps you **stop spending money running and maintaining data centers** (and data management resources in AWS), so you can focus on your customers.

> Commvault helps you leverage and benefit from AWS variable expense compute, economy of scale pricing, speed of resource deployment, agility, and massively global resource availability. Experiment, test, and adopt AWS services knowing you can stay protected with Commvault.

The AWS Cloud Architecture Guide (this document) leverages the **AWS Well-Architected Framework**, Amazon and Commvault best practices, and learnings from customer environments to streamline, simplify, and secure the protection of your AWS services and more.

Commvault has been integrated with AWS since the initial integration of Amazon S3 in 2008 and is a Storage ISV Competency partner in the **AWS Partner Network (APN)**. Commvault® software has been validated with many APN competencies, partner programs, and service-ready designations.

- Public Sector
- Amazon Linux Ready
- AWS Outposts Ready
- AWS Graviton Ready
- Amazon Redshift Ready
- AWS PrivateLink Ready
- AWS Marketplace Seller
- Storage Software Competency

Commvault® software has been validated by Amazon Solution Architects and deemed well-architected.

FOREWORD ON COST OPTIMIZATION IN AWS

Without a doubt, 2023 is the year of cost optimization, cost controls, and right-sizing as businesses look to adapt to changing economic conditions and market fluctuations. But **cloud financial management** has always been an underlying tenet of a <u>well-architected</u> cloud environment and key to any modern and mature *data strategy.*

So how can you cost-optimize with Commvault - Let us keep it simple:

Step 1: Start with the AWS Well-Architected Framework

Ensure you are architecting, designing, and operating your AWS and Commvault solutions per **AWS Well-Architected** architectural best practices. Get experience in the Well-Architected **hands-on labs**, and use the **AWS Well-Architected Tool** to periodically review and tune your environment as your needs or technology changes.

Commvault helps you build a Well-Architected *data strategy* built on the Commvault Backup & Recovery platform by providing prescriptive guidance in the **Well-Architected Framework** section of this guide. This includes guidance in the **Cost Optimization Pillar** section on how to consider cost during initial design, but also operate a cost-optimized Commvault-protected environment.

Step 2: Build for today, scale on observed demand.

Any *modern data strategy* needs to respond to the needs of the business to secure, defend, and recover applications per business and or regulatory requirements. Your business will be continually moving data from active operational use to inactive archival data stores, and finally destruction per business and industry regulations. Your data management infrastructure must therefore be built and managed for your *immediate data management needs*.

Put simply – deploy <u>minimum viable</u> data management resources and scale only when business data management objectives cannot be met. Commvault helps you complete a one-time or periodic task (i.e., perform deep data indexing of backup data for eDiscovery), then you can rightsize the deployed resources for business-as-usual (BAU) observed loads.

With the elasticity of AWS services, you do not need to place weeks, months, or years in advance. **Pay only for what you use and use that for which you are paying.**

Step 3: Measure continually for data-driven right-sizing

Staying in control of your AWS monthly invoice means continually measuring, reporting, and analyzing your resource usage, performance, and of course costs. Amazon provides detailed **Cost & Usage Reporting**, and **Cost Explorer**, and can notify you when resources exceed predefined business thresholds with **AWS Budgets**.

When resource consumption drops, you want to be aware so you can right-size resources. When resource consumption increases, you need to be able to communicate where the growth is occurring, and the corresponding benefit to the business.

Consider enhancing *operational runbooks* to review and apply right-sizing recommendations like **Amazon Reserved Instance recommendations**, **AWS Compute Optimizer recommendations**, and Amazon EBS improvements (from Compute Optimizer).

Remember, any <u>cost control exercise</u> will only be successful if there is adequate data to understand if a change has occurred and if a proposed change will have a positive or negative impact.

From Mathew Ericson (Pr. Product Manager, Commvault)

NOTICES

This document is provided for informational purposes only. It represents Commvault's current product offerings and practices as of the date of issue of this document, which is subject to change without notice. The responsibilities and liabilities of Commvault® to its customers are controlled by Commvault agreements, and this document is not part of, nor does it modify, any agreement between Commvault and its customers.

Table of Contents

Cloud Architecture Guide for AWS

This guide serves as an architecture guide for solutions architects and Commvault® customers who are building data protection and management solutions utilizing Amazon Web Services (AWS) and Commvault® software.

It includes cloud concepts, architectural considerations, and sizing recommendations to support Commvault® software in Amazon Web Services (AWS). The approach defined in this guide applies to both running Commvault solely in AWS and extending existing edge-based Commvault deployments into hybrid cloud architectures. The guide covers several common use cases for protecting AWS compute and container-based resources, cloud databases, and storage services. This guide also addresses how to seamlessly and securely migrate your applications to AWS and adopt an on-demand disaster recovery capability to AWS.

Currently, this guide delivers architecture considerations and sizing recommendations for Amazon Web Services (AWS). Guides for other public cloud environments are available as well at **docs.commvault.com**.

Acknowledgments

How this guide is structured

The Cloud Architecture Guide is structured to progress from well-architected architectural guidance to selecting a reference architecture, and then tuning your implementation using data management design principles and best practices.

You can consider the guide split into four (4) major sections or tasks you will perform while architecting for your cloud and/or hybrid data management with Commvault® and/or **Metallic.io**.

1	**Cloud Protection Coverage** *Get started quickly by leveraging existing AWS Marketplace solutions to protect your AWS services. Understand your data management responsibilities in Cloud.*	**Why Commvault?** **Getting started in AWS Marketplace** **Commvault protection of AWS Cloud Products** **Metallic protection of AWS Cloud Products** **Cloud Shared Responsibility**
2	**Well-Architected Guidance** *Review the AWS Well-Architected Framework architectural principles with guidance from Commvault and Commvault partner real-world experience. Select the Reference Architecture that meets your business resiliency and performance needs.*	**Zero trust architecture** **Well-Architected Framework** **Reference Architectures** **Ransomware**
3	**Designing and optimizing for AWS Cloud** *Take your data management to the next level with design best practices that improve performance, increase resiliency, and reduce cost.*	**Design and best practices**
4	**Intelligent Data Management** *Explore the intelligent data management use-cases available with Commvault & Metallic. Adopt automation for hands-off/lights-off management at scale.*	**Data Management Use-Cases** **Adopting APIs for Automation** **Multi-cloud mobility**

Terminology

Commvault is a multi-faceted data management solution that contains multiple components that can be consolidated for reduced infrastructure footprint or separated for improved scalability. The following are some common terms that will appear within this document and their definition.

AN **Access Node** refers to the component that is responsible for connecting to your primary application and capturing the data to be protected. An Access Node runs a Commvault software package for accessing the source application (i.e., Virtual Server Agent, MySQL Agent, PostgreSQL Agent, Cloud Apps Agent).

MA **MediaAgent** refers to the component that is responsible for data transfer and indexing to facilitate optimized backup and recovery operations. A MediaAgent runs the Commvault MediaAgent software package which communicates directly with secondary storage libraries (disk, cloud, tape).

Any reference to an Access Node (within this document) refers to a system that performs the role of 'Access Node' and 'MediaAgent' unless stated otherwise.

Why Commvault?

Consider the following capabilities when assessing your **data management** needs in AWS.

Broad protection for cloud-native, SaaS, and traditional workloads

*Commvault has the **broadest industry support** for cloud-native, SaaS, and traditional applications, hypervisors, and storage arrays. Backup is not often the first capability productized by new service providers, Commvault is there to perform protection for your current and future applications.*

Protection for data, regardless of location

*Commvault protects all AWS **regions** and availability zones including **AWS Outposts** and **AWS Local Zones**. Commvault automatically discovers your cloud workloads by tag, then manages the AWS snapshot lifecycle to meet your business rules and cost objectives.*

Cloud-native protection – by default

Commvault orchestrates the creation of cloud-native snapshots (Amazon EC2, Amazon EKS, Amazon EBS, Amazon RDS, Amazon DynamoDB, Amazon DocumentDB, Amazon Redshift) with encryption. Commvault automates snapshot copies within and across regions and accounts.

Cloud mobility without compromise

Workloads in AWS may be ideal today, back on-premises tomorrow (AWS Outposts) and in another cloud provider for new dev/test initiatives (Azure, Google, Oracle). Commvault provides mobility for Containers, VMs, Databases, and Application data across clouds – meaning **flexibility** for your business.

Self-service backup, recovery, and disaster recovery

*Let **authorized** end users perform the recovery using the **Commvault Command Center™** to self-service simple and complex recovery needs, without specialist AWS skills or knowledge.*

Recovery readiness and insight

*Commvault Command Center™ provides visibility into **SLA compliance, backup/recovery history, and data access requests** (eDiscovery). Your business will know what data it has, whether it is protected, and whether it represents a risk (PII data, insecure data, orphaned data).*

Information Management across your entire data estate

Commvault provides insight into your data through Risk Analysis (formerly File Storage Optimization and Data Governance), and eDiscovery and Compliance capabilities. Visualize and identify data risks, and inefficiencies across protected and live data sources.

Quick Links to common tasks

Looking to get started quickly? Refer to **Cloud Feature Support for AWS** for protected services and best practices.

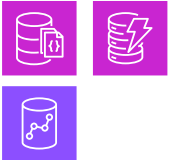|	
AWS Identity and Access Management (IAM)	• IAM Role Templates **JSON Templates for IAM Role Definition and User Permissions** o Amazon EC2 and RDS protection **amazon_restricted_role_permissions.json** o Amazon S3 backup & cloud libraries **amazon_s3.json** o Amazon S3 on Outposts backup **amazon_outposts.json** o VM Conversion **amazon_permission_conversion.json** (enhanced VM migration) o Amazon VM Import/Export **amazon_permission_vmimportexport** • Amazon RDS **amazon_rds_backup_restore_permissions.json** • Amazon Redshift **amazon_redshift_backup_restore_permissions.json** • Amazon DocumentDB **amazon_documentdb_backup_restore_permissions.json** • Amazon EC2 **amazon_DB_FS_backup_restore_permissions.json** (in-guest agent)
Amazon S3 – Cloud Libraries	• How to configure cloud storage – **Configuring Cloud Storage** • Required Amazon IAM permissions – **amazon_s3.json** • Using an IAM Role – **Adding an S3 cloud library with an attached IAM Role Policy** • Using STS:AssumeRole – **Adding an S3 cloud library with STS Assume Role** • Enabling Amazon S3 Object Lock – **Enabling S3 Object Lock on Amazon cloud libraries**
Amazon EC2 Protection	• Protecting Amazon EC2 Instances – **Creating a VM Group for AWS** • How IAM actions are used for protection – **Amazon Web Services Permissions Usage** • Replication for Amazon EC2 backups – **Replication**
Amazon Container Protection (EKS, EKS-D, Red Hat OpenShift)	• **Protecting Elastic Kubernetes Service** (containers, persistent volumes, secrets) • Dynamic application discovery **Auto-protecting containers by label selector** • Application migration **Application and data migration (cross-cluster, cross-region)**
VMware Cloud on AWS	• Protecting VMware Cloud – **VMware Cloud on AWS**

Amazon RDS Protection	• Protecting Amazon RDS – **Amazon RDS**
	• **Using native snapshots** Aurora, MariaDB, Microsoft SQL Server, MySQL, Oracle, Postgres
	• **Using export/dump** Aurora – MySQL/PostgreSQL, MariaDB, Microsoft SQL Server, MySQL, PostgreSQL
Amazon DocumentDB / Amazon DynamoDB / Amazon Redshift Protection	• Amazon DocumentDB protection – **Amazon DocumentDB**
	• Amazon DynamoDB protection – **Amazon DynamoDB**
	• Amazon RedShift protection – **Amazon Redshift**
Amazon S3 Protection	• Amazon S3 protection – **Getting Started with Amazon S3**
	• Commvault + Amazon S3 best practices – **Best Practices for Amazon S3**
AWS Outposts Protection	• AWS Outposts protection – **AWS Outposts** (includes Amazon EC2, Amazon RDS, Amazon EKS, and Amazon S3 on Outposts)
Amazon FSx / Amazon EFS Protection	• Protecting Amazon FSx for Windows – **Amazon FSx for Windows File Server**
	• Protecting Amazon EFS – **AWS EFS (Amazon Elastic File System)**
Amazon VM Import/Export Integration	• Migrating to AWS with VM Import/Export – **VM Conversion Using the Import Method**

Commvault-enhanced VM Conversion **recommended.**	• Accelerated migration to AWS with VM Conversion – **Cross-Hypervisor Restores**
	• VM Conversion pre-requisites – **Converting to Amazon**

Disaster Recovery	• Replicating Amazon EC2 Instances – **Live Sync Replication for Amazon**
	• Replicating Applications in AWS – **Application-Aware Backups**
	• Replication pre-requisites – **Considerations for Amazon Replication**

| **App migration**	• Database migration to Amazon EC2 – **Oracle** / **SQL Server**
	• Database migration to Amazon RDS – **Oracle Database Application Migration to an Amazon RDS Database**
	• **Postgres database migration**, and **MySQL/MariaDB database migration**

Getting started in AWS Marketplace

Getting started with Amazon Marketplace

Commvault publishes AMI images for both a CommServe® instance and Access Node (MediaAgent + Virtual Server Agent) within the Amazon Marketplace.

Navigate to the **Commvault Marketplace Page** to obtain the image for your environment.

Need some help getting started, view the following quick start – **Getting Started with Commvault in Amazon Marketplace**

Commvault publishes the following products within **AWS Marketplace Infrastructure Software Backup & Recovery**
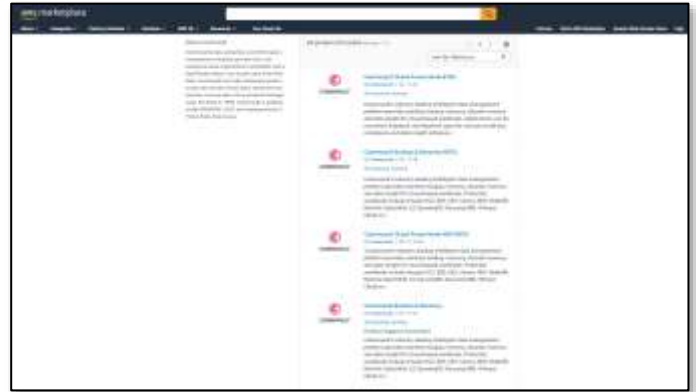
CloudFormation Template

- **Commvault Backup & Recovery** deploys Commvault Backup & Recovery on a single Amazon EC2 Microsoft Windows Server instance within an existing VPC with all dependencies. You may choose from a pre-defined set of annual software subscription sizes during setup. Additional usage will be observed and metered daily.

- **Commvault Backup & Recovery BYOL** deploys Commvault Backup & Recovery on a single Amazon EC2 Microsoft Windows Server instance within an existing VPC with all dependencies. The solution provides a Commvault 60-day license for testing, trials, and proof of concept (POC) initiatives. You may purchase a license after the 60-day trial expires or deploy a Commvault Backup & Recovery instance (see above).

> ⓘ **Note**
>
> Commvault Backup & Recovery BYOL will deploy Amazon EC2, Amazon EBS, AWS KMS, Amazon S3, Amazon VPC resources (optional), and Amazon CloudWatch alarms. Review the pricing page for each service to understand expected costs.

Amazon Machine Image

- **Commvault Cloud Access Node ARM BYOL** deploys a consolidated MediaAgent, Virtual Server Agent, and Cloud Apps installation running on Amazon Linux 2 on AWS Graviton instances. This instance is the default instance type for **Automatic scaling for Amazon Access Nodes**

- **Commvault Cloud Access Node BYOL** deploys a consolidated MediaAgent, Virtual Server Agent, and Cloud Apps installation running on Red Hat Enterprise Linux (RHEL) 7.9 on AWS EC2 Intel or AMD instances.

Purchasing Commvault Backup & Recovery via AWS Marketplace

Commvault is available for purchase in the **AWS Marketplace** via the **Commvault Backup & Recovery** and **Commvault Professional Services** product(s). AWS Marketplace modernizes your software acquisition, testing, and governance across your organization. With Commvault Bring Your Own License (BYOL), AMI usage, and

Professional Services products, you can take control of your cloud data, with the added benefit of Commvault Subject Matter Experts (SMEs) assisting you via Technology Consulting, Enterprise Support, and/or Managed Services.

The following is a list of available products.

Product Name	Product Type	Description
Commvault Backup & Recovery	Server	**Commvault Backup & Recovery** is part of Commvault's Intelligent Data Services Platform that enables organizations to proactively simplify and manage the complexity of enterprise data. Broadest coverage for cloud, SaaS, and on-prem protection
Commvault Technology Consulting	Professional Services	Commvault helps customers assess, design, implement, and maintain data management solutions that deliver immediate value and sustainable results. Commvault technical consultants ensure that your data management environment is designed for optimal results, configured quickly, and easy to maintain.
Commvault Enterprise Support	Professional Services	Commvault's Enterprise Support is designed for those enterprise customers looking to gain improved business value from their Commvault deployment and help ensure full confidence that their data management environment will deliver when it is needed most.
Commvault Managed Services	Professional Services	Commvault Remote Managed Services provides secure, reliable, and cost-effective remote monitoring and management of your data protection environment. Retain full ownership of your data management infrastructure, including hardware and software, while we provide secure service delivery.
Commvault Training	Professional Services	We get it – your time is valuable. Learn skills to effectively manage your Commvault environment and give your career a boost. We offer content for learners at all levels. Our On-Demand Learning Library is free for customers and partners, or you can purchase credits for self-paced or instructor-led training.

Automating Deployment with AWS CloudFormation

Commvault automates the deployment of **Commvault Backup & Recovery** (BYOL, AMI usage) using **AWS CloudFormation** Infrastructure as Code (IaC) automation.

Commvault deploys within an existing VPC and creates/configures the following resources during deployment.
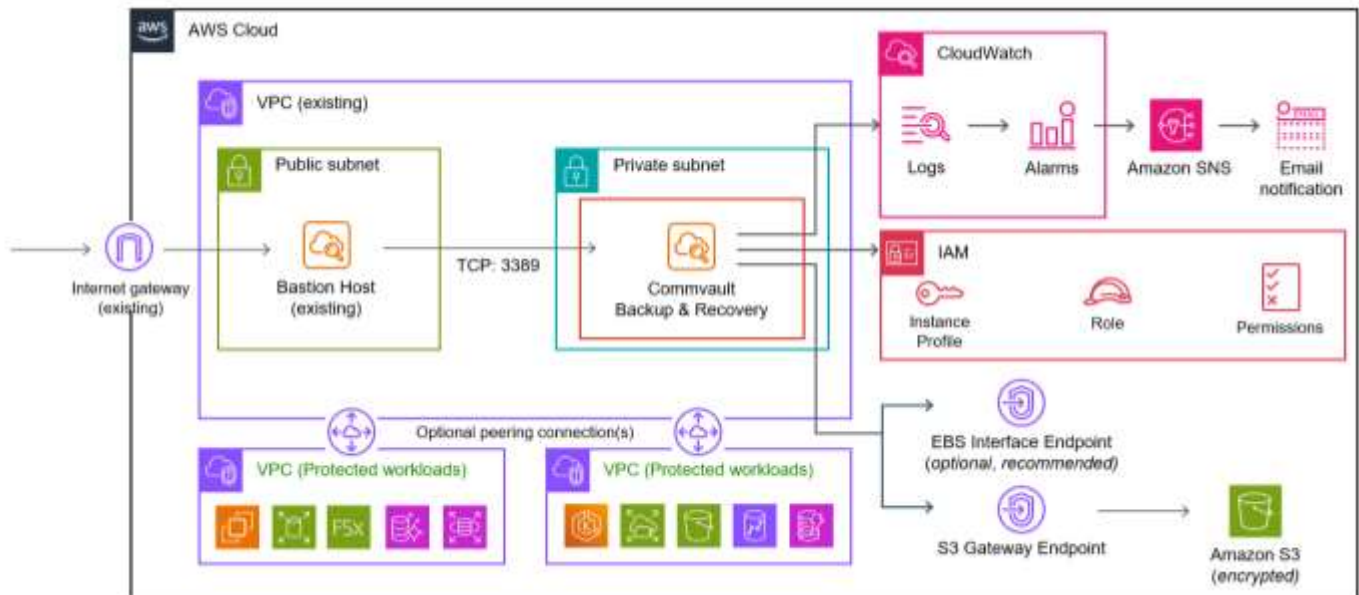
Commvault deploys the following components:

- An encrypted **Amazon EC2 instance** running Commvault Backup & Recovery within an existing VPC and Subnet, with associated **Security Group** controls.

- All require **AWS Identity & Access Management Roles, Instance Profiles, and Policies** to enable Commvault data protection and management activities (customizable to customer protection needs).

- An encrypted **Amazon S3 bucket** for the creation of a Commvault Cloud Library for backup copies.

- Multiple **Amazon CloudWatch** alarms monitor instance inactivity and disk space consumption, including email alerts to the administrator.

- (optional) An **Amazon VPC endpoint** (gateway type) for accessing Amazon S3 service.

- (optional) An **Amazon Elastic IP** (EIP) for accessing the Amazon EC2 via a static public IP address.

Remote Access / Bring Your Own Software

Commvault software may be installed in an Amazon EC2 instance that matches the hardware and operating system requirements detailed in Commvault documentation (docs.commvault.com). The most current procedures for software deployed are listed in the **documentation**.

Installation Basics

The following links cover the steps to install the Commvault CommServe, MediaAgents, and Access Nodes (Virtual Server Agents) on one or multiple Amazon EC2 instances. Commvault recommends using pre-built and configured AWS Marketplace images to accelerate deployment and alignment to Commvault and AWS best practices.

- **Installation Overview**
- **Installing the CommServe**
- **Installing the MediaAgent**
- **Installing the Virtual Server Agent (AWS)**

CommServe Disaster Recovery Solution Comparison

Commvault provides multiple solutions to recover from a CommServe instance loss. A Disaster Recovery of a Commvault CommServe® instance may range from building new and restoring previous data to periodic replication and failover to a DR instance, Refer to **CommServe Disaster Recovery** for a comparison of the options.

Pre-Packing Commvault® Software Within a VM Template

Commvault® software may be packaged inside Standard Operating Environment (SOE) images or Amazon Machine Images (AMIs), using a Decoupled install. This means that the software agents are deployed within the instance but will only be activated upon first-boot and registration with the CommServe instance.

For more information, please refer to the Installing the Custom Package instructions within the documentation:

Custom Packages

Infrastructure as Code provisioning and configuration management

For environments using Infrastructure as Code (IaC) toolsets such as Puppet, Chef, or Ansible, Commvault® supports deployment methods that allow administrators to both control agent deployment and configuration to provide an automated deploy-and-protect outcome for applications and servers.

Refer to the Commvault ansible library to automate your Commvault operations **github.com/Commvault/ansible**

Automate your RESTful development and testing with the Commvault POSTMAN collection **github.com/Commvault/Rest-API-Postman-Collection**.

Use the Commvault Python SDK to automate repeatable Commvault operational tasks **github.com/Commvault/cvpysdk**.

For more information on creating an unattended installation package for inclusion in a recipe, please refer to the Unattended Installation guide within Commvault documentation:
Unattended Installation

For more information on using Commvault® software's XML / REST API interface to control configuration post-deployment, please refer to the online documentation links below to review the options available for each iDataAgent:

- **REST API – Overview**

- **Command Line – Overview**

Getting started with Amazon S3 for backup data

Amazon S3 provides a scalable, highly available, secure, and performant object storage service to house your backup data from protecting AWS resources and edge-based on-premises locations. Amazon S3 provides a mixture of storage classes aimed at frequent use and infrequent use and very infrequent use (archival) usage.

In hybrid environments, Commvault recommends that your primary backup copy is written to a local storage device for optimal resource performance (i.e., **Commvault HyperScale X**). Amazon S3 can then be configured as a secondary or auxiliary copy location, providing offsite protection while benefiting from S3 scalability and durability. The secondary copy is replicated periodically as an encrypted network-optimized **Deduplicated Accelerated Streaming Hash (DASH) Copy**.

The link below lists all the supported cloud storage targets, including Amazon S3 and each supported storage class:

- **Supported Cloud Storage**

The link below covers cloud storage target setup and management.

- **Cloud Storage – Overview**

Unsupported Cloud Storage Configurations

If a Cloud Storage target is not listed in the Cloud Storage Support table, but the cloud storage endpoints are publicly accessible and provide either an Amazon S3-compatible or OpenStack-compatible REST API, you can verify the compatibility of the storage offering with Commvault.

Depending upon your cloud device type you may choose to verify the compatibility between:

- **Amazon S3 supported vendors and Commvault®**
- **OpenStack object storage supported vendors and Commvault®**

If you don't see your preferred Cloud Object Storage provider listed, please see **Verifying Cloud Storage Product Compatibility with Commvault**. Commvault supports *Amazon S3* and *OpenStack Swift* API interfaces for object storage but requires a partner validation test matrix to be completed before listing a new object storage solution supported.

Commvault protection of AWS Cloud Products

Commvault protects the following **AWS Products** and integrated services. Protecting your business-critical workloads is part of your **shared responsibility** in AWS. Note that each AWS product will only appear once.

Products will be listed with the following protection coverage; a **blank entry** indicates no protection support.

P the service is **protected** with native integration with the service APIs.
(*i.e., Use Amazon RDS* `CreateDBSnapshot`*, and* `CopyDBSnapshot` *to create backups*).

I the service is protected with native application **integration**.
(*i.e., Using the* `mysqldump` *binary to connect to the database and create a DB export*).

+ the service data is stored or exported to another protected AWS service.
(i.e., *Protection* **AWS CloudTrail logs** *storage in Amazon S3 using the* **Commvault Amazon S3 integration**).

Analytics

Amazon Athena	+	Amazon OpenSearch Service	+
Amazon CloudSearch	+	Amazon QuickSight	+
Amazon DataZone	+	Amazon Redshift	**P**+
Amazon EMR	**PI**+	AWS Clean Rooms	+
Amazon FinSpace	+	AWS Data Exchange	+
Amazon Kinesis	+	AWS Data Pipeline	+
Amazon Kinesis Data Analytics	+	AWS Glue	+
Amazon Kinesis Data Firehose	+	AWS Glue DataBrew	+
Amazon Kinesis Data Streams	+	AWS Glue Elastic Views	+
Amazon Kinesis Video Streams	+	AWS Lake Formation	+
Amazon Managed Streaming for Apache Kafka	+		

Application Integration

Amazon API Gateway	+	Amazon Simple Queue Service (SQS)	+
Amazon AppFlow	+	AWS AppSync	+
Amazon EventBridge	+	AWS Console Mobile Application	
Amazon Managed Workflows for Airflow	+	AWS Express Workflows	+
Amazon MQ	+	AWS Step Functions	+
Amazon Simple Notification Service (SNS)	+		

Blockchain

Amazon Managed Blockchain	+	
Amazon Quantum Ledger Database (QLDB)	+	

Business Applications

Alexa for Business	+	Amazon Pinpoint APIs	+
Amazon Chime	+	Amazon Simple Email Service (SES)	+
Amazon Chime SDK	+	Amazon WorkDocs	+
Amazon Connect	+	Amazon WorkDocs SDK	+
Amazon Honeycode	+	Amazon WorkMail	+
Amazon Pinpoint	+	AWS Wickr	

Cloud Financial Management

AWS Application Cost Profiler	+	AWS Cost Explorer	+
AWS Billing Conductor	+	Reserved Instance Reporting	+
AWS Budgets	+	Savings Plans	
AWS Cost and Usage Report	+		

Compute

Amazon EC2 (including Spot Instances)	PI+	AWS ParallelCluster	I+
Amazon EC2 Auto Scaling	PI+	AWS Serverless Application Repo	+
Amazon EC2 Image Builder	+	AWS SimSpace Weaver	I+
Amazon Genomics CLI	I+	AWS Thinkbox Deadline	I+
Amazon Lightsail	+	AWS Thinkbox Frost	I+
Amazon Lightsail for Research	+	AWS Thinkbox Krakatoa	I+
AWS App Runner	+	AWS Thinkbox Sequoia	I+
AWS Batch	+	AWS Thinkbox Stoke	I+
AWS Compute Optimizer	+	AWS Thinkbox Xmesh	I+
AWS Elastic Beanstalk	+	AWS Wavelength	PI+
AWS Lambda	+	Bottlerocket	

	AWS Local Zones	PI+		Elastic Fabric Adapter	PI+
	AWS Nitro Enclaves	PI+		AWS HPC (NICE DVC)	
	AWS Outposts family	PI+		AWS HPC (NICE EnginFrame)	
	AWS Outposts rack	PI+		VMware Cloud on AWS (VMC) inc. Outposts	PI+
	AWS Outposts servers	PI+			

Contact Center

	Amazon Connect	+			

Containers

	Amazon ECS Anywhere	+		Amazon Elastic Container Service	+
	Amazon EKS Anywhere	PI+		Amazon Elastic Kubernetes Service	PI+
	Amazon EKS Cloud	PI+		AWS Fargate	+
	Amazon EKS Distro (EKS-D)	PI+		Red Hat OpenShift Service on AWS	PI+
	Amazon Elastic Container Registry	+			

Customer Enablement

	AWS IQ			AWS re:Post	
	AWS Activate			AWS Support	
	AWS Managed Services	PI+		AWS Training and Certification	
	AWS Professional Services				

Database

	Amazon Aurora	PI+		Amazon Neptune	+
	Amazon DocumentDB (MongoDB)	P+		Amazon RDS	PI+
	Amazon DynamoDB	P+		Amazon RDS Custom	PI+
	Amazon ElastiCache for Redis	I+		Amazon RDS on VMware	PI+
	Amazon Keyspaces (Cassandra)	+		Amazon Timestream	+
	Amazon MemoryDB for Redis	I+		AWS Database Migration Service	PI+

Developer Tools

Amazon CodeCatalyst	+	AWS CodeBuild	+
Amazon Corretto (Open JDK distribution)		AWS CodeCommit	+
AWS Application Composer	+	AWS CodeDeploy	PI+
AWS Cloud Control API	+	AWS CodePipeline	I+
AWS Cloud Development Kit (CDK)	+	AWS CodeStar	+
AWS Cloud9	P+	AWS Command Line Interface	+
AWS CloudShell	+	AWS Tools and SDKs	+
AWS CodeArtifact	+	AWS X-Ray	+

End User Computing

Amazon AppStream 2.0	+	Amazon WorkLink (WorkSpaces Web)	+
Amazon WorkSpaces	I+		

Front-End Web & Mobile

Amazon Location Service	+	Device Farm	I+
AWS Amplify	+		

Games

Amazon GameLift	+	AWS GameKit	+
Amazon GameSparks	+	Open 3D Engine	

Internet of Things

AWS IoT 1-Click	+	AWS IoT ExpressLink	+
AWS IoT Analytics	+	AWS IoT FleetWise	+
AWS IoT Button	+	AWS IoT Greengrass	I+
AWS IoT Core	+	AWS IoT RoboRunner	+
AWS IoT Device Defender	+	AWS IoT SiteWise	I+
AWS IoT Device Management	+	AWS IoT Things Graph	+
AWS IoT EduKit	+	AWS IoT TwinMaker	+
AWS IoT Events	+	FreeRTOS	+

Machine Learning

Amazon Augmented AI (A2I)	+		Amazon Polly	+
Amazon CodeGuru	+		Amazon Rekognition	+
Amazon CodeWhisperer	+		Amazon SageMaker	+
Amazon Comprehend	+		Amazon SageMaker Ground Truth	+
Amazon Comprehend Medical	+		Amazon SageMaker Studio Lab	+
Amazon DevOps Guru	+		Amazon Textract	+
Amazon Elastic Inference	PI+		Amazon Transcribe	+
Amazon Forecast	+		Amazon Translate	+
Amazon Fraud Detector	+		Apache MXNet on AWS	+
Amazon HealthLake	+		AWS Deep Learning AMIs	PI+
Amazon Kendra	+		AWS Deep Learning Containers	PI+
Amazon Lex	+		AWS DeepComposer	+
Amazon Lookout for Equipment	+		AWS DeepLens	+
Amazon Lookout for Metrics	+		AWS DeepRacer	+
Amazon Lookout for Vision	+		AWS Neuron	+
Amazon Monitron	+		AWS Panorama	+
Amazon Omics	+		TensorFlow on AWS	+
Amazon Personalize	+		TorchServe	+

Management & Governance

Amazon CloudWatch	+		AWS Health Dashboard	+
Amazon Managed Grafana	+		AWS Launch Wizard	+
Amazon Managed Service for Prometheus	+		AWS License Manager	+
AWS AppConfig	+		AWS Management Console	+
AWS Application Auto Scaling	I+		AWS OpsWorks (Chef, Puppet)	+
AWS Auto Scaling	I+		AWS Organizations	+
AWS Backint Agent			AWS Proton	+

AWS Chatbot	+	AWS Resilience Hub	+
AWS CloudFormation	+	AWS Service Catalog	+
AWS CloudTrail	+	AWS Service Management Connector (SMC)	
AWS Config	+	AWS Systems Manager	I+
AWS Control Tower	+	AWS Telco Network Builder	+
AWS Distro for OpenTelemetry	+	AWS Trusted Adviser	+
AWS Fault Injection Simulator	+	AWS Well-Architected Tool	+

Media Services

Amazon Elastic Transcoder	+	AWS Elemental Live	+
Amazon Interactive Video Service	+	AWS Elemental MediaConnect	+
Amazon Kinesis Video Streams	+	AWS Elemental MediaConvert	+
Amazon Nimble Studio	I+	AWS Elemental MediaLive	+
AWS Elemental Appliances & Software	I+	AWS Elemental MediaPackage	+
AWS Elemental Conductor	I+	AWS Elemental MediaStore	+
AWS Elemental Delta	I+	AWS Elemental MediaTailor	+
AWS Elemental Link	+	AWS Elemental Server	+

Migration & Transfer

AWS Application Discovery Service	I+	AWS Application Migration Service	+
AWS Transfer Family	+	AWS Migration Evaluator	+
AWS DataSync	I+	AWS Server Migration Service	+
AWS Migration Hub	+		

Networking & Content Delivery

Amazon CloudFront	+	AWS Direct Connect	+
Amazon Route 53	+	AWS Global Accelerator	+
Amazon VPC **new protection coverage**	P+	AWS Private 5G	+
Amazon VPC Lattice	+	AWS PrivateLink	+

24

AWS App Mesh	+		AWS Site-to-Site BPN	+
AWS Client VPN	+		AWS Transit Gateway	+
AWS Cloud Map	+		AWS Verified Access	+
AWS Cloud WAN	+		Elastic Load Balancing	+

Quantum Technologies

Amazon Bracket	+

Robotics

AWS RoboMaker	+

Satellite

AWS Ground Station	+

Security, Identity, & Compliance

Amazon Cloud Directory		AWS Firewall Manager	+
Amazon Cognito	+	AWS IAM Identity Center	+
Amazon Detective	+	AWS Identity & Access Management (IAM)	+
Amazon GuardDuty	+	AWS Key Management Service (KMS)	+
Amazon Inspector	+	AWS Network Firewall	+
Amazon Macie	+	AWS Private Certificate Authority	+
Amazon Security Lake	+	AWS Resource Access Manager	+
Amazon Verified Permissions	+	AWS Secrets Manager	+
AWS Artifact	+	AWS Security Hub	+
AWS Audit Manager	+	AWS Shield	+
AWS Certificate Manager	+	AWS Signer	+
AWS CloudHSM	+	AWS WAF	+
AWS Directory Service	I+		

Serverless

AWS Lambda	+	AWS Fargate	+

Amazon API Gateway	+	Amazon Simple Queue Service	+
Amazon EventBridge	+	AWS Step Functions	+
Amazon SNS	+	AWS AppSync	+
AWS Simple Storage Service S3	PI+	Amazon DynamoDB	P+
Amazon Elastic File Storage (EFS)	I+	Amazon RDS Proxy	I+
Amazon Aurora Serverless	P	Amazon Redshift Serverless	+
Amazon Neptune Serverless	+	Amazon OpenSearch Serverless	+

Storage

Amazon EBS	PI+	Amazon S3 on Outposts	PI+
Amazon Elastic File Storage (EFS)	I+	Amazon Simple Storage Service Glacier	PI+
Amazon File Cache	I+	AWS Backup	
Amazon FSx	I+	AWS Elastic Disaster Recovery	
Amazon FSx for Lustre	I+	AWS Snowball	PI+
Amazon FSx for NetApp ONTAP	PI+	AWS Snowball Edge	PI+
Amazon FSx for OpenZFS	I+	AWS Snowcone	PI+
Amazon FSx for Windows Server	I+	AWS Snowmobile	PI+
Amazon S3	PI+	AWS Storage Gateway	I+

AWS CloudTrail and Amazon CloudWatch Support

Amazon provides **AWS CloudTrail** to log user activity and API usage, and **Amazon CloudWatch** to provide observability over AWS resources and applications. Commvault protects AWS CloudTrail and Amazon CloudWatch logs, events, and metrics stored in Amazon S3, and optionally to immutable storage in **Amazon S3 Object Lock storage**.

- **Amazon S3 (Simple Storage Service) protection**

In the detailed per-service protection coverage map (below), AWS CloudTrail and Amazon CloudWatch protection is indicated below the service icon with a **supported tick** (✓) or an **unsupported dash** (-), for example, the **Amazon RedShift** service:

- Supports AWS CloudTrail **https://docs.aws.amazon.com/redshift/latest/mgmt/logging-using-cloudtrail.html**
- Supports Amazon CloudWatch **https://docs.aws.amazon.com/redshift/latest/mgmt/metrics-listing.html**

Amazon Redshift

✓ / ✓

Analytics

Analytics	
Amazon Athena ✓ / ✓	Commvault protects the Amazon S3 data sources that **Amazon Athena** analyzes, and Amazon Athena **query output files** stored in S3. • **Amazon S3**.
Amazon EMR ✓ / ✓	Commvault protects the **Amazon EMR** analytics services (provisioned) that leverage open-source Apache Spark, Hive, Presto, and other big data analytics frameworks. Commvault protects Amazon EMR compute instances for the following deployment types: • **Amazon EC2 and Amazon EBS volumes**. • **Amazon EKS**. • **AWS Outposts**.
Amazon Redshift ✓ / ✓	Commvault protects **Amazon Redshift** instances (provisioned) using AWS-native snapshots in multi-account, multi-region deployments. • **Amazon Redshift**.
Amazon Kinesis ✓ / ✓	Commvault protects **Amazon Kinesis** (including **Amazon Kinesis Video Streams**, and **Amazon Kinesis Data Streams**) sources stored in AWS Services and on-premises in traditional Network Attached Storage (NAS). Some common **data producers** Commvault protects include: • **Amazon DynamoDB**. • **Amazon Aurora**. • **Amazon CloudWatch** logs, metrics & events are stored in Amazon S3. • **Network Attached Storage** (NAS). Additionally, Commvault protects **Amazon Kinesis Data Firehose** destinations, such as: • **Amazon S3**. • **Amazon Redshift**. • **Amazon OpenSearch Service**. • **Splunk**. • **MongoDB Cloud**.
Amazon OpenSearch Service ✓ / ✓	Commvault protects **Amazon OpenSearch Service** instances, built on open-source Elasticsearch, by protecting **index snapshots** (automated, manual) stored in Amazon S3. • **Amazon S3**.

Amazon QuickSight ✓ / ✓	Commvault protects **Amazon QuickSight** serverless Business Intelligence (BI) dashboards and reports that are **exported as PDFs**, **exported as CSV/Excel** for sharing and long-term compliance retention. Commvault protects data from Standard and Enterprise editions. Additionally, Commvault protects data sources, such as **Amazon S3 analytics data** located in the Amazon S3 service. • **Amazon S3**.
AWS Glue DataBrew ✓ / ✓	Commvault protects both the source and *brewed* output of **AWS Glue DataBrew** data preparation workflows. AWS Glue supports ingest of `.csv`, `.xlsx`, JSON, Apache ORC, and Apache Parquet files that may be stored on Commvault-protected Amazon S3, Amazon FSx, or Amazon EFS, or on-premises. Additionally, AWS Glue DataBrew exports prepared datasets to Commvault-protected Amazon S3. • **Amazon S3**. • **Amazon FSx**. • **Amazon EFS**. • **Network Attached Storage** (NAS).

Data movement

AWS Glue ✓ / ✓	Commvault protects both structured and semi-structured data stored in **AWS Glue**-supported services including: • **Amazon Aurora** • **Amazon RDS for MySQL, Oracle, PostgreSQL, SQL Server** • **Amazon RedShift** • **Amazon DynamoDB**. • **Amazon S3**. • MySQL, Oracle, Microsoft SQL Server, PostgreSQL databases on **Amazon EC2**. AWS Glue also integrates with **Amazon Athena** and **Amazon Redshift Spectrum** which both consume and store results in Commvault-protected Amazon S3.
Amazon Managed Streaming for Apache Kafka (MSK) ✓ / ✓	Commvault protects fully managed **Amazon Managed Streaming for Apache Kafka** (serverless, provisioned) broker log destinations including: • **Amazon S3**. • **Amazon Kinesis Data Firehose** destinations. ⓘ **Note**: Amazon EC2 compute instances provisioned by Amazon MSK are not visible or accessible for data protection (see **Amazon MSK FAQs – Clusters**).

Data lake	
AWS Lake Formation ✓ / ✓	Commvault protects **AWS Lake Formation** secure data lake integrated services, including: • **Amazon S3**. • **AWS Glue**. • **Amazon Athena**. • **Amazon Redshift**. • **Amazon QuickSight**. • **Amazon EMR**.
AWS Data Exchange ✓ / ✓	Commvault protects **AWS Data Exchange** subscribed datasets **exported to** or stored in Amazon S3. • **Amazon S3**.

Predictive analytics and machine learning	
Amazon Deep Learning AMIs ✓ / ✓	Commvault protects pre-configured **Amazon Deep Learning AMIs** infrastructure and tools through native Amazon EC2 and Amazon EBS protection of your deep-learning instances. Additionally, Commvault protects deep learning results or output stored in Amazon S3. • **Amazon EC2 and Amazon EBS**. • **Amazon S3**.
Amazon SageMaker ✓ / ✓	Commvault protects **Amazon SageMaker** files stored in your `/home/sagemaker-user` folder (for Jupyter) and `/root` for your kernel. These files are stored in Amazon EFS. • **Amazon EFS**.

Application Integration

API Management	
Amazon API Gateway ✓ / ✓	Commvault protects **Amazon API Gateway** downstream services and **persistent data stores**, such as: • **AWS Lambda**. • **Amazon EC2**. • **Amazon Kinesis**. • **Amazon DynamoDB**.

AWS AppSync ✓ / ✓	Commvault protects **AWS AppSync** serverless GraphQL and Pub/Sub APIs integrated services and **persistent data stores**, such as: • **Amazon DynamoDB**. • **Amazon Aurora**. • **Amazon OpenSearch Service**. • **AWS Lambda**.
Event Bus	
Amazon EventBridge ✓ / ✓	Commvault protects your **Amazon EventBridge** event-driven applications by protecting both your source applications and EventBridge destinations (targets). This includes supported targets such as: • Amazon CloudWatch log groups that store logs in **Amazon S3**. • **AWS Glue** workflow and persistent storage locations. • **AWS Lambda** functions and their persistent storage locations. • **Amazon Kinesis** stream, and persistent storage locations. • **Amazon Redshift**.
Messaging	
Amazon Simple Notification Service (SNS) ✓ / ✓	Commvault protects Amazon pub/sub messaging, SMS, email, and mobile push notifications provided by **Amazon Simple Notification Service** (Amazon SNS), by protecting Amazon SNS target locations. This includes: • **AWS Lambda** functions and their **persistent data store** locations. • **Amazon Kinesis Data Firehose** event storage and **backup pipeline** locations, which stores notifications to Amazon S3, Amazon Redshift, Amazon OpenSearch Service, and other **persistent data stores** for backup and archival purposes.
Amazon Simple Queue Service (SQS) ✓ / ✓	Commvault protects Amazon SQS when message objects are persisted or use Amazon S3 as a message store. This is particularly useful when **storing and consuming messages up to 2GB**. • **Amazon S3**.
Amazon MQ ✓ / ✓	Commvault protects Amazon MQ services by protecting the shared storage (Amazon EFS) used by the Active MQ engine. • **Amazon EFS**.

No-code API integration	
Amazon AppFlow ✓ / ✓	Commvault protects the output **persistent data store** of Amazon AppFlow integrations, including: • **Amazon Redshift**. • **Amazon S3**. • **Salesforce**.

Workflows	
AWS Step Functions ✓ / ✓	Commvault protects the output of AWS Step Functions, some examples include: • **Data Processing and ETL Orchestration** to Amazon S3 or Amazon RedShift. • **Machine Learning Operations to Amazon S3**. • **Media Processing to Amazon S3**.
Amazon Managed Workflows for Apache Airflow (MWAA) ✓ / ✓	Commvault protects the output of Amazon Managed Workflows for Apache Airflow, supported plug-ins include: • **Amazon DynamoDB**. • **Amazon EKS**. • **Amazon Redshift**. • **Amazon S3**.

Blockchain

Amazon Managed Blockchain	Commvault protects Amazon Managed Blockchain EC2 infrastructure and certificates stored in S3, that are required to participate in the blockchain network (see **Set Up The Hyperledger Fabric Client**). • **Amazon EC2**. • **Amazon S3**.
Amazon Quantum Ledger Database (QLDB) ✓ / ✓	Commvault protects the output of Amazon QLDB System-of-record implementations that export transaction audit records into downstream systems such as: • **Amazon Redshift**. • **Amazon S3**.

Business Applications

Line of Business Applications

Amazon Pinpoint

Commvault protects **Amazon Pinpoint** data exported in flat-file or streamed to **Amazon Kinesis Dara Firehose streams** for analysis or long-term retention. This includes the following supported **Kinesis destinations**:

- **Amazon S3**.
- **Amazon RedShift**.
- **MongoDB Cloud**.

✓ / ✓

Productivity Applications

Amazon Honeycode

Commvault protects data from Amazon Honeycode zero-code applications by protecting AWS data stores accessed via Zapier, Amazon AppFlow, or direct API integration:

✓ / -

- **Amazon DocumentDB**.
- **Amazon DynamoDB**.
- **Amazon RDS**.
- **Amazon Redshift**.
- **Amazon S3**.

- **Azure DevOps**.
- **Github**.
- **Microsoft Dynamics 365 CRM**.
- **Microsoft Exchange**.
- **Microsoft Office 365**.
- **Microsoft SQL Server**.
- **Microsoft Teams**.
- **MongoDB**.
- **MySQL Database**.
- **PostgreSQL Database**.
- **Salesforce**.

Amazon Chime

Commvault protects Amazon Chime Call Detail Records, Log files, and optional CloudTrail **Audit Logs** located in Amazon S3.

- **Amazon S3**.

✓ / ✓

Amazon WorkDocs

Commvault protects Amazon WorkDocs when a synced copy of important WorkDocs folders is replicated to a supported storage location (e.g., **Automatically sync files from Amazon WorkDocs to Amazon S3**) or resides in a supported storage location (e.g., **Using Amazon FSx for Windows File Server with Amazon WorkSpaces**)

✓ / -

- **Amazon S3**.
- **Amazon FSx for Windows File Server**.

Amazon WorkMail

Commvault protects Amazon WorkMail CloudTrail audit logs and **Exported Mailbox content** stored in Amazon S3.

- **Amazon S3**.

✓ / ✓

Alexa for Business	Commvault may be monitored using Amazon Alexa voice-based commands.
	• **Using Voice-Based Monitoring for CommCell Environments with Amazon Alexa**.
✓ / ✓	

Cloud Financial Management

Organize	
AWS Billing Conductor	Resources created by Commvault may be easily identified for best practice cloud financial management using a mixture of:
AWS Cost Allocation Tags	• Shared data management **accounts and users**.
	• Standard **tags** are applied to Commvault-created resources.
AWS Cost Categories	• Categorization of Commvault-created resources based on IAM user, role, or other AWS service.
	AWS Cost and Usage Reports stores output in Amazon S3 which may also be protected:
	• **Amazon S3**.
✓ / ✓	

Report	
AWS Cost Explorer	Commvault data management infrastructure costs may be observed within AWS Cost Explorer, and the AWS Cost and Usage Report. Common costs incurred when utilizing Commvault multi-tenanted data management include:
AWS Cost and Usage Report	• Amazon EC2 runtime and network transfer costs for **power-managed MediaAgents**.
	• Amazon EC2 runtime and network transfer costs for **auto-scaled Access Nodes**.
	• **Amazon S3 and Amazon S3 Glacier** deduplicated backup and archive storage.
✓ / ✓	
	Commvault can also protect AWS Cost reports exported to Amazon S3 and Amazon S3 Glacier:
	• **Amazon S3**.

Access	
AWS Consolidated Billing	Commvault easily deploys within an existing AWS Organization and can leverage AWS Consolidated Billing, AWS Purchase Order Management, and AWS credit management.
AWS Purchase Order Management	Commvault can also protect AWS billing reports (**monthly cost allocation reports**) exported to Amazon S3 and Amazon S3 Glacier:
	• **Amazon S3**.
AWS Credits	
- / -	

Control	
AWS Cost Anomaly Detection - / -	Commvault is a consumer of AWS services within your AWS accounts and/or AWS Organization. Commvault compute, storage and network resources may be monitored, visualized, and alerted by the AWS Cost Anomaly Detection service. Alarms sent to an Amazon SNS topic and persisted for traceability may be protected when stored in: • **Amazon S3**. • **Amazon Redshift**. • AWS Lambda persistent data stores. • Amazon Kinesis Data Firehose data stores.
Forecast	
AWS Cost Explorer ✓ / ✓	Commvault is a consumer of AWS services within your AWS accounts and/or AWS Organization and may be visualized and forecasted using AWS Cost Explorer. Cost Explorer Reports and Forecasts may be exported into comma-separated value (CSV) files and stored in any of the following Commvault-protected storage services: • **Amazon S3**. • **Amazon FSx for Windows File Server**. • **Amazon EFS**.
Budget	
AWS Budgets - / -	Commvault is a consumer of AWS services within your AWS accounts and/or AWS Organization and may have event-driven and interactive budgets set, forecast, and automated **AWS Budgets Actions** configured with AWS Budgets. Budgets may be downloaded to a `.CSV` file and stored in any of the following Commvault-protected storage services: • **Amazon S3**. • **Amazon FSx for Windows File Server**. • **Amazon EFS**.
Purchase	
AWS Free Tier - / -	Commvault can consume and protect **AWS Free Tier** resources to minimize cloud resource fees where free tier resources are available.
AWS Reserved Instances (Reporting) - / -	Commvault supports **and recommends** the use and protection of **Amazon EC2** and **Amazon RDS** Reserved Instances. • **Amazon EC2**. • **Amazon RDS**.

AWS Savings Plans - / -	Commvault supports **and recommends** the use and **AWS Savings Plans** to achieve lower runtime costs for Amazon EC2, AWS Lambda, and AWS Fargate instances. Commvault recommends that **data management infrastructure** leverage a combination of **AWS Reserved Instances** and **AWS Savings Plans** for the best financial outcome for your critical data management services. A note on Commvault Cloud Cost Optimization Commvault power manages (shutdown) its **Cloud MediaAgents** for reduced runtime cost. Commvault auto-scales **Cloud Access Nodes** during backup activities and then terminates those resources when no longer required.

Rightsize	
AWS Cost Explorer Right Sizing Recommendations	Commvault supports **and recommends** the use of **AWS Cost Explorer Rightsizing Recommendations** for Commvault data management compute resources. Rightsizing recommendations may be downloaded to a `.CSV` file and stored in any of the following Commvault-protected storage services: • **Amazon S3**. • **Amazon FSx for Windows File Server**. • **Amazon EFS**.
Amazon S3 Intelligent-Tiering 	Commvault supports and recommends the use of **Amazon S3 Intelligent-Tiering storage class** for active and infrequently accessed backup data. Commvault recommends using the Frequent Access, Infrequent Access, and Archive Instant Access Tiers only. Commvault provides additional cost optimization for long-term retention and archival data via **Commvault Combined Storage Tier**, which should be used for archive or deep-archive use-cases. Commvault can also protect data residing in Amazon S3 Intelligent-Tiering buckets. • **Amazon S3**.

Compute

Instances	
Amazon EC2 **Amazon EC2 Spot** ✓ / ✓	Commvault protects, replicates, and recovers Amazon EC2 and **Amazon EC2 Spot** Instances with crash-consistency and/or application-consistency. A backup may consist of Amazon EBS snapshots, streamed backup copies, or a combination of both. **Agentless recovery** of files and folders into an EC2 instance is supported using Amazon Systems Manager. • **Amazon EC2**. • **Supported agents for Amazon** (application consistent protection with Amazon EBS snapshots). • **Application-Aware Backups** (agent-less application-aware protection). • **Disaster Recovery for Amazon EC2** (replicate and recover cross region).

Amazon Lightsail	Commvault protects **Amazon Lightsail** applications and websites that utilize object storage to store images, videos, or HTML files (i.e., Lightsail files **copied out to Amazon S3)**.
✓ / ✓	• **Amazon S3**.

AWS Batch	Commvault protects the results of AWS Batch jobs by protecting the persistent data store that results are written to. Commvault can also protect AWS Batch compute instances running on Amazon EC2.
✓ / ✓	• **Amazon EC2 instances and Amazon EBS volumes**. • **Amazon S3**. • **Amazon EFS**. • **Amazon FSx for Windows File Server**.

Containers

AWS App Runner	Commvault protects **AWS App Runner** generated **logging and monitoring** artifacts, including user action logs generated by AWS CloudTrail and service metrics and application logs generated by Amazon CloudWatch, stored in Amazon S3.
✓ / ✓	• **Amazon S3**.

Serverless

AWS Lambda	Commvault protects the output of **AWS Lambda** event-driven compute services, including
✓ / ✓	• File processing from/to **Amazon S3**. • Document processing from Amazon S3 to **Amazon DocumentDB**. • Stream Processing to **Amazon DynamoDB**. • IoT backends to **Amazon Redshift**.

Edge and hybrid

AWS Outposts	Commvault protects, migrates, and replicates Amazon EC2, Amazon EBS, Amazon RDS, Amazon EKS, and Amazon S3 data to, from, and between the AWS region and AWS Outposts.
✓ / ✓	• **Protecting EC2 Data in AWS Outposts**. • **Migrating EC2 Instances Between AWS Region and Outposts**.

AWS Snow Family	Commvault protects persistent storage locations on AWS Snow family devices via Amazon S3 or Network File System (NFS) protocol.
✓ / ✓	• **Protecting Amazon EC2 Instances**. • **Network Attached Storage (NAS)**.

AWS Wavelength	Commvault protects **AWS Wavelength** Amazon EC2 instances and Amazon EBS volumes running in a Wavelength zone, using AWS-native snapshots and Commvault streaming backups.

 ✓ / ✓	• **Amazon EC2 instances and Amazon EBS volumes**. Additionally, Commvault protects **AWS Wavelength compatible services** in the region: • **Amazon EKS**. • **AWS CloudTrail** buckets. • **AWS CloudFormation** buckets. • **Amazon DynamoDB**. • **Amazon RDS**.
VMware Cloud on AWS (VMC) ✓ / ✓	Commvault protects, migrates, and replicates VMware vSphere workloads within VMware Cloud on AWS. You can use Commvault one-time VM conversion or ongoing periodic replication to perform Disaster Recovery (DR) failover and failback between your VMC and edge-based vSphere clusters. • **VMware Cloud on AWS**. • **Disaster Recovery for VMware**.
AWS Local Zones ✓ / ✓	Commvault protects. replicates, and recovers **Amazon Local Zones** workloads running on Amazon EC2 and Amazon EKS compute instances and **many other services** (detailed below) within the Local Zone. • **Amazon EC2 instances and Amazon EBS volumes**. • **Amazon EC2 Application-integrated backup** (agent in-guest snapshot protection). • **Amazon EC2 Application-aware backups** (agent-less application-aware protection) • **Amazon EC2 Disaster Recovery** (between Local Zones and/or Regions). • **Amazon ECS** persistent storage locations. • **Amazon EKS** applications and persistent storage. • **Amazon FSx for Windows File Server**. • **Amazon RDS**.
Cost and capacity management	
AWS Compute Optimizer - / -	Commvault intelligent data management services running on Amazon EC2 compute instances and Amazon EBS persistent storage may be tuned based on **AWS Compute Optimizer** recommendations. Commvault recommends the use of C, M, R, and T class instances to meet Commvault minimum CPU and RAM requirements for core CommServe® instance, MediaAgent(s), and Access Nodes. Commvault recommends leveraging AWS Compute Optimizer recommendations calculated over the maximum amount of history (3 months) to ensure cyclic data management activities are considered in optimization recommendations. Optimization recommendations may be **exported** to comma-separated values (`.CSV`) files stored in Commvault-protected **Amazon S3** for historical auditing and reporting.

AWS Elastic Beanstalk ✓ / ✓	Commvault protects **AWS Elastic Beanstalk** persistent **storage and database** locations, including: • **Amazon S3** (application files, server log files). • **Amazon DynamoDB**. • **Amazon RDS**. • Any other third-party application or database running on **Amazon EC2**.
Amazon EKS on AWS Outposts ✓ / ✓	Commvault protects **Amazon EKS** containerized applications and persistent data (containers, persistent volumes, secrets) running on AWS Outposts. • **Protecting EKS Data in AWS Outposts**. • **Migrating EKS Applications Between AWS Outposts, AWS Regions, and On-Premises**.

Contact Center

Amazon Connect ✓ / ✓	Commvault protects **Amazon Connect** call recordings of customer interactions and scheduled metrics reports which are stored in Amazon S3. • **Amazon S3** Commvault can also protect Amazon Connect **integrated AWS services**, such as: • **AWS Lambda** connected persistent storage locations. • **Amazon Kinesis Data Firehose** destinations, such as **Amazon S3** and **Amazon Redshift**. • Conversation transcripts created by **Amazon Lex** are stored in **Amazon S3**. • **Amazon CloudWatch** events are stored in **Amazon S3**. • **Synthesized text-to-speed** generated by **Amazon Polly**, stored in **Amazon S3**.

Containers

Container orchestration

Amazon Elastic Container Service (ECS) ✓ / ✓	Commvault protects the **persistent storage locations** that are used by your **Amazon ECS** and **Amazon ECS Anywhere** containerized applications. Depending on are application architecture this may be one or more of the following: • **Amazon EFS**. • **Amazon FSx for Windows File Server**. • **Amazon S3**. • **Amazon EBS**. See **Choosing the right store type for your containers** for more information.

Amazon Elastic Kubernetes Service (EKS) ✓ / ✓	Commvault protects **Amazon EKS** containerized applications and persistent data (pods, deployments, statefulsets, daemonsets, persistent volumes, secrets, crds) • **Auto-protecting Amazon EKS containers by label selector** • **Amazon EKS application and data migration (cross-cluster, cross-region)**

Compute options	
AWS Fargate ✓ / ✓	Commvault protects the **persistent storage locations** that are used by your **AWS Fargate** containerized applications and tasks. Depending on your application architecture this may be one or more of the following: • **Amazon EFS**. • **Amazon FSx for Windows File Server**. • **Amazon S3**. • **Amazon EBS**. See **Choosing the right store type for your containers** for more information.

Tools & services with containers support	
AWS Copilot - / -	Commvault protects persistent storage locations for **Amazon ECS** and **AWS Fargate** containerized apps deployed with **AWS Copilot**.
Amazon Elastic Container Registry (ECR) ✓ / ✓	Commvault protects container images stored in **Amazon ECR** when a container image pull-through cache is configured and running on your Commvault-protected Kubernetes cluster. Commvault will protect your pull-through cache as a cloud-native Kubernetes application: • **Kubernetes protection**. For more information on configuring a pull-through cache on your Kubernetes cluster, see **Registry as a pull-through cache**.
AWS App2Container ✓ / -	Commvault protects modern Java (JBoss, Apache Tomcat, Spring Boot, IBM WebSphere, Oracle WebLogic) and .NET/ASP.NET web applications that are migrated to **Amazon EKS** using **AWS App2Container**. Commvault protects application manifests and persistent storage (if present). See Amazon Elastic Container Service (above) for persistent storage locations protected with applications migrated to ECS. See **Automate AWS App2Container workflow using Ansible** for more information.

On-premises	
Amazon ECS Anywhere ✓ / ✓	Commvault protects the **persistent storage locations** that are used by your **Amazon ECS** and **Amazon ECS Anywhere** containerized applications. Depending on are application architecture this may be one or more of the following: • **Amazon EFS**. • **Amazon FSx for Windows File Server**. • **Amazon S3**. • **Amazon EBS**. See **Choosing the right store type for your containers** for more information.
Amazon EKS Anywhere ✓ / ✓	Commvault protects **Amazon EKS Anywhere** containerized applications and persistent data (containers, persistent volumes, secrets) • **Amazon EKS, Amazon EKS Anywhere**. • **Amazon EKS on AWS Outposts**.

Enterprise-scale container management	
Red Hat OpenShift Service on AWS (ROSA) - / ✓	Commvault protects containerized applications and persistent storage residing on **Red Hat OpenShift Service on AWS (ROSA)** clusters. Additionally, Commvault can migrate edge and hybrid Red Hat OpenShift containerized applications to the AWS region and vice-versa. • **Red Hat OpenShift Service on AWS (ROSA)** (Pods, Deployments, Statefulsets, Daemonsets, Persistent volumes, Secrets, CRDs). • **Application and data migration (cross-cluster, cross-region)**.

Open-source	
Amazon EKS Distro (EKS-D) - / -	Commvault protects containerized applications and persistent storage residing on **Amazon EKS Distro** clusters. Additionally, Commvault can migrate on-premises Amazon EKS-D containerized applications to the AWS region and vice-versa. • **Amazon EKS Distro (EKS-D)** (Pods, Deployments, Statefulsets, Daemonsets, Persistent volumes, Secrets, CRDs). • **Application and data migration (cross-cluster, cross-region)**.

Database

Relational	
Amazon Aurora ✓ / ✓	Commvault protects **Amazon Aurora** databases (MySQL, PostgreSQL) by creating, copying, sharing, replicating, and recovering serverless (v1 and v2) and provisioned Aurora snapshots cross-region and cross-account. Commvault also provides full database-native exports for long-term retention recovery needs. • **Amazon RDS (snapshot backup)**. • **Amazon RDS (export backup)**. • **Copying Amazon RDS snapshots across AWS accounts**. • **Copying Amazon RDS snapshots across regions**.
Amazon Relational Database Service (RDS) ✓ / ✓	Commvault protects **Amazon RDS** databases (including Amazon Aurora, RDS for MySQL, RDS for PostgreSQL, RDS for MariaDB, RDS for Oracle, RDS for SQL Server, and Amazon RDS Custom) by creating, copying, sharing, replicating, and recovering RDS snapshots cross-region and cross-account. Commvault also provides full database-native exports for long-term retention recovery needs. • **Amazon RDS (snapshot backup)**. • **Amazon RDS (export backup)**. • **Copying Amazon RDS snapshots across AWS accounts**. • **Copying Amazon RDS snapshots across regions**.
Amazon RDS Custom ✓ / ✓	Commvault protects **Amazon RDS Custom** instances (Amazon RDS Custom for Oracle) using database-native application agents installed on Amazon RDS Custom instances. See **Amazon RDS** (above) for additional details. Commvault provides deep application-integrated protection using Commvault Oracle in-guest agents, which provides the following additional backup granularity from snapshot-based RDS protection: • **Oracle** database files (data files, tablespaces), archive logs, and control files. Agent-based protection includes advanced protection options like differential backup and transaction log backups for granular control of protection and recovery scenarios.
Amazon RDS on Outposts ✓ / ✓	Commvault protects **Amazon RDS on Outposts** instances (including MySQL, PostgreSQL, and SQL Server) using multi-region, multi-account cloud-native snapshot management. Commvault also provides full database exports for long-term retention recovery needs. • **Protecting RDS Data in AWS Outposts**. • **Replicating Data Between AWS Outposts, AWS Regions, and On-Premises**. • **Migrate RDS Instances From AWS Region to Outposts**.

Amazon RDS on VMware	Commvault protects **Amazon RDS on VMware** instances (including MySQL, and PostgreSQL) using multi-region, multi-account cloud-native snapshot management. Commvault also provides full database exports for long-term retention recovery needs.
✓ / ✓	See **Amazon RDS** (above) for additional details.

Key-value

Amazon DynamoDB	Commvault protects **Amazon DynamoDB** tables using DynamoDB data access APIs to take full and incremental backups across regions and accounts. Recovery may occur across regions, accounts, and new table names for dev/test seeding activities.
✓ / ✓	• **Amazon DynamoDB**.

In-memory

Amazon ElastiCache	Commvault protects **Amazon ElastiCache for Redis** in-memory cache database backups that are exported to Amazon S3.
✓ / ✓	• **Amazon S3**. **Amazon ElastiCache for Memcached** does not support backup & recovery.
Amazon MemoryDB for Redis	**Amazon MemoryDB for Redis** provides automated and manual backup capability which can be exported to Amazon S3 for use outside of the Amazon ElastiCache service
✓ / ✓	Commvault can protect the **Amazon S3 export** location.

Document

Amazon DocumentDB (with MongoDB compatibility)	Commvault protects fully managed NoSQL **Amazon DocumentDB** cluster groups using multi-region, multi-account cloud native snapshot management. Commvault auto-discovers and protects DocumentDB instances by tag and optionally copies snapshots to another region or AWS account.
✓ / ✓	• **Amazon DocumentDB (with MongoDB compatibility)**.

Ledger

Amazon Ledger Database Services (QLDB)	Commvault protects Amazon Quantum Ledger Database (QLDB) System-of-record implementations that **export the contents of the QLDB journal** to Amazon S3 or stream journal data into downstream systems such as:
✓ / ✓	• **Amazon Redshift**. • **Amazon S3**.

Migrations	

AWS Database Migration Service (AWS DMS)

✓ / ✓

Commvault protects all supported **AWS DMS** database **sources** and **targets** (destinations), including

Sources

- **Oracle**.
- **Microsoft SQL Server**.
- **MySQL**.
- **MariaDB**.
- **PostgreSQL**.
- **MongoDB**.
- **SAP**.
- **IBM DB2**.
- **Microsoft Azure SQL Database**.
- **Google Cloud for MySQL**.
- **Amazon RDS instances**.
- **Amazon S3**.

Destinations

- On-premises and Amazon EC2 instance databases.
- Amazon EDS instance databases.
- Amazon Aurora Serverless.
- Amazon Redshift.
- Amazon S3.
- Amazon DynamoDB.
- Amazon OpenSearch Service.
- Amazon ElastiCache for Redis.
- Amazon Kinesis Data Streams.
- Amazon DocumentDB (with MongoDB compatibility).
- Amazon Managed Streaming for Apache Kafka.

Developer Tools

DevOps and Automation	

AWS Cloud9

✓ / -

Commvault protects **Amazon Cloud9** IDE environments using Amazon EC2 backup and recovery. The protection uses Amazon EBS snapshots and optional Commvault streamed backups.

See **Amazon EC2**.

AWS CodeBuild

✓ / ✓

Commvault protects **AWS CodeBuild** artifacts from built code stored in Amazon S3 buckets.

- **Amazon S3**.

AWS CodeCommit

✓ / ✓

Commvault protects **AWS CodeCommit** Git-compatible repositories if synchronized with a **GitHub** or **Azure DevOps**-compatible location (i.e., **A Serverless Solution to Keeping Git Repositories Synchronized**).

AWS CodeDeploy

✓ / ✓

Commvault protects **AWS CodeDeploy** deployment configurations by protecting the deployment targets like Amazon EC2 instances or edge-based virtual machines running the CodeDeploy agent.

Commvault does not protect, recreate, or recover deployment configurations.

AWS CodePipeline	Commvault protects **AWS CodePipeline** continuous delivery pipeline **sources** that are stored in Amazon S3.
✓ / ✓	• **Amazon S3**.

AWS CodeStar	Commvault protects **AWS CodeStar** resources used to develop, build, and deploy your CodeStar project. Some example resources protected natively by Commvault, and deployed and used by CodeStar templates include:
✓ / ✓	• **Amazon EC2**. • **Amazon S3**.

IDE and IDE Toolkits

Azure DevOps	Commvault protects Azure DevOps instances with native Azure DevOps integration.
- / -	• **Azure DevOps**.

Monitoring and Tracing

AWS X-Ray	Commvault protects **Amazon X-Ray**-related or **integrated services** for applications being analyzed, including:
✓ / ✓	• AWS Lambda persistent storage locations (see **Compute**). • AWS Elastic Beanstalk persistent storage locations (see **Compute**). • Amazon EventBridge events are stored in **Amazon S3**.

End User Computing

Cloud-native persistent desktops

Amazon WorkSpaces Family	Commvault protects cloud-native Amazon Linux and Microsoft Windows 10 **Amazon WorkSpaces Family** persistent desktops using in-guest agents that protect data and monitor and alert on malicious or anomalous activity that may indicate ransomware or malware infection.
✓ / ✓	• **Amazon WorkSpaces Family**.

Cloud-native non-persistent desktops and applications

Amazon AppStream 2.0	Commvault protects cloud-native Amazon Linux 2 and Microsoft Windows 10 **Amazon AppStream 2.0** application and Windows settings (users' plugins, toolbar settings, browser favorites, application connection profiles, and other settings) stored in Amazon S3.
✓ / ✓	• **Amazon S3**.

Cloud-native secure web access	

Amazon WorkSpaces Web
 	
 ✓ / ✓
	Commvault protects persistent storage locations accessible by **Amazon WorkSpaces Web** instances, such as:
	
	 • **Amazon S3**. • **Microsoft Office 365**.
	                                                    • **Microsoft Dynamics 365**.
	                                                    • **Salesforce**.
	                                                    • **Azure DevOps**.
	                                                    • **Github**.

Front-End Web & Mobile

Lifecycle - Develop	
AWS Amplify ✓ / ✓	Commvault protects underlying cloud services that **AWS Amplify** libraries utilize to store and retrieve data, including: • AWS Lambda persistent storage locations (see **Compute**). • **Amazon S3**. • Amazon Lex transcripts are stored in **Amazon S3**.
Lifecycle – Test & monitor	
AWS Device Farm ✓ / -	Commvault protects **AWS Device Farm** generated **logging and monitoring** artifacts, including AWS Device Farm API calls generated by AWS CloudTrail and stored in Amazon S3. • **Amazon S3**

Games

Lifecycle – Engage	
Amazon GameLift ✓ / ✓	Commvault protects **Amazon Gamelift** FleetIQ compute instances using native Amazon EC2 and Amazon EBS snapshots and network streamed backups. See **Amazon EC2**.
Amazon GameKit ✓ / ✓	Commvault protects **Amazon Gamekit** resources deployed with the AWS Gamekit plugin, which deploys the required infrastructure for your game. Some examples of resources that Commvault protects include: • **Amazon EC2** instances. • **Amazon EKS** applications. • **Amazon S3** storage.

Amazon GameSparks	Commvault protects game backends developed with **Amazon GameSparks** that persist their data to only of the following persistent data stores:
✓ / ✓	• **Amazon DynamoDB**. • **AWS Lambda** persistent data store locations.

Internet of Things

Device software	
FreeRTOS ✓ / ✓	Commvault protects FreeRTOS real-time operating system-generated files that are uploaded to Amazon S3 (see **coreHTTP basic Amazon S3 upload)**. • **Amazon S3**.
AWS IoT Greengrass ✓ / ✓	Commvault protects **AWS IoT Greengrass devices** and deployments that exchange and write their data to **Amazon S3**, and **Amazon Kinesis** locations.
AWS IoT ExpressLink - / -	Commvault protects **AWS IoT ExpressLink**-connected devices by protecting the Amazon storage (**Amazon S3**), and compute (**Amazon EC2**) locations that ExpressLink-connected devices utilize to store important data.
Connectivity and control services	
AWS IoT Core ✓ / ✓	Commvault protects devices that use the **AWS IoT Core** to run remote actions on IoT devices and upload results (files) to **Amazon S3** (e.g., **Create and run the job in AWS IoT**).
AWS IoT Device Defender - / -	Commvault protects **AWS IoT Device Defender** security alerts and anomalies sent to Amazon SNS and then stored using **AWS Lambda**, Amazon SQS, or **Amazon Kinesis**. Alerts are typically stored in **Amazon S3** and protected by Commvault, providing a long-term retention audit location for the security posture of a device over its lifetime.
Analytics services	
AWS IoT SiteWise	Commvault protects **AWS IoT SiteWise** analytics data from industrial equipment by protecting **exported AWS IoT SiteWise data** stored in a Commvault-protected **Amazon S3** bucket.

✓ / ✓	Commvault can also protect data residing on AWS IoT SiteWise gateways on-premises using Linux file-system agents or Virtual Server Agent for virtualized gateways.
AWS IoT Analytics ✓ / ✓	Commvault protects **AWS IoT Analytics** raw messages stored in JSON and Apache Parquet format in a Commvault-protected **Amazon S3** bucket.
AWS IoT TwinMaker ✓ / ✓	Commvault protects **AWS IoT TwinMaker** digital twins of real-world systems by protecting connected data stores, including video data (in Commvault-protected **Amazon Kinesis Video Streams**) and document data (in Commvault-protected **Amazon S3**).

Machine Learning

Amazon SageMaker ✓ / ✓	Commvault protects **Amazon SageMaker** machine learning (ML) models by protecting input data such as **Amazon S3** and **Amazon Redshift**. Once SageMaker has analyzed data or evaluated a machine learning model, results are written to a Commvault-protected Amazon S3 bucket.
Amazon Augmented AI (Amazon A2I) - / -	Commvault protects **Amazon Augmented AI** by protecting the **Amazon S3** bucket that contains consolidated reviews that human reviewers have reviewed, weighted, and scored for further machine-learning model training.
Amazon CodeGuru ✓ / ✓	Commvault protects **Amazon CodeGuru**-related and **integrated services**, including: • **GitHub, GitHub Enterprise Cloud, and GitHub Enterprise Server Repositories**. • **Amazon S3**.
Amazon Comprehend ✓ / ✓	Commvault protects **Amazon Comprehend** AI-driven insights from text and documents by protecting both the source content stored in **Amazon S3** (email, chat, social, phone calls), and the resulting **Amazon Redshift** data warehouse with extracted phrases, entries, and sentiment analysis.
Amazon Comprehend Medical ✓ / -	Commvault protects **Amazon Comprehend for Medical** AI-driven medical insights by protecting both the source medical information stored in **Amazon S3** (unstructured text, doctors' notes, clinical trial reports) and the resulting recommendations and relationships identified during analysis, also written to Commvault-protected Amazon S3 storage.

Amazon Forecast ✓ / ✓	Commvault protects **Amazon Forecast** forecasts by protecting **exported forecasting models** stored in CSV format in Commvault-protected **Amazon S3** buckets.
Amazon Fraud Detector ✓ / ✓	Commvault protects **Amazon Fraud Detector's** fraudulent event predictions by protecting the source event data located in a Commvault-protected **Amazon S3** bucket, and the generated batch predictions stored in Amazon S3. Storing and protecting historical event data and predictions allows future training and re-training of the machine learning (ML) models used by Amazon Fraud Detector.
Amazon HealthLake ✓ / ✓	Commvault protects **Amazon HealthLake** healthcare and life sciences patient population health data by protecting source patient medical records stored in **Amazon S3** and protecting periodic HealthLake FHIR format **exports**, also stored in Amazon S3.
Amazon Kendra ✓ / ✓	Commvault protects Amazon Kendra enterprise search instances by protecting source data stored in services like **Amazon S3**, **Amazon FSx**, **Amazon WorkDocs**, and **Amazon Redshift**. Commvault also protects non-AWS data sources like OneDrive, Salesforce, SharePoint, and GitHub. Commvault also protects search analytics metrics in CSV format when written to a Commvault-protected storage location (Amazon S3, Amazon FSx, Amazon EFS).
Amazon Lex ✓ / ✓	Commvault protects **Amazon Lex** chatbots, built with conversational AI, by protecting the **Amazon S3** bucket used to store Amazon Lex bot **exports** (bots, bot locales, customer vocabularies) or backups.
Amazon Lookout for Equipment ✓ / ✓	Commvault protects **Amazon Lookout or Equipment** industrial equipment monitoring by protecting source sensor data stored in **Amazon S3** and protecting Lookout inference results exported to Amazon S3 in **JSON format**.
Amazon Lookout for Metrics ✓ / ✓	Commvault protects **Amazon Lookout of Metrics** anomaly detection by protecting Commvault-supported data sources, such as **Amazon S3**, **Amazon RDS**, and **Amazon RedShift**. Anomaly detectors that write out to **AWS Lambda** can also protect detection events to Lambda and Commvault-supported persistent data stores.

Amazon Lookout for Vision ✓ / ✓	Commvault protects **Amazon Lookout for Vision** ML-powered defect detection by protecting the regional Lookout **console bucket** stored in **Amazon S3** (images, training results, trial detection results))
Amazon Monitron ✓ / ✓	Commvault protects **Amazon Monitron** industrial equipment monitoring data and insights by protecting **periodic exports** of Amazon Monitron data to Commvault-protected **Amazon S3**.
Amazon Personalize ✓ / ✓	Commvault protects **Amazon Personalize** ML-powered personalization recommendations by protecting the customer profile, catalog, and event data stored in **Amazon S3**. Commvault also protects batch item recommendations and identified user segments exported to Amazon S3.
Amazon Polly ✓ / ✓	Commvault protects **Amazon Polly** synthesized natural-sounding human speech files by protecting Polly-generated audio files stored in **Amazon S3** and/or **Amazon DynamoDB**.
Amazon Rekognition ✓ / ✓	Commvault protects **Amazon Rekognition** detected labels, faces, celebrities, and inappropriate content found in images and videos, and stored in **Amazon RDS** and **Amazon DynamoDB** (e.g., **Storing Amazon Rekognition Data with Amazon RDS and DynamoDB**).
Amazon Textract ✓ / ✓	Commvault protects **Amazon Textract** extracted text and structured data from PNG. JPEG, TIFF, and PDF files. Results of document, invoice, receipt, and identity document analysis are stored in Commvault-protected **Amazon S3** buckets.
Amazon Translate ✓ / ✓	Commvault protects **Amazon Translate** natural language text translations to and from English that are stored in Commvault-protected Amazon S3. Commvault protects both source (input) files and generated translations (output).
Amazon Transcribe ✓ / ✓	Commvault protects **Amazon Transcribe** transcripts of audio and video files and the resulting output JSON transcripts. Commvault protects both source (input) and generated transcripts (output) stored in Commvault-protected **Amazon S3** buckets.

AWS Deep Learning AMIs	Commvault protects, replicates, and recovers **Amazon EC2** instances deployed from **AWS Deep Learning AMIs** using native Amazon EB2 snapshots and Commvault streamed backup copies.
✓ / ✓	
AWS Deep Learning Containers	Commvault protects **Amazon EKS** container-based applications and persistent volumes deployed from **AWS Deep Learning Containers**. Commvault uses cloud-native integration with the Kubernetes API server and the CSI storage driver to protect deep learning containers and attached Amazon EBS volumes.
✓ / ✓	

Management & Governance

Management & Governance - Enable built-in governance	
AWS Control Tower	Commvault protects **AWS Control Tower** configuration and customizations stored within Commvault-protected **Amazon S3** buckets. AWS Control Tower provides a simplified method for setting up a multi-account landing zone (MALZ) and security best practices.
✓ / ✓	
AWS Organizations	Commvault protects **AWS Organizations**' centralized AWS account governance by protecting periodic **exports of all AWS accounts** in your organization to a Commvault-protected **Amazon S3** bucket.
✓ / ✓	Commvault does create or modify AWS Organizations managed accounts, organizations, or policies.
AWS Well-Architected Tool	Commvault protects **AWS Well-Architected Tool** workload reports and milestone reports to provide long-term and compliance retention copies of the application's architectural state. **Workload reports** and **milestone reports** are exported and saved to a Commvault-protected **Amazon S3** bucket or supported AWS storage service.
✓ / -	
AWS License Manager	Commvault protects **AWS License Manager** software license governance and usage reports by protecting **Usage reports** that are stored in Commvault-protected **Amazon S3**.
✓ / -	
Management & Governance - Provision resources	
AWS CloudFormation	Commvault protects **AWS CloudFormation** Infrastructure as Code (IaC) provisioning templates stored in Commvault-protected **Amazon S3** buckets. AWS CloudFormation templates are generated using **AWS CloudFormation Designer** or equivalent tools and saved to an Amazon S3 bucket to perform stack updates.
✓ / ✓	Commvault does not save, create, or delete Stacks, StackSets, Parameters, or Exports.

AWS Service Catalog ✓ / ✓	Commvault protects **AWS Service Catalog** curated IaC templates or product definitions stored in Commvault-protected **Amazon S3** with bucket prefix `cf-templates-`. As new versions of your products are developed, Commvault protects previous AWS CloudFormation product templates.
AWS OpsWorks ✓ / ✓	Commvault protects **AWS OpsWorks**-managed instances of Chef and Puppet by protecting your configuration management automation scripts for **AWS Opsworks for Chef Automate**, **AWS OpsWorks for Puppet Enterprise**, and **AWS OpsWorks Stacks.** Commvault protects *AWS OpsWorks for Chef Automate Server* **backups** and *OpsWorks for Puppet Enterprise Server* **backups** located in **Amazon S3**. Commvault protects *AWS OpsWorks Stacks* **archives** located in Amazon S3.

Management & Governance - Operate with speed

Amazon CloudWatch ✓ / ✓	Commvault protects **Amazon CloudWatch** monitoring metrics, alarms, and statistics for your AWS and edge-based workloads by protecting **Amazon CloudWatch Logs data exports**, Log insights, and **CloudWatch Events log files** stored in **Amazon S3**.
Amazon Managed Grafana ✓ / ✓	Commvault protects **Amazon Managed Grafana** dashboard definitions by protecting periodic Grafana dashboard **exports** located in Commvault-protected **Amazon S3** or supported AWS storage solutions. Commvault does not create, modify, or delete Amazon Managed Grafana workspaces.
Amazon CloudTrail ✓ / ✓	Commvault protects **Amazon CloudTrail** user and API activity logging by protecting **CloudTrail logs** and AWS CloudTrail Lake **query results** written to Commvault-protected **Amazon S3** buckets.
AWS Config ✓ / ✓	Commvault protects **AWS Config** configurations and relationships by protecting **Configuration Snapshots** stored in Commvault-protected **Amazon S3** buckets. Protecting configuration snapshots provide a provable, protected audit history of managed resources.
AWS Systems Manager ✓ / ✓	Commvault protects **AWS Systems Manager** operational insights and automation by protecting **OpsData Explorer exports**, **Patch compliance** reports, and **automation scripts** stored in Commvault-protected **Amazon S3**.

AWS Managed Services (AMS)	Commvault protects resources deployed and managed in **AWS Managed Services** environments. Commvault protects resources in *AWS Managed Services Accelerate* and *AWS Managed Services Advanced* environments using a mixture of AWS native snapshots and agent-based protection allowing tight integration with AMS infrastructure automation solutions. Commvault protects a broad range of AWS **Compute**, **Containers**, **Databases**, and **Storage** services that AMS supports. Commvault also protects **AMS aggregated service** logs located in Commvault-protected **Amazon S3**.
AWS Proton	Commvault protects containerized applications deployed by **AWS Proton** to **Amazon EKS**, **Amazon EKS Anywhere**, and **Amazon EKS Distro** (EKS-D). Additionally, Commvault protects persistent storage locations for containerized applications deployed by AWS Proton to **Amazon ECS**, **Amazon ECS Anywhere**, and **AWS Fargate**.
AWS DevOps Guru	Commvault protects **AWS DevOps Guru** insights and anomaly detections by protecting insights delivered to Amazon SNS topics and then stored in Commvault-protected **Amazon S3** or **Amazon Redshift**.

Media Services

Media Services - Content Creation
Amazon Nimble Studio ✓ / ✓	Commvault protects **Amazon Nimble Studio** creative studio resources including **Amazon EC2** instances, **Amazon FSx** file-systems, and **Amazon EFS** Linux home directories used by Nimble Studio.
AWS Thinkbox Deadline - / -	Commvault protects **AWS Thinkbox Deadline** management software for render farms by protecting on-premises physical, virtualized or Amazon EC2 Thinkbox compute instances. Additionally, Thinkbox uses **MongoDB** for its database which is also natively protected by Commvault.
AWS Thinkbox Frost - / -	Commvault protects **AWS Thinkbox Frost** particle mesh rendering solutions by protecting on-premises physical, virtualized or Amazon EC2 Frost compute render instances.
AWS Thinkbox Krakatoa - / -	Commvault protects **Amazon Thinkbox Krakatoa** volumetrics rendering solutions by protecting on-premises physical, virtualized or Amazon EC2 Krakatoa compute render instances.

AWS Thinkbox Sequoia - / -	Commvault protects **Amazon Thinkbox Sequoia** point cloud processing and meshing solutions by protecting on-premises physical, virtualized or Amazon EC2 Sequoia compute render instances.
AWS Thinkbox Stoke - / -	Commvault protects **Amazon Thinkbox Stoke** particle simulations by protecting on-premises physical, virtualized or Amazon EC2 Stoke compute render instances.
AWS Thinkbox XMesh - / -	Commvault protects **Amazon Thinkbox Xmesh** animated scene generation solutions by protecting on-premises physical, virtualized or Amazon EC2 XMesh compute render instances.

Media Services - Content Distribution

AWS Elemental MediaConvert ✓ / ✓	Commvault protects **AWS Elemental MediaConvert** video transcoding solutions by protecting input files, multi-format output files, and job exports stored in Commvault-protected **Amazon S3** buckets.
AWS Elemental MediaLive ✓ / ✓	Commvault protects **AWS Elemental MediaLive** live broadcast video processing solutions by protecting input files, HLS output group, and archive or frame capture output files stored in Commvault-protected **Amazon S3** buckets.
AWS Elemental MediaPackage ✓ / ✓	Commvault protects **AWS Elemental MediaPackage** video packaging and distribution output by protecting input Video on Demand (VOD) assets and MediaPackage packaged content stored in Commvault-protected **Amazon S3** buckets.
AWS Elemental MediaTailor ✓ / ✓	Commvault protects **AWS Elemental MediaTailor** over-the-top (OTT) ad insertion enhanced video by protecting input source files, bumper files, and output files stored in Commvault-protected **Amazon S3**.

Amazon Interactive Video Service (IVS)	Commvault protects **Amazon Interactive Video Service (IVS)** live and interactive video streams by protecting IVS **auto-recorded** channels stored in Commvault-protected **Amazon S3**.
✓ / ✓	

Media Services - AWS Elemental Appliances and Software	
AWS Elemental Live - / -	Commvault protects edge-based **AWS Elemental Live** physical and virtual live video processing appliances using agents installed on qualified hardware, or virtualization crash-consistent snapshots. Commvault also protects input and output files stored in Commvault-protected **Amazon S3**.
AWS Elemental Conductor - / -	Commvault protects edge-based **AWS Elemental Conductor** management software using agents installed on qualified hardware, or virtualization crash-consistent snapshots.
AWS Elemental Statmux - / -	Commvault protects edge-based **AWS Elemental Statmux** physical and virtual live video processing appliances using agents installed on qualified hardware, or virtualization crash-consistent snapshots. Commvault also protects input and output files stored in Commvault-protected **Amazon S3**.

Migration & Transfer

Migration & Transfer - Assess and mobilize	
Migration Evaluator - / -	Commvault protects **Migration Evaluator** backups and nightly synchronization to Amazon S3 through Commvault-protected virtualization machine (VM) backup, file-system backup, and **Amazon S3** protection.
AWS Application Discovery Service - / -	Commvault protects **AWS Application Discovery Service** on-premises service inventories exported in CSV format. Commvault protects the Application Discovery collector VM, file-system, and collected data relating to servers and VMs may be **exported** to **Amazon S3** or an equivalent Commvault-protected storage service.

Migration & Transfer - Migrate your applications	
AWS Database Migration Service ✓ / ✓	Commvault protects AWS Data Migration Service databases in their source location using agent-based database integrated protection and then protects resulting databases located on **Amazon EC2**, **Amazon RDS**, **Amazon Redshift**, **Amazon DynamoDB**, and **Amazon S3** exports.

AWS Mainframe Modernization	Commvault protects **AWS Mainframe Modernization** workloads by protecting **Amazon EC2** runtime environments and **Amazon EFS** and **Amazon FSx** file-systems used by re-factored applications.
✓ / ✓	

Migration & Transfer - Migrate your storage

AWS DataSync	Commvault protects **AWS DataSync** source and destination locations, including NFS, SMB, HDFS, Object storage, Amazon EFS, Amazon FSx for Windows, Amazon FSx for NetApp ONTAP, and Amazon S3.
✓ / ✓	

AWS Snow family	Commvault supports **AWS Snow family** storage transfer devices for offline migration of Commvault backup and archival data from on-premises to Amazon S3. Commvault supports writing data to Snow devices as **Cloud libraries** (S3) or **Disk libraries** (NFS).
✓ / ✓	

Satellite

AWS Ground Station	Commvault protects **AWS Ground Station** satellite communications and data **delivered** asynchronously to Commvault-protected **Amazon S3** buckets. Alternatively, AWS Ground Station contact data may be synchronously delivered to Commvault-protected **Amazon EC2** instances.
✓ / ✓	

Networking & Content Delivery

Networking & Content Delivery – Network foundations

Amazon VPC	Commvault protects **Amazon VPC** resource configuration related to your protected **Amazon EC2** instances. As part of Amazon EC2 recovery, Commvault will detect and recreate missing Amazon VPCs, Subnets, Elastic Network Interfaces (ENIs), and Security groups. Additionally, Amazon VPC and related resource configurations may be restored as JSON for forensic investigations.
✓ / ✓	

AWS Transit Gateway	Commvault protects **AWS Transit Gateway** flow logs that are **published** to Commvault-protected **Amazon S3**.
✓ / ✓	

Networking & Content Delivery - Application networking

| AWS App Mesh | Commvault protects cloud-native containerized applications residing on Amazon EKS, Amazon EKS Anywhere, and Amazon EKS Distro Kubernetes clusters that are part of **AWS App Mesh** service meshes.

Commvault does not protect or recreate AWS App Mesh resources. |
| --- | --- |

AWS Cloud Map	Commvault protects AWS services that can have service discovery provided by **AWS Cloud Map**. Commvault protects the following AWS services that receive simplified service naming via AWS Cloud Map:

- **Amazon EC2** instances.
- Amazon ECS and **Amazon EKS** applications and services.
- **Amazon S3** buckets.
- **Amazon DynamoDB** tables.

Commvault does not protect or recreate AWS Cloud Map services or namespaces.

Networking & Content Delivery – Edge networking

AWS Global Accelerator	Commvault protects **AWS Global Accelerator** flow logs published to Commvault-protected **Amazon S3** buckets.

✓ / ✓

Quantum Technologies

Amazon Braket	Commvault protects **Amazon Braket** quantum computing-enabled research, circuit simulations, and proof of concept applications by protecting the **Amazon S3** output location of Braket jobs.

✓ / ✓

Robotics

AWS RoboMaker	Commvault protects **Amazon Robomaker** robot and robotics fleet simulation applications which write simulation results to Commvault-protected **Amazon S3** buckets. Additionally, Commvault protects **Amazon EC2** Cloud9 IDE instances used during simulation creation.

✓ / ✓

Security, Identity & Compliance

Identity & access management

AWS Identity & Access Management (IAM)	Commvault requires an AWS IAM user or role to perform protection of AWS resources and supports a **least privilege** approach to refining permissions to only the AWS products that will be protected by Commvault software. Commvault supports **AWS Secure Token Service (STS)** to obtain temporary credentials during protection activities.

Commvault does not recover users, groups, roles, or policies definitions.

✓ / ✓

AWS Directory Service	Commvault protects **AWS Directory Service** (AWS Managed Microsoft AD) instances by creating an Amazon EC2-based Replica Domain Controller and installing the **Commvault Active Directory agent** to protect the AWS Directory Service objects and attributes.
✓ / ✓	

Detection	

AWS Security Hub	Commvault protects **AWS Security Hub** security, risk management, and compliance findings when exported to a Commvault-protected **Amazon S3** bucket. See **How to export AWS Security Hub findings to CSV format** for details on the solution.
✓ / ✓	

Amazon GuardDuty	Commvault protects **Amazon GuardDuty**'s continuous security monitoring findings when **exported** to a Commvault-protected **Amazon S3** bucket for historical auditing purposes. Commvault does not create, modify, or restore detectors, filters, members, destinations, or threat intel sets.
✓ / ✓	

Amazon Inspector	Commvault protects **Amazon Inspector**'s continual vulnerability scanning and reporting services by protecting **exported findings** (CSV, JSON), stored in Commvault-protected **Amazon S3** buckets.
✓ / ✓	

Network and application protection	

AWS Network Firewall	Commvault protects **AWS Network Firewall** services that provide stateful firewall and intrusion detection protection to Amazon VPC workloads. Commvault protects Network Firewall flow logs delivered to Commvault-protected **Amazon S3** buckets.
✓ / ✓	Commvault does not create or modify firewalls, policies, or rule groups.

AWS Web Application Firewall (WAF)	Commvault protects **AWS Web Application Firewalls** by protecting **web ACL traffic** written to Commvault-protected **Amazon S3** buckets for historical auditing. Commvault does not create or modify web ACLs, rule groups, rules, and managed protections.
✓ / ✓	

Data Protection	

Amazon Macie	Commvault protects **Amazon Macie**'s sensitive data discovery and reporting by protecting source data locations, and the sensitive data discovery job results stored in Commvault-protected **Amazon S3** buckets.
✓ / ✓	

AWS CloudHSM	Commvault protects **Amazon CloudHSM** single-tenant hardware security modules (HSMs) on AWS by protecting **HSM backups** stored in Commvault-protected **Amazon S3** buckets.

✓ / ✓

Compliance	
AWS Audit Manager	Commvault protects **AWS Audit Manager** risk and compliance assessment and evidence collection services by protecting **assessment reports** and **manual evidence uploads** published to Commvault-protected **Amazon S3** buckets.

✓ / ✓

Serverless

Serverless	
Serverless	Commvault protects key Serverless building blocks, including: • **AWS Lambda** persistent data stores. • **AWS Fargate** persistent data stores. • **Amazon S3**. • **Amazon DynamoDB**. • **Amazon Aurora Serverless**.

✓ / ✓

Storage

Object, file, and block storage	
Amazon Simple Storage Service (Amazon S3) ✓ / ✓	Commvault protects, recovers, and replicates Amazon S3 objects including metadata, ACLs, and tags within and across regions and accounts. Commvault protects objects residing in all Amazon S3 storage classes, including Amazon S3 Object Lock. • **Amazon S3 backup**. • **Replication for Amazon S3**. Commvault supports the use of Amazon S3 (including **S3 Object Lock**) for backup and archival data in deduplicated, compressed, and encrypted format referred to as **Cloud Storage**.
Amazon S3 on Outposts ✓ / ✓	Commvault protects, recovers, and replicates **Amazon S3 on Outposts** objects including metadata, ACLs, and tags within and across regions and accounts. Commvault protects objects residing in all Amazon S3 on Outposts storage classes, including Amazon S3 Object Lock. • **Amazon S3 on AWS Outposts backup**. • **Replication for Amazon S3 on Outposts**. Commvault supports the use of Amazon S3 on Outposts for backup and archival data in deduplicated, compressed, and encrypted format referred to as **Cloud Storage**.

Amazon S3 Glacier ✓ / ✓	Commvault protects, recovers, and replicates Amazon S3 Glacier objects including metadata, ACLs, and tags within and across regions and accounts. Commvault protects objects residing in all Amazon S3 Glacier storage classes. • . **Amazon S3 Glacier backup**. Commvault supports the use of Amazon S3 Glacier (including **S3 Glacier Vault Lock**, and **S3 Object Lock**) for infrequent-use backup and archival data in a deduplicated, compressed, and encrypted format referred to as **Cloud Storage**. Commvault recommends using **Commvault Combined Storage Tier** to streamline data recall from **Amazon S3 Glacier Flexible Archive** and **Amazon S3 Glacier Deep Archive** storage classes. Do <u>not</u> write backups directly to S3 Glacier Flexible Archive / Deep Archive. Commvault supports S3 Batch Operations for faster restores from S3 Glacier Flexible Retrieval, S3 Glacier Deep Archive, and S3 Intelligent-Tiering Archive Access and Deep Archive Access Tiers (see **Reduce recovery time and optimize storage costs with faster restores from Amazon S3 Glacier and Commvault**).
Amazon Elastic Block Store (Amazon EBS) ✓ / ✓	Commvault protects **Amazon EBS** high-performance block volumes using EBS snapshots and **APIs** in crash-consistent and application-consistent modes. Commvault also creates long-term retention or archival copies of EBS volumes on optimized deduplicated Amazon S3 storage. • **Amazon EC2 backup** (crash-consistent). • **Amazon EBS backup** (application consistent). • **Replicating EBS Snapshots Cross-region and Cross-account)**.
Amazon Elastic File System (Amazon EFS) ✓ / ✓	Commvault protects **Amazon EFS** (all storage classes) shared NFSv4.1 and v4.0 file-systems by mounting the file-system on a Commvault Access Node and protecting files, symbolic links, and ACLs. • **Amazon EFS**
Amazon FSx for NetApp ONTAP ✓ / ✓	Commvault protects, **migrates**, and replicates **Amazon FSx for NetApp ONTAP (FSxN)** fully managed NFS and SMB file-systems using Commvault Network Attached Storage (NAS) protection. • **NAS File Server backup**. • **Replication for File System Agents** (Disaster Recovery). • **Configuring NetApp SnapLock Compliance and SnapLock Enterprise on Amazon FSx for NetApp ONTAP disk libraries**.
Amazon FSx for OpenZFS	Commvault protects, **migrates**, and replicates **Amazon FSx for OpenZFS** fully managed shared storage accessible by NFS protocol, using Commvault Network Attached Storage (NAS) protection. • **NAS File Server backup**.

✓ / ✓	• **Replication for File System Agents** (Disaster Recovery).
Amazon FSx for Windows File Server FSX ✓ / ✓	Commvault protects, **migrates**, and replicates **Amazon FSx for Windows File Server** fully managed file storage accessible by SMB protocol, using Commvault Network Attached Storage (NAS) protection (including ACLs). • **Amazon FSx for Windows File Server**. • **Replication for File System Agents** (Disaster Recovery).
Amazon FSx for Lustre FSX ✓ / ✓	Commvault protects and replicates **Amazon FSx for Lustre** fully managed shared storage built on Lustre's high-performance file system, by mounting the file-system on a Commvault Access Node and protecting files, symbolic links, and ACLs. • **Amazon FSx for Lustre** (Linux file-system backup). • **Replication for File System Agents** (Disaster Recovery).

Data Migration

AWS Snow Family - / -	Commvault supports **AWS Snow family** storage transfer devices for offline data transfer and migration of Commvault backup and archival data from on-premises to Amazon S3. Commvault supports writing data to Snow devices as **Cloud libraries** (S3) or **Disk libraries** (NFS). • **Migrating Data to Amazon S3 Using Snow Family Devices**. • **Seeding a Cloud Storage Library**. Commvault writes data to the Snow device in compressed, deduplicated, and encrypted format meaning more data can be securely migrated with a single Snowball device.

Hybrid cloud storage and edge computing

Amazon Storage Gateway (File Gateway) ✓ / ✓	Commvault protects **Amazon FSx File Gateway** file shares accessible by SMB protocol, on a VMware ESXi, Microsoft Hyper-V, Linux KVM, Amazon EC2, or hardware appliance. using Commvault Network Attached Storage (NAS) protection (including ACLs). Commvault protects your on-premises File Gateway (including local cache folders) before data is uploaded to your Amazon FSx file-share in the AWS region. • **Virtual Machine backup and recovery** • **Linux File-system backup** (hardware appliance only) ⓘ **Note**: Commvault supports the protection of Amazon FSx for Windows File Server shares directly in AWS. File Gateway protection is intended to protect data in-flight to AWS.

Amazon Storage Gateway (Tape Gateway) ✓ / ✓	Commvault supports writing on-premises backup and archive data to **Amazon Storage Gateway** (**Tape Gateway**), accessible via iSCSI Virtual Tape Library (VTL) protocol. Tape Gateway can also be combined with AWS Snowball to offline perform mass migration of physical tapes to logical or 'virtual' tape libraries in Amazon S3 with Commvault. • **Configuring AWS Tape Gateway** Commvault can protect the *Cache storage* and *Upload buffer* on the Tape Gateway using **Virtual Machine snapshot and backup**, or **agent-based file-system backup**. ⓘ **Note**: Commvault also supports writing directly to Amazon S3 from on-premises, removing the need to deploy and manage Tape Gateways or Virtual Tape Library (VTL) configurations.
Amazon Storage Gateway (Volume Gateway) ✓ / ✓	Commvault protects **Amazon FSx Volume Gateway** iSCSI volumes presented from an on-premises server, **virtual machine (VM)**, or hardware appliance. Commvault protects your *cached volumes* and *stored volumes* before data is uploaded or synchronized with Amazon S3. • **Virtual Machine backup and recovery**. • **Linux File-system backup** (servers and/or hardware appliance only). ⓘ **Note**: Commvault supports the protection of Amazon S3 directly in AWS. Volume Gateway protection is intended to protect volume data in-flight to Amazon S3.

Disaster Recovery and Backup

AWS Elastic Disaster Recovery (DRS) **Commvault Auto Recovery**	Amazon Elastic Disaster Recovery (DRS) provides fully managed Disaster Recovery for AWS cloud workloads. Commvault recommends leveraging **Commvault Auto Recovery** to perform DR replication, recovery, and automated testing for Amazon EC2 instances, applications, databases, and storage in hybrid cloud environments. Commvault flexible RPO and RTO-based solutions performing replication and recovery of the following Amazon workloads: • **Amazon EC2** instances. • **Amazon RDS Custom** databases (Oracle). • **SAP HANA running on Amazon EC2**. • **Amazon FSx for Windows Server** file-systems. • **Amazon EFS** file-systems. • **Amazon S3** object storage. • **Hadoop**.

AWS Backup	**AWS Backup** provides backup & recovery services for a subset of AWS services.
Commvault Backup & Recovery	Commvault recommends leveraging **Commvault Backup & Recovery** to perform holistic backup & recovery across your entire AWS and hybrid application landscape. Commvault performs backup & recovery services for the following Amazon services both in the region and on-premises via AWS Outposts support. In additional Commvault has the widest set of **protected technologies** outside of the AWS region.

- Amazon EC2
- Amazon EKS
- Amazon EKS Distro
- Amazon EKS Anywhere
- AWS Outposts
- Amazon S3

- Amazon RDS
- Amazon RDS Custom
- Amazon Redshift
- Amazon DocumentDB
- Amazon DynamoDB
- Amazon FSx
- Amazon EFS

See **www.commvault.com/aws** for additional details

Metallic protection of AWS Cloud Products

Commvault software is also available as a Software as a Service (SaaS) offering called **Metallic** (see **metallic.io**).

Metallic has multiple service offerings tailored to the AWS product to be protected, namely:

- **Metallic VM & Kubernetes Backup**
  - o Amazon EC2, Amazon EBS, and Amazon EKS compute workload protection.
- **Metallic Database Backup**
  - o Amazon RDS (including Aurora serverless/provisioned), Amazon Redshift, Amazon DynamoDB, and Amazon DocumentDB protection.
- **Metallic File & Object Backup**
  - o Amazon S3 protection.
  - o Amazon EBS file-system protection via an agent in guest.

Metallic File & Object Backup allows the protection of Amazon CloudWatch and Amazon CloudTrail logs, metrics, and events that are stored in Amazon S3. See **Commvault protection of AWS Cloud Products** for detailed per-service protection that utilizes Commvault/Metallic native Amazon S3 protection.

Metallic can also be used to migrate your existing Hyper-V, VMware, or Azure virtual machines into Amazon EC2 using **Cross Hypervisor Restores (VM Conversions)**.

Metallic is capable of protecting the following services at the time of writing (August 2022) via a customer-supplied **Metallic Backup Gateway** deployed within the customer-owned and operated VPC and native integration to the AWS service API/endpoint.

Compute

Instances (virtual machines)	
Amazon EC2	Metallic auto-discovers and protects Amazon EC2 instances and attached Amazon EBS volumes using native incremental Amazon EBS snapshots and optional streaming backup copy. • **Backups for Amazon EC2**. Instances may reside within any of the global AWS **regions**, including GovCloud.
Amazon EC2 Spot	Metallic auto-discovers and protects Amazon EC2 Spot instances and attached Amazon EBS volumes using native incremental Amazon EBS snapshots and optional streaming backup copy. • **Backups for Amazon EC2**. Instances may reside within any of the global AWS **regions**, including GovCloud.
Edge and hybrid	
VMware Cloud on AWS	Metallic auto-discovers and protects VMware Cloud on AWS (VMC) Virtual Machines, VM templates, VMDK files, and virtual Raw Disk Mapping (RDM) volumes using native VMware VDDK snapshots and optional streaming backup copy. • **VMware**.

Containers

Container orchestration	
Amazon Elastic Container Service (ECS)	Metallic auto-discovers and protects the **persistent storage locations** (listed below) that are used by your Amazon ECS containerized applications. • **Amazon S3**. See **Choosing the right store type for your containers** for more information.
Amazon Elastic Kubernetes Service (EKS)	Metallic auto-discovers and protects Amazon EKS containerized applications and persistent data (containers, persistent volumes, secrets) • **Kubernetes**.
Compute options	
AWS Fargate	Metallic protects the **persistent storage locations** (listed below) that are used by your Amazon Fargate containerized applications/tasks. • **Amazon S3**. See **Choosing the right store type for your containers** for more information
Tools & services with containers support	
AWS Copilot	Metallic protects persistent storage locations for Amazon ECS and AWS Fargate containerized apps deployed with AWS Copilot.
Amazon Elastic Container Registry (ECR)	Metallic protects container images stored in Amazon ECR when a container image pull-through cache is configured and running on your Kubernetes cluster. Metallic will protect your pull-through cache as a cloud-native Kubernetes application: • **Kubernetes**. For more information on configuring a pull-through cache on your Kubernetes cluster, see **Registry as a pull-through cache**.
AWS App2Container	Metallic auto-discovers and protects modern Java (JBoss, Apache Tomcat, Spring Boot, IBM WebSphere, Oracle WebLogic) and .NET/ASP.NET web applications that are migrated to Amazon EKS using AWS App2Container. Commvault protects application manifests and persistent storage. • **Kubernetes**. See **Automate AWS App2Container workflow using Ansible** for more information.

On-premises	
Amazon ECS Anywhere	Metallic protects the **persistent storage locations** (listed below) that are used by your Amazon ECS Anywhere containerized applications.
	• **Amazon S3**.
	See **Choosing the right store type for your containers** for more information.
Amazon EKS Anywhere	Metallic auto-discovers and protects Amazon EKS Anywhere containerized applications and persistent data (containers, persistent volumes, secrets)
	• **Kubernetes**.

Enterprise-scale container management	
Red Hat OpenShift Service on AWS (ROSA)	Metallic auto-discovers and protects containerized applications and persistent storage residing on Red Hat OpenShift Service on AWS (ROSA) fully-managed clusters. Additionally, Metallic can migrate on-premises Red Hat OpenShift containerized applications to the AWS region and vice-versa.
	• **Kubernetes**.

Open-source	
Amazon EKS Distro (EKS-D)	Metallic auto-discovers and protects containerized applications and persistent storage residing on Amazon EKS Distro self-managed clusters. Additionally, Metallic can migrate Amazon EKS-D containerized applications to the AWS region and vice-versa. Consider protecting your Amazon EKS-D on-premises applications to a Commvault HyperScale X device for rapid recovery.
	• **Kubernetes**.
	• **Commvault HyperScale X**.

Database

Relational	
Amazon Aurora	Metallic auto-discovers and protects serverless (v1 and v2) and provisioned Amazon Aurora MySQL and PostgreSQL databases using multi-region, multi-account cloud-native snapshot management.
	• **Amazon RDS**.
Amazon Relational Database Service (RDS)	Metallic auto-discovers and protects Amazon RDS instances (including Amazon Aurora, RDS for MySQL, RDS for PostgreSQL, RDS for MariaDB, RDS for Oracle, and RDS for SQL Server) using multi-region, multi-account cloud-native snapshot management.
	• **Amazon RDS**.

Amazon RDS Custom	Metallic protects Amazon RDS Custom instances (Amazon RDS Custom for Oracle) using database-native application agents installed on the Amazon RDS Custom instances.
	• **Amazon RDS**.
	Metallic also provides deep application integrated protection using Metallic Oracle and SQL Server agents, which include:
	• **Oracle** database files (data files, tablespaces), archive logs, and control files.
	• **SQL Server** database files and log files.
	Agent-based protection includes advanced protection options like differential backup and transaction log backups for granular control of protection and recovery scenarios.
Amazon RDS on VMware	Metallic auto-discovers and protects Amazon RDS on VMware instances (including MySQL, PostgreSQL, and Amazon RDS Custom) using multi-region, multi-account cloud-native snapshot management.
	• **Amazon RDS**.
Amazon Redshift	Metallic auto-discovers and protects Amazon Redshift data warehouse instances using multi-region, multi-account cloud-native snapshot management.
	• **Amazon Redshift**.

Key-value

Amazon DynamoDB	Metallic auto-discovers and protects Amazon DynamoDB tables using multi-region, multi-account cloud-native snapshot management.
	• **Amazon DynamoDB**.

Document

Amazon DocumentDB (with MongoDB compatibility)	Metallic auto-discovers and protects fully managed NoSQL Amazon DocumentDB cluster groups using multi-region, multi-account cloud native snapshot management.
	• **Amazon DocumentDB (with MongoDB compatibility)**.

Ledger

Amazon Ledger Database Services (QLDB)	Metallic protects Amazon Quantum Ledger Database (QLDB) System-of-record implementations that **export the contents of the QLDB journal** to Amazon S3 or stream journal data into downstream systems such as:
	• **Amazon Redshift**.
	• **Amazon S3**.

Storage

Metallic recommends using Amazon S3 as your primary backup location (**Cloud Copy**) for the backup of AWS workloads. Amazon S3 offers the most durable, secure, and scalable storage for the rapid recovery of your AWS workloads. Metallic supports all **Amazon S3 storage classes** and recommends using **Amazon S3 Infrequent-Access (S3-IA)** as the default for *primary backup*s stored in Amazon S3.

Additionally, Metallic supports **combined storage** locations that provide optimized retrieval from archival storage classes such as Amazon S3 Glacier Flexible Retrieval and Amazon S3 Glacier Deep Archive. Metallic recommends the use of combined storage for all archive data.

Object, file, and block storage	
Amazon Simple Storage Service (Amazon S3)	Metallic protects Amazon S3 objects in a Metallic-controlled Amazon S3 backup location in encrypted, compressed, and deduplicated format. Replication of Amazon S3 backups reduces cross-region replication fees due to deduplicated data storage format. • **Amazon S3**. Metallic supports the deployment of customer-managed storage and backup gateway(s) within your Amazon VPCs to keep your AWS backups residing adjacent to your protected workloads. • **Storage and Backup Gateway – Amazon S3**.
Amazon Simple Storage Service Glacier	Metallic protects and writes long-term retention and archival data to Amazon S3 Glacier buckets. Be sure to review Amazon S3 Glacier features and pricing to ensure Glacier meets your access time and cost of retrieval expectations.
Amazon Elastic Block Store (Amazon EBS)	Metallic protects Amazon EBS volumes as part of protecting region-based Amazon EC2 instances, see **Compute**.
Hybrid cloud storage and edge computing	
Amazon Storage Gateway (File and Volume Gateway)	Metallic can protect your Amazon S3, Amazon FSx, and iSCSI Volume Storage Gateway data. Metallic protects gateway data by installing agents on the Storage Gateway Appliance to protect the cache and upload buffer files and folders. Additionally, Metallic can protect Storage Gateway files life-cycled to Amazon S3. • **File server** (Linux and Windows). • **Amazon S3**.
Amazon Storage Gateway (Tape Gateway)	Metallic can write backup and archival data to Virtual Tape Libraries (VTLs) presented via Amazon Storage Gateway (Tape Gateways) for businesses that have a requirement to continue the usage of tape-based data management. • **On-Premises Deployment of a Metallic Backup Gateway**. • **Configuring a Tape Storage**.

Managed File Transfer	
AWS Transfer Family	Metallic protects the destination for your data transferred using AWS Transfer Family SFTP, FTPS, FTP, and AS2 transfers. • **Amazon S3**

Disaster Recovery and Backup	
AWS Elastic Disaster Recovery (DRS) **Commvault Auto Recovery**	Amazon Elastic Disaster Recovery (DRS) provides fully managed Disaster Recovery for AWS cloud workloads. Commvault recommends leveraging **Commvault Auto Recovery** to perform DR replication, recovery, and automated testing for Amazon EC2 instances, applications, databases, and storage in hybrid cloud environments. Metallic and Commvault work seamlessly together from within **Commvault Command Center™**. Commvault flexible RPO and RTO-based solutions perform replication and recovery of the following Amazon workloads: • **Amazon EC2** instances. • **Amazon RDS Custom** databases (Oracle). • **SAP HANA running on Amazon EC2**. • **Amazon FSx for Windows Server** file-systems. • **Amazon EFS** file-systems. • **Amazon S3** object storage. • **Hadoop**.
AWS Backup **Commvault Backup & Recovery**	AWS Backup provides backup & recovery services for a subset of AWS services. Commvault recommends leveraging **Commvault Backup & Recovery** to perform holistic backup & recovery across your entire hybrid application landscape. Commvault performs backup & recovery services for the following Amazon services both in the region and on-premises via AWS Outposts support. In additional Commvault has the widest set of **protected technologies** outside of the AWS region. • Amazon EC2. • Amazon EKS. • Amazon EKS Distro. • Amazon EKS Anywhere. • AWS Outposts. • Amazon S3. • Amazon RDS. • Amazon RDS Custom. • Amazon Redshift. • Amazon DocumentDB. • Amazon DynamoDB. • Amazon FSx. • Amazon EFS.

Cloud Shared Responsibility

Security and Compliance

Security and Compliance of your data stored and handled in Amazon Web Services (AWS) is a shared responsibility. AWS is responsible for protecting the infrastructure that runs all AWS services. Infrastructure refers to hardware, software, and physical locations represented by regions, availability zones, and edge locations. This is often referred to as security "of the cloud."

AWS takes its security responsibility seriously with external validation to the leading industry regulations as found at **aws.amazon.com/compliance/programs.**

Customers are responsible for protecting their data stored in AWS services by configuring and hardening the platform, application, and identity and access management controls that govern access to AWS services. This extends to operating system configuration, network access control lists (ACLs), and firewall or security group configuration. Handling data both in-transit ("on the wire") or at-rest ("on disk") with additional protections like encryption is also the responsibility of the customer.

This document is intended to inform and instruct Commvault customers on how to comply with their **Security 'IN' the Cloud** responsibility as detailed by the AWS Shared Responsibility Model **aws.amazon.com/compliance/shared-responsibility-model**.
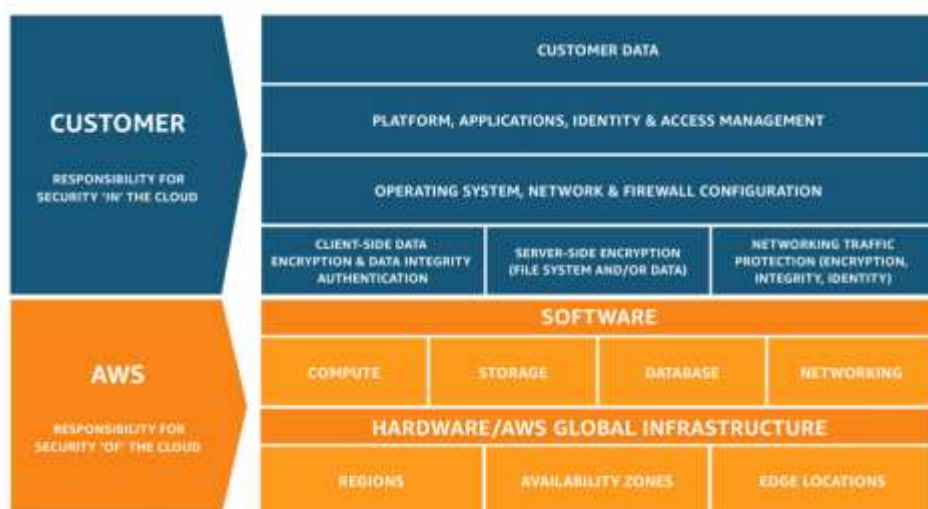
This document will help **you** secure your Commvault and AWS infrastructure and provide *Recovery Readiness* for your AWS-hosted applications and **data**. Applying Commvault to your AWS data landscape you will gain recoverability for the following AWS services both within and across the region.

- Compute:
    - Amazon Elastic Compute 2 (EC2) instances.
    - Amazon EC2 Spot instances.
    - Amazon Elastic Kubernetes Service (EKS) applications.
    - AWS Outposts – EC2 instances, EKS applications.
    - Amazon Local Zones – EC2 instances, EKS applications.
    - Red Hat OpenShift on AWS (ROSA) applications.
    - Amazon Elastic Kubernetes Distribution (EKS-D) applications.
    - VMware Cloud on AWS instances.

- o VMware Cloud on Outposts.
- Storage
  - o Amazon Elastic Block Store (EBS) volumes.
  - o AWS Outposts – EBS volumes, S3 buckets.
  - o Amazon Local Zones – EBS volumes.
  - o Amazon Elastic File System (EFS) file-systems.
  - o Amazon Elastic FSx for Windows file-systems.
  - o Amazon FSx for Lustre file-systems.
  - o Amazon Simple Storage Service (S3) – as a source, as target (all storage classes).
  - o Amazon S3 Glacier (as target).
  - o AWS Storage Gateway (Tape Gateway) (as target)
  - o AWS Storage Gateway (File Gateway, Volume Gateway) (as source).
  - o AWS Snowball Edge, Snowmobile, and Snowcone (as migration devices).
- Database
  - o Amazon Aurora – MySQL, PostgreSQL (serverless, provisioned).
  - o Amazon RDS – Oracle, MariaDB, MySQL, PostgreSQL, Microsoft SQL Server.
  - o Amazon RDS Custom – Oracle, Microsoft SQL Server.
  - o Amazon RDS on VMware - MySQL, PostgreSQL, and Microsoft SQL Server.
  - o Amazon DocumentDB / MongoDB.
  - o Amazon DocumentDB.
  - o Amazon Redshift.
  - o Amazon RDS on Outposts.
  - o Amazon RDS on Local Zones.

Sustainability

Environmental sustainability is another shared responsibility between AWS and its customers.
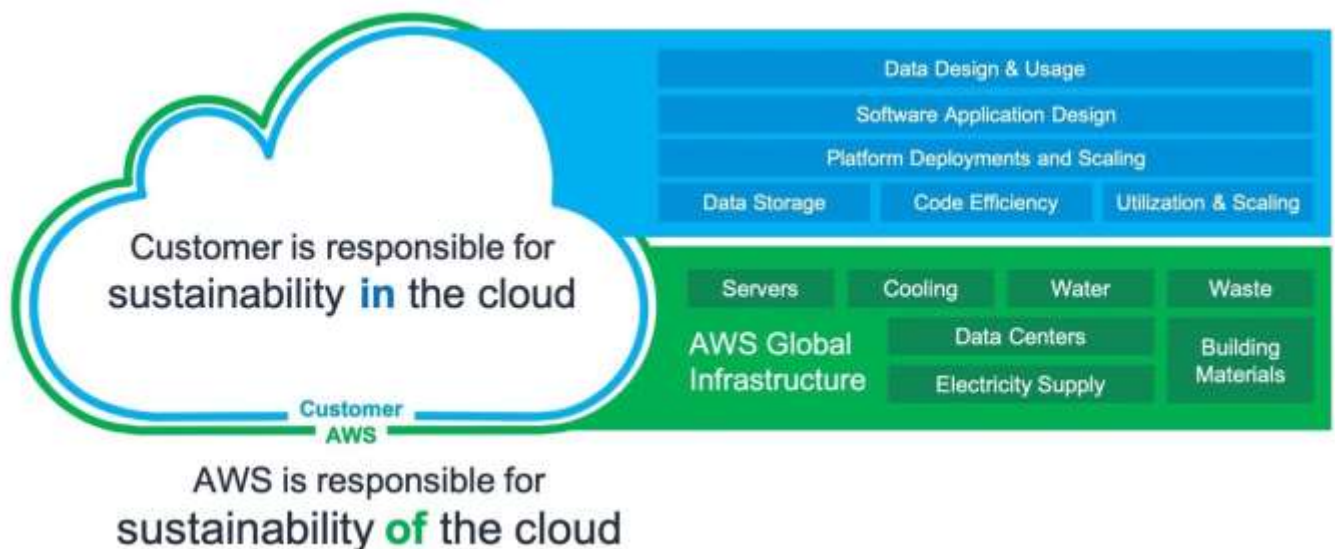
Source: The shared responsibility model, **docs.aws.amazon.com/wellarchitected/latest/sustainability-pillar/the-shared-responsibility-model.html**

AWS is responsible for optimizing the sustainability of the AWS cloud by optimizing the physical resources that contribute to the carbon footprint of AWS cloud services. This includes delivering next-generation compute, storage, and networking infrastructure that drives efficient consumption, sourcing renewable power, and using sustainable water practices.

AWS treats the responsibility seriously, tracking direct emissions (I.e., fuel combustion by backup generators), indirect emissions (i.e., consumed electricity), and all other indirect emissions (i.e., transport costs of delivering hardware to AWS data centers).

Customers are responsible for sustainable practices "IN" the cloud. This includes optimizing workloads to avoid waste, terminating unused workloads, and minimizing the resources uses to the minimum required to service the business need.

Adopting Commvault as your intelligent data management platform in AWS allows you to address your shared sustainability responsibility by:

- Minimizing hardware required to perform multi-region, multi-account data protection.
- Powering down Amazon EC2 instances when they are not actively required, with auto-Power on when needed (**MediaAgent Power Management**).
- Provisioning and then terminating Amazon EC2 instances only when data management activities require them (**Auto-scaling Access Nodes**).

Zero trust architecture

The adoption of the cloud has meant that traditional perimeter defense and defense-in-depth approaches to securing an organization's applications and data are no longer acceptable. **Zero trust** (ZT) cybersecurity models are built on the following base assumption:

> Assume that an attacker is always in your environment and that an enterprise-owned or operated environment is no different or trustworthy than any non-enterprise-owned environment.

NIST Special Publication 800-207 – Zero Trust Architecture
https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-207.pdf

This approach drives a fundamental change in how applications are architected, designed, and operated and is built on five (5) to eight (8) pillars – Commvault has features and functions that allow the implementation of a Zero trust architecture which are detailed below.
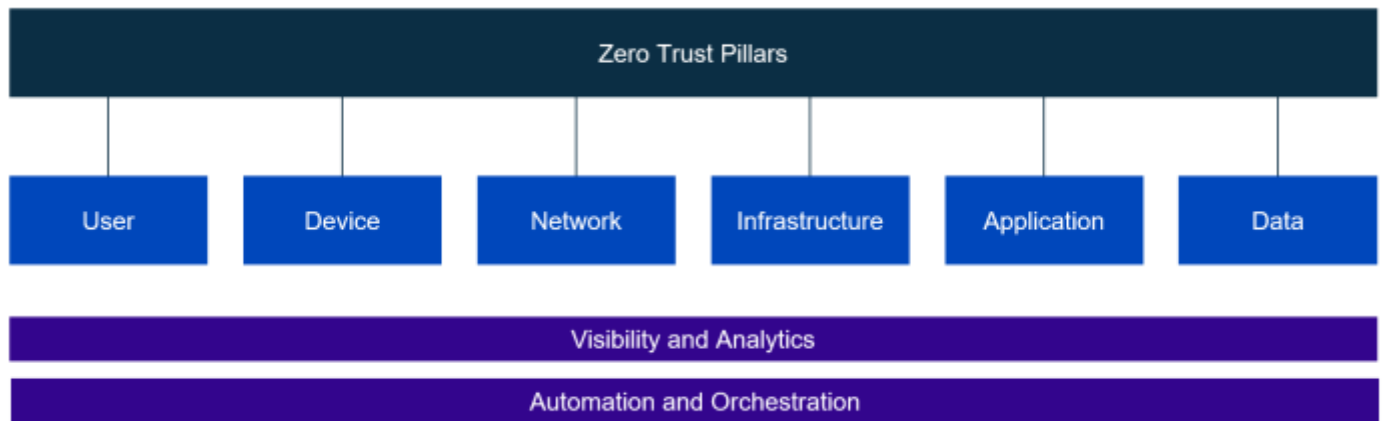
User Trust

Commvault software provides a centralized identity store that identifies individual users and groups permitted to interact and operate the multi-tenanted Commvault data management platform.

Authentication of users may occur utilizing Single Sign-On (SSO) centralized identity stores such as Active Directory (AD), secure LDAP, and externalized identity and access management (IAM) systems accessed using SAML 2.0 and OAuth (i.e., Okta, Ping, SecureAuth). The use of externalized IAM systems allows dynamic authentication rules that can incorporate location, time of day, and other metadata to determine if a user authentication request should be granted.

Additionally, access to Commvault administrative interfaces supports Multi-Factor Authentication (MFA) / Two-Factor Authentication (2FA) with support for Commvault and industry-leading web authentication applications (Google and Microsoft Authenticator). Hardware-based devices supporting Fast Identity Online (FIDO2) like Yubico Yubikey can be used as a second-factor device.

Common Access Cards (CAC) may also be used to perform password-less authentication.

Once authenticated, Commvault has a privileged access management system that combines the user/group, a role, and one or many entities (application, virtual machine, containers) that the user is permitted to act upon. All sessions are authenticated with a configurable default timeout of 30 minutes.

Events within the Commvault system that indicate a penetration or threat is present can programmatically modify a user or user group access dynamically by utilizing the Commvault REST API. Commvault has multiple anomaly detection and machine-learning algorithms that track and automatically respond to events and user behavior that is indicative of a threat (i.e., multiple failed login attempts).

Commvault software implements a secure multi-tenancy model where tenants are created as securely separated as logical entities called **companies**. Commvault considers tenant isolation and granular role-based access control and a specialized form of macro-segmentation between companies. Individual tenant user rights and access controls are considered by Commvault as micro-segmentation within the tenant.

Device Trust

Commvault performs active vulnerability management and reporting, for issues that impact Commvault products at **documentation.commvault.com/2022e/essential/146231_security_vulnerability_and_reporting.html**. Vulnerabilities may require an update to Commvault software and/or Operating system and third-party software on client devices.

Commvault endpoint protection software provides device security and enterprise mobility management for the protection of high-risk mobile data (i.e., laptops, tablets, phones). Commvault endpoint and edge protection include mechanisms to perform data-loss prevention periodic encryption, identify device location, and initiate automated remote wipes in response to lost devices.

Commvault software maintains a unique device identity by enrolling all protected devices with a device-specific cryptographic certificate that is managed and rotated periodically by Commvault. All communications (control and data plane traffic) implement device authentication using the certificate to establish and validate identity. All communications between the device and Commvault data management infrastructure are encrypted using an AES256 cipher.

Commvault software assists in device management by providing centralized device configuration reporting and insights. Additionally, Commvault software automatically downloads and deploys software patches and updates per the defined policy on all client and core data management devices/appliances. Commvault can also manage the deployment of Windows Operating System (OS) updates if required.

Commvault centralized reporting provides a device inventory of all protected hardware and details all installed protection modules and configuration.

Network Trust

Commvault software allows deployment into any communications network topology and enforces the network rules or network access control policy for the organization. Commvault software supports direct connections, port-forwarding network gateways, DMZ-based network gateways, authenticated HTTP proxies, and advanced network topologies that enforce one-way, two-way, and bi-directional tunneling on discrete ports.

Commvault software effectively provides a software-defined network topology and/or software-defined perimeter that mimics or mirrors the network used for data protection. Network topologies may be modified programmatically using CLI, SDK, and REST APIs to provide dynamic network controls and configuration in response to detected threats.

Commvault software provides micro-segmentation at the workload level. Each device running the Commvault software core package includes an application-level firewall that allows discrete access control on defined ports and protocols. Additionally, macro-segmentation is provided via network topologies and network gateways/firewalls that dictate how data can flow between Commvault data management components.

Commvault encrypts all data in-transit (control plane, data plane) using per-device symmetric cryptography where the same key is used for encryption and decryption, and AES-256 cipher suite is used by default. Communication is session-based with re-authentication required periodically to ensure devices and users reestablish their privileges. Commvault software crypto library implementation has been certified as **FIPS-140-3 cryptographic module validation program compliant** pending CMVP review.

Infrastructure Trust

Commvault keeps enterprise workloads secure while migrating between cloud environments by encrypting control and data-plane traffic in-transit. Additionally, VM conversion activities that utilize temporary cloud storage locations may write to encrypted-only object storage buckets, with cloud-provider or customer-managed keys (CMKs) specified.

Commvault is a cloud access security broker (CASB), with multi-cloud, multi-account permissions to protect (read) and restore or migrate (write) data from and to cloud environments. Commvault uses privileged cloud access credentials to perform cloud data management and provides an access control layer that authenticates and authorizes users before allowing access to cloud resources. All actions attempted and executed against a cloud are logged in the Commvault immutable audit log for traceability and forensic analysis.

Application Trust

Commvault software utilizes multiple web application components to service HyperText Transfer Protocol (HTTP) requests to Commvault Command Center™, WebConsole, and REST API endpoint(s). Commvault software supports the use of an Operating System (OS) firewall and **attack surface reduction rules** to automatically block known malicious behaviors.

Commvault recommends implementing cloud provider web application firewalls like **AWS Web Application Firewall (WAF)** in front of Commvault web-services to provide an additional level of threat detection and mitigation.

Commvault development practices employ multiple methods to develop and maintain secure data handling at all stages of data management. Commvault development practices require peer review from multiple parties including security domain specialists. Commvault performs quarterly static vulnerability scanning and remediation on Commvault software and third-party libraries and performs penetration testing both internally and via third-party engagements. Commvault is committed to detecting and resolving security issues rapidly and provides methods for individuals and organizations to report security defects for prioritized resolution.

Commvault utilizes the **Quay/Clair vulnerability scanner** on container images utilized by the Commvault software and ensures zero (0) issues are reported. Commvault software uses the `:latest` tag for actively maintained official docker images to ensure Commvault software is always using the most current and patched OS image.

Commvault software and by extension Metallic Data Management as a Service (DMaaS) both provide a secure access cloud layer that provides a privileged access control layer to authenticated and authorized organizational users and user groups. Granular access to perform read-only reporting restores to new locations or restores to the original location for service recovery can all be granted.

Commvault software employs a least privilege approach to cloud data management requesting only the minimum permissions required to protect an AWS service. Commvault role-based access controls (RBAC) are then overlayed to further restrict the individual user and/or user group.

Commvault software is accessible via the HTTPS portal and provides an 'any device access' approach to enterprise backup and archival data. Data may be accessed from any device (PC, tablet, mobile phone) and downloaded or restored to any location by authorized users.

Data Trust

Commvault software provides software and hardware encryption for data-in-transit and data-at-rest in the cloud and on-premises data storage locations. Commvault has a built-in FIPS-140-3 compliant cryptographic library ^{pending CMVP review} for generating and rotating encryption keys stored securely within the Commvault Database (CSDB). Alternatively, customers may choose to utilize Cloud-provider Key Management Services like **AWS Key Management Services (KMS)**. Customers may stay in control of their keys with a cloud-based or on-premises KMIP-compliant hardware security module (HSM) like **AWS CloudHSM**.

Commvault integrates natively with cloud-provider key management services to transparently access encrypted data-in-use in unencrypted form. Commvault accesses encrypted application data in unencrypted format and transfers securely via an encrypted tunnel to Commvault encrypted storage. When transferring cloud-native snapshots between encryption boundaries, Commvault decrypts and then re-encrypts snapshots with customer-selected encryption keys.

Commvault software provides data security by allowing data classification and tagging by detecting sensitive content, data protection based on data classification, and data masking to prevent data leakage/spillage in cross-environment data restore (i.e., production to development seeding). Commvault prevents data leaking by performing redaction for sensitive email and indexed data fields during export to PDF for sharing with external parties.

Commvault software helps businesses comply with relevant industry, geography, and organizational regulations and policies. Commvault has been successfully deployed and certified to many industry standards, a published list of relevant standards where Commvault implementations have been certified is here:

Certifications and Compliance
https://documentation.commvault.com/2022e/expert/110316_certifications_and_compliance.html

Commvault is responsible for the integrity of data persistent within Commvault-controlled cloud, HyperScale™, disk, and tape storage locations. Commvault performs periodic data verification jobs to validate stored data has not been modified since the initial backup. Data verification in cloud storage locations is not recommended due to the recall or retrieval cost. The durability of cloud storage and storage of multiple independent copies prevents the need to perform periodic costly data verification.

Commvault **Risk Analysis** (formerly File Storage Optimization and Data Governance), and **eDiscovery and Compliance** index and classify backup, archival and live data to organize and optimize data by risk and value. Data may be analyzed based on access and modification frequency, access control settings, and the presence of personally identifiable information. Legal and compliance requests can perform granular keyword searches and exports to respond to external and internal forensic investigations.

Visibility and Analytics

Commvault software provides threat intelligence via a granular audit log of all activities occurring within the data management platform. Events may be forwarded to external Security Incident and Event Management (SIEM) or Security Orchestration Automation and Response (SOAR) systems via Syslog, SDK, or custom action.

Commvault recommends that audit events are forwarded to **Amazon CloudWatch** infrastructure and application monitoring to visualize, report, and automate responses to critical events and threats. Additionally, Amazon CloudWatch integrates with Amazon Inspector, Amazon GuardDuty, and AWS Security Hub (with event transform and load) so end-to-end visibility of threat progression can be observed.

Commvault provides continuous diagnostics and mitigation capability via centralized reporting and alerts for infrastructure (servers, virtual machines, desktop/laptops) that are lagging from the most current software patches.

Automation and Orchestration

Commvault software represents a centralized policy engine (PE) that is responsible for evaluating a user or user group request for access to a resource. The policy is implemented as a three-way relationship between a user/user group, role, and a Commvault entity (server, VM, database, application, etc.).

Commvault provides the Commvault Firewall Daemon (CVFWD) on all data management infrastructure, which performs the role of a policy administrator (PA) establishing and terminating encrypted communication tunnels for authorized data management activities.

Commvault software also performs policy enforcement (PE) from the centralized CommServe® instance which monitors, initiates, and terminates communication as required to complete data management activities. The CommServe Job Manager communicates with individual policy administrator instances to monitor and terminate connections on-demand.

Additional resources

- NIST Special Publication 800-207 Zero Trust Architecture
  https://csrc.nist.gov/publications/detail/sp/800-207/final
- NIST Special Publication 800-63-3 Digital Identity Guidelines
  https://pages.nist.gov/800-63-3/
- Whitehouse.gov – Memorandum For The Heads of Executive Departments And Agencies
  https://www.whitehouse.gov/wp-content/uploads/2022/01/M-22-09.pdf
- U.S Cybersecurity & Infrastructure Security Agency (CISA) Cloud Security Technical Reference Architecture
  https://www.cisa.gov/cloud-security-technical-reference-architecture
- U.S General Services Administration – Zero Trust Architecture Buyers Guide
  https://www.gsa.gov/cdnstatic/Zero%20Trust%20Architecture%20Buyers%20Guide%20v11%202020210610%20(2).pdf
- How to think about Zero Trust architectures on AWS **https://aws.amazon.com/blogs/publicsector/how-to-think-about-zero-trust-architectures-on-aws/**
- How to build a Zero Trust Recovery Solution with Commvault and Metallic
  https://www.commvault.com/blogs/build-a-zero-trust-recovery-solution-with-commvault-and-metallic
- Zero Trust Networks, Evan Gilman, Doug Barth, O'Reilly Publishing
  https://oreilly.com/library/view/zero-trust-networks/9781491962183/ch01.html

Well-Architected Framework

The **AWS Well-Architected framework** helps cloud architects learn, design, and build secure, performant, and resilient cloud solutions. The AWS Well-Architected framework consists of six pillars –, Security, Reliability, Performance Efficiency, Cost optimization, and Sustainability.

Commvault recommends using the AWS Well-Architected framework when designing and building your Commvault data management platform in the AWS cloud. This section provides the Commvault software features and functions that align with well-architected guidance and provides reference architectures, design patterns, and best practices that should be used to deliver optimal price, performance, and resilience in AWS.

Operational Excellence Pillar

The **Operational Excellence** pillar guides how to run your workloads effectively to address your unique business objectives and service level agreements (SLAs). Operational Excellence provides the guardrails for operating your Commvault software to meet your recovery point objectives (RPOs) and recovery time objectives (RTOs) in a repeatable fashion, regardless of the constant change in your application landscape.

Amazon publishes a set of best practices for the Operational Excellence Pillar. The sections below detail how to implement these best practices with Commvault software, refer to the **Operational Excellence Pillar Documentation** and **Labs** for the latest Amazon recommendations.

Organization

To align your ongoing data management operational activities, it is key to understand your role in addressing the organizational priorities that support your business. Understanding your organization includes:

Organization Priorities

OPS01-BP01 Evaluate external customer needs.

Commvault recommends building a cross-functional team across IT, operations, development, finance, legal, governance risk & compliance, and application owners to establish and maintain data management needs. Maintain a central request or **backlog** of customer feature requests for future improvements.

Use your service desk (i.e., **ServiceNow**) to understand the products, services, and incidents customers are experiencing with data protection, recovery, and lifecycle management. Explore services that are experiencing high levels of escalation for opportunities to optimize delivery to meet business needs (i.e., after-hours mission-critical recovery assist).

Commvault provides **Recovery Readiness**, **SLA**, and **Strike Count** (repeated failure) insights to isolate data protection hotspots that are not meeting business needs.

Related documents: Operational Excellence – OPS01-BP01 Evaluate external customer needs.

OPS01-BP02 Evaluate internal customer needs.

Commvault recommends that data management teams engage cross-functional shared services teams including network, security, operations, and development to focus and coordinate improvement efforts where the greatest business value will be achieved.

Understand the business goals, technology needs, and immediate priorities impacting the ability to deliver. Promote a **data management as a service** approach with stated outcomes and service levels to internal stakeholders so that improvements can be shared across teams and workloads.

Create a central repository of workload knowledge including but not limited to what business function the workload provides, teams responsible for building/operating/maintaining the workload, and business classification for the workload (business-critical, tier1, tier2, tier3).

Commvault **entity tags** and **AWS Resource tags** should be used to document workload owners, operators, and business-classification or importance.

Related documents: Operational Excellence – OPS01-BP02 Evaluate internal customer needs.

OPS01-BP03 Evaluate governance requirements.

Commvault recommends engaging your legal counsel and governance, risk & compliance team to understand and potentially establish internal policy, standards, and procedures for data protection, retention, and destruction. **Risk Analysis** can be used to help find and visualize your critical and sensitive data for analysis against internal policies and procedures.

Ensure that your internal policies and procedures include an **eDiscovery and Compliance** function to streamline data collection concerning internal or external discovery requests. Consider how fast, efficient, and scale collection of relevant data in emails, file-servers, laptops, and documents can be performed across AWS and edge-based locations. Work with legal counsel to understand your requirements for ensuring a chain-of-custody for eDiscovery cases and requirements for marking data under investigation as immutable (i.e., legal hold).

Considering using Commvault reports like the **CommCell Growth Details in the Growth and Trends Report** to educate legal counsel on overall data retained for the business. Internal policies that dictate the data retention period, and the number of copies to retain will benefit from *data-driven decisions* based on governance and cost requirements.

Commvault provides prescriptive guidance on the **ports and protocols** required to secure Commvault communications which may require review against instance security policies before releasing to production.

Related documents: Operational Excellence – OPS01-BP03 Evaluate governance requirements.

OPS01-BP04 Evaluate compliance requirements.

Commvault recommends engaging your legal counsel and governance, risk & compliance team to understand your external data management requirements, including regulatory compliance, industry standards, and relevant **certifications** for third-party proof of compliance.

Ensure your team has a clear understanding of each workload under protection, all data and services used by the workload, and their external regulatory responsibilities to protect and optionally recover the workload.

Commvault recommends a **data classification strategy** that prioritizes external (customer-facing) then internal (business-facing) workloads. Operational efforts to improve compliance, and performance, reduce cost or automate complex operations should be prioritized based on the value of the application/service to the business.

Commvault publishes a list of certifications and standards that Commvault software confirms with:- **Certifications and Compliance**.

Related documents: Operational Excellence – Evaluate compliance requirements.

OPS01-BP05 Evaluate threat landscape.

Evaluate threats to your data management practices, operational risks, and informational security risks. Maintain a **threat model** and develop mitigations to threats that could adversely impact business operations (i.e., minimize insider threats with **active approvals**, proactively monitor and **alarm unusual activity and ransomware-like events**).

Monitor **Commvault Security Vulnerability and Reporting** updates to ensure you are running the most current and secure software version.

Review Commvault-recommended controls for **Ransomware protection** which includes mitigations for multiple components within your Commvault data management platform.

Commvault will monitor and surface security recommendations in the **Health Report – Security Assessment Tile**, issues may also be proactively alarmed to SecOps teams for action. The **Security assessment tile** includes suggested security controls that could improve the security of the Commvault data management platform, like Two-factor authentication, Single-Sign-On, Password complexity, actions on failed login, and many others.

Commvault Threat Scan allows organizations to scan backup files for malware and analyze backup files for high levels of entropy and change, which can be indicative of ransomware activity. If using Metallic Data Management as a Service (DMaaS), consider utilizing **Metallic ThreatWise** to provide active cyber deception while uncovering, containing, and minimizing threats in your environment.

Related documents: Operational Excellence – OPS01-BP05 Evaluate threat landscape.

OPS01-BP06 Evaluate tradeoffs.

Evaluate the impact of tradeoffs in your data management system and maintain a register of architectural designs and design decisions that trace reasoning for as-built systems. Take a *data-driven approach* to assess the impact of a trade-off, for example, use **Commvault Risk Analysis** to identify whether a workload stores sensitive data before deciding on changes its access network, security, or data protection processes.

Use Commvault **docs**, **community**, **maintenance advantage**, and **knowledgebase** to stay informed on the latest advancements and generate active discussion on the pros/cons of trade-offs in data management.

Related documents: Operational Excellence – OPS01-BP06 Evaluate tradeoffs.

OPS01-BP07 Manage benefits and risks.

Maintain a central register of the business goals, needs, and priorities and assess initiatives for benefit and risk to those goals. Maintain a registry of known risks, architectural decisions, and design decisions as a living system for operational personnel, architects, and auditors.

For example, you may be able to improve time-to-market, security, reliability, performance, and cost but at the cost of increasing risk such as slowing application restore times. Commvault recommends security as job #1 in all proposed changes, rank your business goals and priorities to help prioritize changes.

Commvault publishes reports on **recovery readiness**, **performance**, and **cost modeling** to help assess initiatives.

Related documents: Operational Excellence – OPS01-BP07 Manage benefits and risks.

Operating Model

It is key that data management teams have a clear understanding of their role and responsibility in protecting workloads, which teams they are dependent upon, how decisions are made, and how escalations are raised when conflict arises.

A team may be responsible for deployment and support for the shared data management platform, the shared data management application, and/or the shared technology standards related to data management.

Regardless of the **Operating Model** selected, the following core principles should be observed:

- Identification of faults should occur automatically, and log incidents for response by the responsible team (see **Alert Targets** for more information) to prevent issue resolution delays.

- The transition of incidents or service requests between teams should occur within a ticketing system providing traceability, activity history, and service level tracking and escalation (see **Creating Incidents on ServiceNow**) to facilitate the timely transition of incidents between teams.

- Visibility of achieved service levels and/or active events for the data management system should be available to all affected teams using a read-only reporting role for owned or operated workloads (i.e., **Recovery Readiness Report**).

- **Owner security** must be applied to all Commvault-protected clients and Commvault-infrastructure to allow timely self-service recovery and alerting. All data management resources must have an identified owner (user or user group) to allow the assignment of responsibility for break/fix activities.

Ownership of protection responsibility, recovery responsibility, and the data management system itself is crucial to ensuring that when a data-loss event occurs, the business responds and recovers without delay. Commvault has a rich **role-based access control (RBAC)** framework that allows deployment and granular permission assignment on a user, user group, and entity (server, client, instance) basis.

Operating model 2 by 2 representations

There are several different operational models encountered across the diverse business landscape using Amazon Web Services.

Commvault may be deployed and operate in a secure, governed, self-service model regardless of the model selected. Commvault provides a rich **role-based access model (RBAC)** based on assigning permissions to users and user groups. Additionally, Commvault can operate in multi-tenant mode, where each line of business, workload, or department is represented as a *company*, with dedicated authentication, dashboards, reporting, and self-service rights.

- **Fully separated operating model**
- **Separated Application Engineering and Operations (AEO) and Infrastructure Engineering and Operations (IEO) with centralized governance**
- **Separated AEO and IEO with centralized governance and a service provider**
- **Separated AEO and IEO with centralized governance and an internal service provider consulting partner**
- **Separated AEO and IEO with decentralized governance**

Relationship and ownership

OPS02-BP01 Resources have identified owners.

Commvault recommends assigning individual or group owners for all protected workloads (entities) from virtual workstations, traditional and containerized compute instances, and cloud databases and storage. Ownership should be set as **owner security setting** to allow user self-service and reflected in **entity tags** and **AWS Resource tags**.

Commvault recommends owner information is reflected in AWS tags and/or resource groups for all AWS accounts, services, and resources. Owner information may be replicated in Commvault with entity tags and optionally **Smart Server groups** to provide facilitate automation of workloads at scale. Commvault will protect, recover, and allow customization of AWS resource tags for some AWS workloads, and restore protected tags unmodified for remaining workloads.

Centralized operational documentation should identify what *ownership* means and record a workload RACI model (see **example RACI model**) indicating the owner's tasks for design, building, operating, securing, and recovery of a workload.

Related documents: Operational Excellence – OPS-BP01 Resources have identified owners.

OPS02-BP02 Processes and procedures have identified owners.

Commvault recommends that data management processes and procedures are stored and shared with authorized individuals. Processes must include a responsible team for the process and the responsible party for executing the process.

Ownership of the activity will help define data management **roles** and **security associations** within Commvault, which allow secure governed self-service data management across your organization.

> 💡 **Pro-Tip**
>
> Do not store sensitive information such as contact email or phone number in AWS Resource tags or Commvault entity tags. Ensure an organizational system to lookup contacts is available to link workload metadata to organizational users.

Related documents: Operational Excellence – OPS02-BP02 Processes and procedures have identified owners.

OPS02-BP03 Operations activities have identified owners responsible for their performance.

Commvault recommends documenting the owner and/or team responsibilities for both protected workloads and the Commvault data management platform itself. Responsibilities and engaged teams may change based on the *criticality* of the workload or recovery use-case (i.e., the recovery of the Commvault data management platform may be processed as a *disaster event* due to organizational impact).

Commvault role-based access can enforce responsibilities by limiting the recovery scope that a user or group may perform (e.g., users may be granted **Out of Place Recover** for testing but not **In Place Recover** which would impact the production application). Operations activities should also identify the individuals or teams required to provide active approval of high-risk activities using **business logic workflow** approvals.

This operational ownership information must be open and discoverable to allow timely identification of individuals, teams, and escalation managers during an unplanned application outage.

> 💡 **Pro-Tip**
>
> Your operational processes and ownership information are crucial during a wide-spread outage event, ensure these processes are stored in a highly-available storage service (e.g., Amazon EFS, Amazon FSx) and backed-up with Commvault.

Related documents: Operational Excellence - OPS02-BP03 Operations activities have identified owners responsible for their performance.

OPS02-BP04 Team members know what they are responsible for.

Commvault recommends that **entity tags** are attached to all data management resources (hypervisors, VM groups, cloud databases, cloud storage, credentials) to indicate the primary owning team or individual.

Team members and workload (application) owners can then use **Views** to understand the scope of the workloads and resources they are responsible for.

Operational processes should include the roles, and responsibilities for the operation of the workload using accepted models like a RACI model (see **example RACI matrix**), tied to the entity tag that identifies to individual or team member responsible. Data management team members must have a clear understanding of their responsibility to protect and recover the workload.

Related documents: Operational Excellence – OPS02-BP04 Team members know what they are responsible for.

OPS02-BP05 Mechanisms exist to identify responsibility and ownership.

Commvault recommends setting a **customHelpUrl** in Commvault Command Center™ to allow self-service users to reach operational teams or service desk personnel without delay, to assist in identifying system owners and operations teams. Traceability to ownership of resources can occur via establishing the owning AWS account, business unit, and user or role who created the resource.

Identification of the owning individual or team is simplified by applying **AWS Resource tags** and **entity tags** on the resource or workload.

Consider using tags to direct users to a URL or email alias that can assist in identifying workload ownership. Do not store sensitive or personally identifiable information in tags. You can maintain ownership visibility by protecting and recovery AWS Resource tags when protecting supported AWS workloads with Commvault.

Related documents: Operational Excellence - OPS02-BP05 Mechanisms exist to identify responsibility and ownership.

OPS02-BP06 Mechanisms exist to request additions, changes, and exceptions.

Commvault recommends that all changes are submitted via a service request system to track, assign, and trace request completion time. Commvault integrates with many industry-leading service desks including **ServiceNow**.

Standardized changes may be automated in the service desk catalog, ad hoc changes may be specifically requested by the user for review by operations teams.

It should be noted that the addition of new workloads for protection with pre-configured **Plans** occurs automatically at backup time and can use AWS Resource tags to auto-discover resources to protect. Likewise, changes to the protection Plan (data retention, frequency, number of copies) for a workload can be modified by authorized cloud users by updating the AWS Resource tag attached to their resource. You can ensure that Commvault-restored AWS workloads are automatically protected or excluded by protecting or customizing AWS Resource tags during or after recovery.

Related documents: Operational Excellence - OPS02-BP06 Mechanisms exist to request additions, changes, and exceptions.

OPS02-BP07 Responsibilities between teams are predefined or negotiated.

Commvault recommends per team operational agreements (OLAs), or third-party service-level agreements (SLAs) are stored in a central secure shared storage area, accessible during day-to-day operation and disaster recovery events.

Formal and informal communication channels should be established between dependent teams to remove delays and drive the timely resolution of requests (i.e., Slack, Microsoft Teams).

An understanding of each team's predefined responsibilities allows the definition of **Custom Roles** within Commvault that enable authorized users to self-service their data management needs in **Commvault Command Center™**.

Related documents: Operational Excellence – OPS02-BP07 Responsibilities between teams are predefined or negotiated.

Organizational Culture

An organizational culture that promotes the adoption of best practices and continual evolution of the organization's shared services, and employees' skills are important for ongoing operational performance. Team members should be empowered to experiment and take risks, especially when business outcomes or SLAs are at risk. Communication within and across teams must be seamless, and senior leadership must be approachable to resolve issues before they impact business objectives or SLAs.

OPS03-BP01 Executive Sponsorship.

Commvault recommends that data management policies and processes that impact business services receive sign-off from a corporate sponsor or signatory. This assists in resolving disparate expectations on service levels across the business (i.e., response times, the data retention time for backup, and operational support hours).

Executive sponsors will expect that **SLA Reports** showing Met, Missed, and Excluded resources from protection, and achieved service-levels are published and assist to promote an environment of continual improvement.

Measuring, monitoring, and publishing organizational **recovery readiness** is recommended to drive the adoption of best practices across the organization and evolve application design, data management, and overall organization cloud capability.

Related documents: OPS03-BP01 Executive Sponsorship.

OPS03-BP02 Team members are empowered to take action when outcomes are at risk.

Commvault recommends empowering team members to propose, test, and implement improvements when the **Strike Count** increases or **SLAs trend downward** . Permission to perform these changes is possible through the Commvault **role-based access control (RBAC)** system.

Commvault recommends **regular restore testing** and **game-day simulated failure** testing that builds an understanding of the tradeoffs between protection cost and recovery performance. Consulting with workload owners on identified risks allows a balanced approach to risk acceptance and mitigation.

Related documents: OPS03-BP02 Team members are empowered to take action when outcomes are at risk.

OPS03-BP03 Escalation is encouraged.

Escalation is promoted within and across teams to help engage key decision-makers and stakeholders. Commvault provides **notification escalation** to repeat and escalate conditions requiring human attention to resolve.

Commvault recommends that all protected workloads or resources have a *business classification* attached via AWS Resource tags and/or entity tags to help inform appropriate urgency during escalations. You can protect these *business classifications* by using Commvault to protect your AWS workloads (including. attached tags).

You can use the **Health Report: Strike Count** to obtain a view of the number of repeated failures a workload has experienced over the last one, two, or three consecutive days. Repeated failures indicate a failure condition that requires immediate attention to return to a *protected* state.

Related documents: OPS03-BP03 Escalation is encouraged.

OPS03-BP04 Communications are timely, clear, and actionable.

Commvault recommends maintaining a central change calendar or diary to inform the organization of planned outages. Commvault software provides the ability to put key components into **Maintenance Mode** or disable **Backup / Restore activity** to help inform operational teams when planned maintenance activities are underway.

Commvault **Webhooks** can automate the push of critical alerts to systems like internal communications platforms like **Microsoft Teams** or **Slack**.

> 🔆 **Pro-Tip**
>
> Commvault provides backup and recovery protection for **Microsoft Teams Chat messages** if an audit log of operational communications is required.

Related documents: OPS03-BP04 Communications are timely, clear, and actionable.

OPS03-BP05 Experimentation is encouraged.

Commvault recommends frequent individual and team **gameday events** to allow teams to experiment and learn from both successful and unsuccessful recovery experiments. Try to constrain experiments to a single goal, like improving recovery time, reducing cost, or improving the security posture of a workload. Maintain a register of experiments and successful and unsuccessful outcomes, these are a valuable resource for the team and organization-wide development.

Commvault can be deployed with free 60-day trial licensing from the **AWS Marketplace**, and use Out of Place Recovery to restore copies of workloads for testing so that Production environments are unaffected.

Commvault recommends engaging finance and line-of-business leaders to ensure that cloud budgets include the resources to support experimentation (e.g., Amazon EC2 runtime, Amazon S3 storage, and network transfer costs).

> ⓘ **Note**
>
> Always ensure **data masking** controls are in place if testing with production data.

Related documents: OPS03-BP05 Experimentation is encouraged.

OPS03-BP06 Team members are enabled and encouraged to maintain and grow their skill sets

Commvault provides multiple learning resources from **documentation**, **education advantage**, **community**, **on-demand learning library**, and **on-demand learning labs**.

Promote continual improvement and pursue **Commvault certifications** and **AWS certifications** to continually better and test member knowledge.

Reach out to your **Commvault sales representative** to find out when the next **Commvault + AWS Immersion Day** is planned for hands-on experience with the latest Commvault data management features.

Check the AWS-provided **development resources** for additional best practices, patterns, and learnings.

Related documents: OPS03-BP06 Team members are encouraged to maintain and grow their skill sets.

OPS03-BP07 Resource teams appropriately

Commvault recommends providing self-service backup and restore capability to application owners to offload effort from operational teams. Commvault simplifies complex recoveries with a self-service interface that does not require specialized knowledge of AWS snapshot-managed, encryption, or cross-region/cross-account data transfers.

Commvault recommends using a service desk so that an analysis of long-running or manual tasks can be performed and investment in automation can occur.

Use business-agreed performance measures for the team (i.e., SLA attainment, cloud costs) to measure the performance of the team and invest in required team resources when required. Data and protected workloads will

continue to grow, consider and investment in new team members focus on **automation skills** that allow managing more workloads with less manual effort.

Related documents: OPS03-BP07 Resource teams appropriately.

OPS03-BP08 Diverse opinions are encouraged and sought within and across teams.

Commvault recommends operating multi-discipline communities of excellence (COE) in data management. This provides opportunities for information sharing and experimentation between application owners and data management teams.

Data Management has evolved from traditional *backup & recovery*, to include the following disciplines that your developers, application owners, and operations team should explore:

- **Data Protection**
- **Foundational Security**
- **Forensics & Monitoring**
- **AI & Automation**
- **Governance and Compliance**
- **Cloud & Storage Optimization**
- **Unified Data Management**
- **API-First**

Related documents: OPS03-BP08 Diverse opinions are encouraged and sought within and across teams.

Prepare

To prepare for unanticipated events, it is key to understand how Commvault software operates and integrates with your protected workloads. To prepare for protecting your workload, you should perform the following:

Design Telemetry

OPS04-BP01 Implement application telemetry.

Commvault writes application events and metrics to a centralized **Log Files** location on all Commvault instances, logs may be viewed for an individual Job or a **specific instance**.

Commvault recommends installing the **Amazon CloudWatch agent** on all Commvault instances and optionally protected workloads, to collect and forward logs, events, and metrics to CloudWatch. Use Amazon CloudWatch **Creating metrics from log events using filters** to extract performance metrics from Commvault Logs (e.g., search for `GB/hr.`, `MB/sec`). Centralizing your workload and Commvault logs, events, and metrics allows troubleshooting to take a complete view of your application and data protection instances.

Commvault logs provide critical event and **resource utilization metrics** used for baselining normal operations and performing root cause analysis (RCA) when an issue occurs. Instance-related performance metrics are stored in `Base\Log file\ResourceMonitor` folder (see **Monitor System and Commvault Resource Logs**) and may be visualized in the **Infrastructure Load** report. Alternatively, Amazon CloudWatch provides **instance metrics** that provide the same metrics plus additional Amazon EC2-specific metrics (i.e., EBS performance, ENA performance).

Commvault log files prefix all log entries with a **<job_id>** which may be used to query logs in **Amazon CloudWatch Logs Insights** and trace an individual activity (transaction) across multiple Commvault components.

Related documents: OPS04-BP01 Implement application telemetry.

OPS04-BP02 Implement and configure workload telemetry.

Commvault recommends installing Amazon CloudWatch agents on Commvault instances (CommServe, MediaAgents, and Access Nodes) and forwarding application logs to **Amazon CloudWatch Logs** for centralized observability and insight.

Commvault recommends configuring additional telemetry for your Commvault data management platform by using a **command-line alert & notification** to **publish custom metrics** to Amazon CloudWatch for any Commvault alert type. Commvault alerts may be generated on SLAs, throughput, or health-based alerts checks and measures requiring attention (see **Predefined Alerts**).

Consider enabling additional AWS-based telemetry by enabling **VPC Flow Logs** for any Commvault data transfers, and **AWS CloudTrail** for all AWS accounts and machine identities configured in Commvault.

Related documents: OPS04-BP02 Implement and configure workload telemetry.

OPS04-BP03 Implement user activity telemetry.

Commvault includes a detailed immutable **audit trail** that shows all user actions performed, helping identify what data management capabilities are being used. Commvault has built-in dial-home telemetry to aid in customer support cases that record anonymized feature activity to provide incident and problem resolution (see **Activating Cloud Metrics**).

Related documents: OPS04-BP03 Implement user activity telemetry.

OPS04-BP04 Implement dependency telemetry.

Commvault components that are dependent on other sub-components include a **check readiness function** (reachability check) to validate the operational state of local and remote components (also available via REST API).

Commvault is also dependent on AWS regional and global **service endpoints**, should a data management activity attempt to contact a service endpoint fail, the attempt will be logged in the relevant application log file (which should be collected and forwarded to CloudWatch for alarming).

Related documents: OPS04-BP04 Implement dependency telemetry.

OPS04-BP05 Implement transaction traceability.

Commvault log files articulate transaction execution flow, including the AWS account or credentials being used. Failed AWS actions are logged with the return code and message from the service endpoint. Use the Commvault **job_id** to trace a transaction or activity across multiple Commvault instances using **Amazon CloudWatch Log Insights** queries.

For example:

```
22659 5883 09/16 03:34:13 4288 CAmazonInfo::Connect() - Connection Successful to source
account [nnnnnnnnnnnn]
```

Related documents: OPS04-BP05 Implement transaction traceability.

Design for Operations

OPS05-BP01 Use version control.

Use version control (where supported) to record all changes made to your Commvault infrastructure and software resources.

Commvault recommends using Amazon CloudFormation or equivalent Infrastructure as Code (IaC) tools (i.e., **Terraform**) to deploy and version control changes to your Commvault resources. Commvault publishes an AWS

Marketplace **all-in-one** AMI and **Cloud Access Node** AMIs for expansion into new regions or to handle more workload protection.

Commvault supports storing versioned configuration, configuration management scripts, and infrastructure as code definitions using **AWS CloudFormation** templates and **Terraform**. Commvault configuration management automation may be achieved using the **Amazon Systems Manager Run Command** executing Commvault **command-line**, **Python SDK**, **PowerShell Module**, or **REST API** commands.

Additionally, Commvault protects **GitHub** and **Azure DevOps** source code repositories to protect your version-controlled repositories from unplanned data loss.

Related documents: OPS05-BP01 Use version control.

OPS05-BP02 Test and validate changes.

Commvault recommends that changes to your data management platform occur in isolated **canary deployments** or **blue/green deployments.** Change proposals may be validated in staged deployment within test, pre-production, and then isolated production resources using AWS resource tags to control workloads selected for the test. Workloads can be directed to use specific MediaAgents and/or Access Nodes without affecting other data management activities. Baselines should be established for key workloads to ensure changes do not adversely affect recovery performance.

Related documents: OPS05-BP02 Test and validate changes.

OPS05-BP03 Use configuration management systems.

Configuration management tools should be utilized to deliver, monitor, and maintain standardized, secure operating system and application configurations across your workloads. Commvault software can be automated and orchestrated by configuration management tools like **Amazon Systems Manager**, Chef, Puppet, CFEngine, and Saltstack leveraging the Commvault **Command Line**, **Python SDK**, **PowerShell**, or **REST API**.

Infrastructure as code and configuration management automation should enforce your design standards to avoid configuration drift from manual or unauthorized changes. Common application configuration settings used across your Commvault infrastructure can use **AWS AppConfig** to centrally store and manage common automation settings (i.e., storing the **web service endpoint** to contact Commvault for REST API automation).

AWS Systems Manage Change Calendar and **AWS Systems Manager Maintenance Windows** can be used in conjunction with Commvault **Blackout Windows** to ensure that configuration changes are not attempted during planned changes.

Commvault recommends using **AWS Config** to capture, monitor, and notify on changes in resource configuration across regions, accounts, and environment types (dev-test, production). In environments that have strict dependence on AWS resource quotas, consider solutions like the **AWS Solutions Library - Quota Monitor for AWS** to be notified when approaching quota limits.

Check **github.com/Commvault** for code snippets and development documentation.

Additionally, Commvault recommends protecting Amazon VPC configurations so that failed or unauthorized changes can be audited and remediations to security posture can be made. See **Auto-discovering, protecting, and recovering Amazon VPC resources** for the AWS IAM permissions required to collect and recover supported Amazon VPC resources. Commvault-recorded network configuration can be used to compare against your system-of-record and identify the root cause of failed changes (see **Restoring Amazon VPC network configuration**).

Related documents: OPS05-BP03 Use configuration management systems.

OPS05-BP04 Use build and deployment management systems

Commvault can be managed by continuous integration/continuous deployment (CI/CD) pipelines for initial provisioning, configuration management, and ongoing patch management using **developer tools**.

Commvault compute, network, and storage resources may be deployed using infrastructure as code (i.e., AWS CloudFormation, **Terraform**) services to reduce effort and change for human error. In fact, underlying IaC builds are stored in Amazon S3, which Commvault also protects to ensure the 'deployment' state is captured for future restore events.

Additionally, Commvault software supports immutable updates and can **repave** CommServe® instance, MediaAgents, and Access Nodes in environments where regular infrastructure refresh is required. Contact your Commvault sales representative to discuss immutable updates with Commvault.

Related documents: OPS05-BP04 Use build and deployment management systems.

OPS05-BP05 Perform patch management.

Commvault recommends regular patch management or software currency to obtain new features, address issues, and remain compliant with your organizational security policies.

Commvault releases **Platform Releases** every 6 months.

Commvault releases **Maintenance Releases** monthly.

Patch management should be automated in pre-production environments to validate that new software releases do not result in a reduction of protection coverage or performance, before pushing to production.

Commvault software may be configured to automatically **download patches** and **deploy patches** periodically. Any updates must update the CommServe, then MediaAgents, then Access Nodes, and then other clients. Deployment may target a subset of instances, so that rollout may be staged. Rollout occurs as a mutable update to existing pre-production or production instances.

Additionally, Commvault software supports immutable updates and can **repave** CommServe® instance, MediaAgents, and Access Nodes in environments where regular infrastructure refresh is required. Contact your Commvault sales representative to discuss immutable updates with Commvault.

Related documents: OPS05-BP05 Perform patch management.

OPS05-BP06 Share design standards.

Commvault recommends establishing a central repository for design standards and best practices that leverage the **Commvault Cloud Architecture Guide** (this document) and the **AWS Well-Architected framework** best practices.

Commvault recommends storing design standards and best practices in an openly accessible area that is securely shared with authorized individuals and/or teams, consider using the following AWS storage services for service design artifacts:

- Amazon Elastic File System
- Amazon FSx for Windows Server
- Amazon FSx for NetApp

Ensure you are protecting your **Amazon EFS**, and **Amazon FSx** file-systems with Commvault.

Related documents: OPS05-BP06 Share design standards.

OPS05-BP07 Implement practices to improve code quality.

Commvault employs multiple industry best practices to improve code quality including test-driven development, fault-injection testing, chaos engineering, code reviews, and standards adoption and certification. Update to the latest maintenance release or feature release to obtain the latest features and enhancements for your Commvault data management platform.

Additionally, customers may log application software defects via **Maintenance Advantage** or via **Reporting a Security Vulnerability** web-form.

Related documents: OPS05-BP07 Implement practices to improve code quality.

OPS05-BP08 Use multiple environments.

Commvault recommends maintaining multiple data management environments to validate the latest software changes, leverage **AWS Marketplace** AMIs to accelerate on-demand deployment of new environments, and then destroy after testing.

Long-running test environments may remain powered down to reduce cost and use infrastructure as code-cloned application workloads to test on-demand. Consider using the **AWS Solutions Library – Instance Scheduler** on AWS to automatically power-down test environments at night. Commvault Marketplace AMIs offer a free 60-day free trial license that can be used to perform testing.

Multiple environments can be easily deployed and maintained using **AWS CloudFormation** or the Commvault **Terraform module** for the management of AWS resources, and Commvault users, tenants, storage, and Amazon EC2 protection.

▲ Note Operating a pre-production Commvault environment for longer than 60 days will need to be licensed. Submit a request via the **Product Registration & License Management Form** to 'Move Licensed Components from one CommServe to another CommServe'. Only a small subset of your licensing needs to be migrated to your pre-production instance.

Related documents: OPS05-BP08 Use multiple environments.

OPS05-BP09 Make frequent, small, reversible changes.

Commvault delivers new quality, performance, and security enhancements monthly via **Maintenance Releases**. Commvault recommends regularly applying Maintenance Releases with minimal scope, to incrementally improve your data management services. Building an organizational skill to apply, assess, and roll out changes is crucial to maintaining a health data management platform.

Related documents: OPS05-BP09 Make frequent, small, reversible changes.

OPS05-BP10 Fully automate integration and deployment.

Commvault **automated software upgrades** can be targeted to a specific set of instances to automate software download and delivery in a staged deployment per environment (i.e., dev-test, sandbox, pre-prod).

Commvault requires that components be upgraded top-down with CommServe, then MediaAgents, then Access Nodes.

Consider automation of deployment and testing tasks in high-impact workloads that drive a large amount of manual effort during integration and upgrade activities.

Related documents: OPS05-BP10 Fully automate integration and deployment.

Mitigate Deployment Risks

OPS06-BP01 Plan for unsuccessful changes

Plan for the ability to return to a known good state after a failed update of any Commvault component, including CommServe, MediaAgents, an Access Nodes. Commvault provides the ability to rapidly roll back from unsuccessful application changes using an **Amazon EC2 Full Instance recovery** from a known good snapshot or backup copy.

Commvault recommends that a **Commvault Disaster Recovery (DR) backup** is taken before any planned platform change. Commvault does not provide the ability to roll back the installation of a **Platform Release** or a **Maintenance Release**, your DR backup will be required to roll back failed database upgrades.

Alternatively, you can **Create an AMI from an Amazon EC2 Instance** for a rapid rollback from a failed CommServe change or upgrade. In high-availability environments, **Maintenance Failovers** can be used to update a CommServe passive instance, while restricting your active instance to restore-only mode.

Related documents: OPS06-BP01 Plan for unsuccessful changes.

OPS06-BP02 Test and validate changes.

Commvault Out of Place Recovery functionality allows you to create temporary copies of your production compute, database, and storage workloads to test changes. Use application copies to test changes on realistic applications and data volumes before deployment to production environments.

Commvault recommends automating the deployment and testing of all critical features and functions to streamline software updates and changes. Changes should be applied via a centralized configuration management system, deployed to temporary test workloads, validate acceptable operations, destroy test workloads, and then flag the change as passed or failed (for rollback).

Changes that test Commvault instances themselves (CommServe, MediaAgents, Access Nodes) can utilize Commvault-published AMIs in the **AWS Marketplace** to streamline instance deployment.

A note on Amazon Marketplace Availability

Commvault software is available for rapid initial deployment from the **Amazon Marketplace** with Commvault and Amazon best practices pre-configured. After initial deployment, Commvault software is patched by automated software download and patch schedules built-into the Commvault Command Center™. New AWS Marketplace images are released for net new deployments but are not used to uplift software versions or perform patching.

Commvault publishes two (2) images which are updated with each platform release (see **Platform Releases**).

- **Commvault Backup & Recovery BYOL**
  Commvault® Backup & Recovery provides an all-in-one instance with CommServe®, MediaAgent, and Access Node software installed. This AMI is recommended as the starting point for dev/test, POC, and Production deployments (Windows 2019, x86_64).
- **Cloud Access Node BYOL**
  Commvault Cloud Access Node BYOL AMI extends an existing Commvault environment by providing a dedicated instance to cost-optimize, store, and transfer your data for recovery, disaster recovery, and migration initiatives. A CommServe® instance is required to utilize the Cloud Access Node for cloud data management (Amazon Linux 2 arm64 and RHEL 7.6 x86_64).

Related documents: OPS06-BP02 Test and validate changes.

OPS06-BP03 Use deployment management systems.

Commvault configuration changes and software updates may be deployed using Continuous Integration/Continuous Deployment (CI/CD) pipelines. Use Commvault **developer tools** and **AWS Systems Manager** to automate change on Amazon EC2 instances running Commvault software.

Related documents: OPS06-BP03 Use deployment management systems.

OPS06-BP04 Test using limited deployments.

Commvault recommends using **All-in-one Commvault** deployments to perform basic functional testing of changes before wider rollout to production. All-in-one CommServe instances deployed from the AWS Marketplace include a 60-day trial license that allows *canary testing* of changes. After initial testing passes, a **blue/green deployment** approach can be used to test alongside production workloads by creating dedicated Access Nodes and/or MediaAgents for the test.
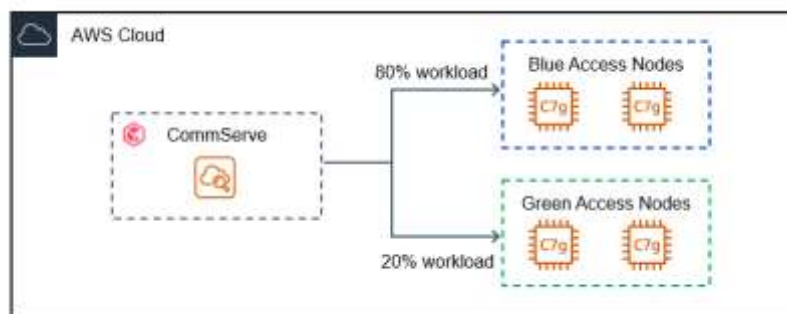
Speak with your Commvault sales representative if you require a long-term test system, your existing licensing rights may be distributed between production and non-production Commvault deployments.

Related documents: OPS06-BP04 Test using limited deployments.

OPS06-BP05 Deploy using parallel environments.

Commvault uses a mutable deployment approach that pushes application and OS updates from a Commvault-maintained **Software Cache**. Changes may be pre-validated in a parallel environment but must be pushed to the existing production environment for widespread rollout.

Alternatively, Commvault software supports immutable infrastructure and can **repave** CommServe® instance, MediaAgents, and Access Nodes in environments where regular infrastructure refresh is required. Speak to your Commvault sales representative about using immutable deployment practices to update Commvault software in AWS.

Related documents: OPS06-BP05 Deploy using parallel environments.

OPS06-BP06 Deploy frequent, small, reversible changes.

Make small, frequent, and reversible changes that are version-controlled to provide change history and permit automated testing, validation, and rollback. This practice can be automated with your CI/CD pipeline using automated operations methodologies like GitOps (see **What is GitOps?** for more information).

Commvault delivers new quality, performance, and security enhancements monthly via **Maintenance Releases**.

Commvault recommends regularly applying Maintenance Releases with minimal scope, to incrementally improve your data management services.

Related documents: OPS06-BP06 Deploy frequent, small, reversible changes.

OPS06-BP07 Fully automate integration and deployment.

Commvault **automated software upgrades** can be targeted to a specific set of instances (server, server groups) to automate staged software upgrades (i.e., dev-test, sandbox, pre-prod).

Instances may be flagged for inclusion in an initial test scope, or particular upgrade phase by applying **entity tags** that group resources in **smart server groups** configured for the upgrade.

Related documents: OPS06-BP07 Fully automate integration and deployment.

OPS06-BP08 Automate testing and rollback.

Commvault automates the recovery or rollback (recreation) of cloud and hybrid applications to a known good state from Amazon snapshots (where available) or Commvault backup copies. Commvault components may also be rolled back using Commvault software or by manual creation and recovery of **Amazon Machine Images (AMIs)**.

Commvault does not recommend automated rollback of Commvault components due to the potential widespread impact on organization-wide data management services.

Related documents: OPS06-BP08 Automate testing and rollback.

Operational Readiness and Change Management

OPS07-BP01 Ensure personnel capability.

Knowing you are ready to 'Go Live!' for a workload involves one or many checklists that validate workload understanding (i.e., business SLA requirements), operational access, operational alerting & response, and personnel readiness. You should ensure that all team members understand the workload, its protection, its recovery, and standard troubleshooting practices.

You can use Commvault **documentation**, **On-Demand Learning Library**, **Certification Program,** and **How to Videos** to continually build awareness and confidence in protecting new workloads. Operational readiness is a living process and must be revisited periodically to ensure best practices and internal and regulatory needs continue to be met. Amazon has multiple tools (Amazon Security Hub, Amazon Inspector, AWS Config) that can monitor the configuration of both protected workloads and the Commvault platform, any identified remediations must be tested and communicated to the operational team(s).

Ensure there are sufficient team members to ensure consistent **SLAs** for business applications. Train and certify team members on operating Commvault software using **Education Advantage** resources. Your service desk will be an invaluable source of information on whether your existing team resources are sufficient to meet business recovery demands.

Related documents: OPS07-BP01 Ensure personnel capability.

OPS07-BP02 Ensure consistent review of operational readiness.

Develop, test, and maintain consistent operational processes for common data management tasks.

Commvault recommends using AWS tags to classify workloads based on business value and the relative impact of a service outage. AWS tags are used by Commvault to automatically discover and protect workloads per business policy. AWS tags can provide critical metadata that allow prescriptive guidance for procedural runbooks or investigative playbooks. You can ensure consistent operational processes by protecting and recovering AWS tags as part of your Commvault Backup & Recovery data management plan for AWS workloads.

For example, a 'business-critical' tag may result in communication to an internal Slack group before commencing service recovery, whereas a 'development' tag may allow immediate recovery from the last known good backup. As you identify hotspots in your environments, tags can be used to <u>automate</u> responses to common events, optimizing and accelerating your response to events that could impact SLAs.

Ensure that data management team members have the time to assess, test, and accept new workloads into operational readiness. At a minimum, operational readiness must include backup checklists, recovery checklists, and validation that RPO/RTO are met per data classification.

Related documents: OPS07-BP02 Ensure a consistent review of operational readiness.

OPS07-BP03 Use runbooks to perform procedures.

Commvault provides detailed procedures for all supported tasks at **docs.commvault.com** and the ability to execute common procedures using **workflows**, **business logic workloads**, or **command line** execution via **AWS Systems Manager Run Command**.

Start with well-documented and tested runbook documentation, as hotspots are identified, trigger runbooks via automation using Commvault **command-line alert notifications** or **Amazon CloudWatch EventBridge**.

Related documents: OPS07-BP03 Use runbooks to perform procedures.

OPS07-BP04 Use playbooks to investigate issues.

Commvault provides detailed procedures for all supported tasks at **docs.commvault.com** for playbook development. Playbooks should be developed to ensure consistent triage and investigative steps for all incidents. Playbooks often have more log analysis and are ideal for automation. You can use Commvault **command line alert notifications** or **Amazon CloudWatch Events** to trigger the collection and presentation of logs, events, and metrics needed to analyze the issue (i.e., create a CloudWatch dashboard with a custom view focused on the issue time period). Any Commvault or CloudWatch event can trigger a playbook and a bespoke playbook data collection process.

Playbooks should ideally be developed and tested in test or pre-production environments, and then re-validated during scheduled **game day** tests of production processes.

Related documents: OPS07-BP04 Use playbooks to investigate issues.

OPS07-BP05 Make informed decisions to deploy systems and changes.

Commvault provides **Risk Analysis** to define, find, and manage critical or sensitive data in your applications. Use findings to assess the risks of deploying new applications. When making high-risk changes, existing **Runbooks**, *business classification,* and *data classification* or an application should be assessed. Runbooks should be enhanced to ensure that the business-value and data classification of a workload recorded on the application as **AWS Resource Tags**, inform the level of risk that is acceptable. As you recover or clone your AWS workloads (i.e., Amazon EC2 instances), Commvault ensures consistent governance can be applied by restoring or customizing your AWS tags on restore.

OPS07-BP06 Create support plans for production workloads.

Commvault recommends that a Commvault support plan is active for all production Commvault instances. Support-level should be equivalent or greater than the Service Level Objectives (SLOs) for the most critical protected application (i.e., if your most critical application requires recovery in 6 hrs., choose a support plan with less than 6 hrs. response time to allow for diagnosis and resolution). Review the **Commvault Support Options** and select the most appropriate service-level for your business.

Ensure that your support coverage, and support contact details are documented for your operational teams, as this information is critical during a recovery event. Ensure operations staff have runbooks that detail how to request support, and ideally what information to collect before logging a customer support issue.

Commvault recommends that production AWS workloads have equivalent support plans active before being moved into Production (consider adding this check to our Operational Readiness Review checklist). Commvault and Amazon (AWS) support teams will work together to resolve reported issues following the support plan active on the account affected.

Related documents: OPS07-BP06 Create support plans for production workloads.

Operate

To operate your Commvault data management platform with continued compliance with business SLAs, the health of the platform must be continually monitored, and issues remediated before affecting service. This requires:

Understanding Workload Health

OPS08-BP01 Identify key performance indicators.

Commvault recommends using a protection Service Level Agreement (SLA) for each of your protected workload types (i.e., business-critical, tier1, tier2, tier3). Data protection success is measured by two primary health metrics or business requirements:

- *Recovery Point Objective* is a measure of the frequency of backup activity or the maximum amount of data loss that is acceptable for the workload.
- *Recovery Time Objective* is a measure of the time it takes to recover the workload to full service.

Commvault reports and monitors your data management KPIs using the **Recovery Readiness Report**. The Recovery Readiness Report displays the configured and achieved RPOs and RTOs for your protected workloads. Workloads that are not meeting their configured SLA may be further investigated with the **SLA and Strikes (failure) report.**

SLA and Strike reporting identifies repeated failures in your Commvault environment, regardless of where a workload resides – the region, AWS Outposts, AWS Local Zones, or on-premises. Drill-down details on workloads that met and missed SLA are possible, along with detailed per-job logs, and the ability to interactively resubmit failed activities.

Commvault recommends KPIs for data management including business-level **Recovery Readiness** and **SLA Health**. Commvault recommends that the second tier of KPIs be used to reflect overall **Commvault Health** and **SLA and Strikes** (per workload). Use Server Groups to report SLAs for individual Resource Groups.

> ⚙ Pro-Tip
>
> It is important to measure both macro-level KPIs and micro-level (per workload) KPIs as a single mission-critical application may not be noticed on an organization-wide report, whereas consistent failures over three days will be highlighted in the SLA and Strikes report.

Related documents: OPS08-BP01 Identify key performance indicators.

OPS08-BP02 Define workload metrics.

Typically, the percentage of workloads that **met or missed SLA** is the most appropriate metric for validating acceptable service delivery for the Commvault workload. Commvault presents the achieved SLA and success/failure of Jobs in the last 24 hours in the **Overview dashboard**. Achieved SLA is also reflected in the **Recovery Readiness** and **SLA Health** reports.

Assessing the achieved SLA at a per-application level may be achieved using the Recovery Readiness report, and detail on the number of failures may be found in the **SLA and Strikes** report. Detailed Job history (backup, restore) provides workload level metrics including success/failure, number of warnings, number of errors, and throughput achieved.

Commvault **Metrics Reporting** records historical backup and recovery metrics across one or many CommServe® instances allowing establishing of baselines and identifying anomalous protection patterns.

Workload metrics should be assessed from the top down:

- Is the organizational SLA healthy?

- Is my workload compliant with business SLA?

- How long has my workload been non-compliant? (strikes)

- How has my workload performed in the past? (success, failure, throughput)

Commvault supports building **custom reports** and accessing report data via export and/or **REST API** for unique KPI development.

Commvault recommends that granular per-job logs, including total elapsed time, data transferred, and data written, are forwarded to Amazon CloudWatch and key metrics extracted and tracked as workload metrics in Amazon CloudWatch dashboards.

Related documents: OPS08-BP02 Define workload metrics.

OPS08-BP03 Collect and analyze workload metrics.

Commvault recommends using business-level **SLA health**, **Service Level, and Strikes** reporting to identify systemic issues, then using detailed telemetry to identify the root cause and resolve systemic or isolated incidents. Commvault simplifies the analysis of missed and met SLAs over time using the **SLA Trend Report**. Analysis can filter based on resource groups (subclient, server group).

Commvault monitors workload SLAs and platform health proactively and uses **alerts and notifications** that alarm on conditions requiring attention. Commvault recommends sending Commvault alerts to Amazon CloudWatch as **custom metrics** for centralized observability and alarming.

Commvault recommends developing runbooks that integrate Amazon CloudWatch instance metrics, **Job Summary Reports**, **Infrastructure Load Reports**, and custom metrics to analyze workload health and issues. Amazon CloudWatch Logs Insights can also be used for Commvault Log Files that have been forwarded to **CloudWatch**.

Related documents: OPS08-BP03 Collect and analyze workload metrics.

OPS08-BP04 Establish workload metrics baselines.

Commvault Reports provides the ability to visualize baselines for **Infrastructure Load**, **Storage Utilization**, **SLA**, **Growth and Trends**, and many others from the Commvault **report store**.

Alternatively, infrastructure and data management metrics extracted from Log Files may be observed and baselined in **Amazon CloudWatch Metrics**. When CloudWatch is monitoring metrics extracted from logs, **CloudWatch anomaly detection** will use statistical and machine-learning algorithms to identify and optionally alarm when metrics anomalies occur. Anomaly detection can be performed on any metric tracked in Amazon CloudWatch including CPU utilization, memory utilization, network utilization, average throughput, RPO, RTO, and so on.

Related documents: OPS08-BP04 Establish workload metrics baselines.

Establish and understand the patterns of normal application behavior and anomalous behavior in your data management platform. Commvault provides **unusual file activity and ransomware detection** on active clients and backup jobs. **CommServe Anomaly Alerts** identify anomalies in the Commvault data management system across infrastructure, events, and jobs.

Alternatively, **CloudWatch anomaly detection** can perform *anomaly detection* for metrics extracted from CloudWatch and Commvault-published logs, events, and metrics.

> 🔅 **Pro-Tip**
>
> Your historical logs, events, and metrics are critical resources to effectively operating your workloads for consistent data protection and recovery. Be sure to review the export and backup of important CloudWatch data to Commvault-protected Amazon S3 bucket(s).

Related documents: OPS08-BP05 Learn expected patterns of activity for workload.

Commvault provides the **CommServe anomaly digest** alert which proactively generates notifications for workloads that may miss their SLA or have missed SLA. **Commvault anomaly digest** alerts identify workloads at risk or already missing SLA and may be integrated by Amazon CloudWatch for centralized metrics and logging.

Commvault recommends publishing Commvault anomaly alerts to CloudWatch via **custom metrics publishing**.

Related documents: OPS08-BP06 Alert when workload outcomes are at risk.

Commvault provides the **CommServe anomaly digest** alert which proactively generates notifications for workloads that may miss their SLA or have missed SLA.

Commvault **Predefined Alerts** identify anomalies in job counts, job completion time, event type counts, and data pruning performance by default. Additionally, the total number of created, deleted, and modified files are detected as potential file activity anomalies which are indicative of ransomware or malware activity.

Commvault is actively analyzing, and alerting anomalies detected across all data management activities. Alternatively, **Alert Notification Types** can be used to push alerts into Amazon CloudWatch using **Command Line**, **Windows Event Log** integration, or **Commvault Workflow** automation. Once ingested, **Amazon CloudWatch Anomaly Detection**, alerting, and event-based alarms and automation may be utilized.

Related documents: OPS08-BP07 Alert when workload anomalies are detected.

Commvault provides a business-level **Executive Summary Presentation** that rolls up SLA success and failure and backup retention growth over the past 12 months. Use the Executive Summary Presentation, and the **Recovery Readiness Report** to report the achievement of business-required protection SLAs.

Alternatively, use **Amazon CloudWatch dashboards** to create aggregated views of business-level and resource-level metrics that represent the achievement of business and technical KPIs.

> ⓘ **Note**
>
> SLA-based percentages reflected in the Executive Summary Presentation, Recovery Readiness Report must be exported or retrieved using **Reports REST API** and forwarded to CloudWatch for visualization.

Related documents: OPS08-BP08 Validate the achievement of outcomes and the effectiveness of KPIs and metrics

Understanding Operational Health

OPS09-BP01 Identify key performance indicators.

Operational health refers to how performant and available the Commvault data platform is when measured by the end user. Key performance indicators (KPIs) for operational health should be established and measured to ensure acceptable service performance.

At a minimum, Operational health should include:
- Infrastructure availability.
- Infrastructure downtime (including and excluding planned maintenance events).
- The total number of customer support cases.
- Service level achievement for protected workloads (**Recovery Readiness**).

Commvault recommends working with business leaders to agree on KPIs for shared data management services. This may include new service availability, performance, feature cadence, support incidents per month, or the number of met/missed SLA workloads.

Related documents: OPS09-BP01 Identify key performance indicators.

OPS09-BP02 Define operations metrics.

Operations metrics are concerned with the efficiency of the shared data management service and whether it is meeting business KPIs to reduce service outage risks. Operational health includes the user's ability to consume the service to restore their workloads in either a self-service or service desk-assisted recovery.

Acceptable downtime will be a key indicator in the availability design and cost of the overall platform. Customer support cases should be triaged periodically to identify areas for improvement both technical and non-technical (e.g., customer education, new self-service feature issues, software defects). The number and duration of customer incidents should be reported regularly against agreed operations metrics (i.e., less than NN incidents per measurement period).

SLAs with the business should be established that identify the required data management platform availability and acceptable outage (planned and unplanned), and incident response and fix times.

Operational metrics should also consider governance and compliance responsibilities. Compliance with internal standards, external regulatory requirements, and assessed and mitigated threats should be tracked

Commvault provides **Certifications and Compliance** standards with which Commvault confirms. AWS provides **Compliance Resources** to help generate proof of compliance documentation.

Related documents: OPS09-BP02 Define operations metrics.

OPS09-BP03 Collect and analyze operations metrics.

Commvault recommends **sending operational log files to Amazon CloudWatch** so that trends and hotspots may be identified across your application landscape.

Operational CloudWatch dashboards should be created that reflect infrastructure and application availability and achieved data protection SLAs across AWS and edge locations. Metrics relating to the volume of health-based alarms, customer incidents, and compliance breaches should be created and visualized for operational reviews.

Related documents: OPS09-BP03 Collect and analyze operations metrics.

OPS09-BP04 Establish operations metrics baselines.

Commvault Reports provides the ability to visualize baselines for **Infrastructure Load**, **Storage Utilization**, **SLA**, **Growth and Trends**, and many others from the Commvault report store.

Alternatively, infrastructure and data management metrics extracted from Log Files may be observed and baselined in **Amazon CloudWatch Metrics**. For metrics tracked by CloudWatch, **CloudWatch anomaly detection** automatically calculates baselines and **alarms** on the identification of anomalies.

Related documents: OPS09-BP04 Establish operations metrics baselines.

OPS09-BP05 Learn the expected patterns of activity for operations.

Commvault will automatically baseline the typical volume and type of events that occur daily. If an anomalous event or events occurs, you will be notified (See **CommServe Anomaly Alert**).

You can learn about normal and anomalous activities in Commvault within the **Anomaly Dashboard**.

Patterns of activity may involve periodic blackout windows, highly active release or trading months, or periods of inactivity due to regional holidays and employee travel. Maintain a shared calendar of the established system and business patterns to help pre-plan for anticipated peaks or drops in activity.

Establishing a shared repository of activity patterns helps onboard and educate new team members and work better with cross-functional teams in the business.

Related documents: OPS09-BP05 Learn the expected patterns of activity for operations.

OPS09-BP06 Alert when operations outcomes are at risk.

Commvault provides the **CommServe anomaly digest** alert which proactively generates notifications for workloads that may miss their SLA or have missed SLA.

Related documents: OPS09-BP06 Alert when operations outcomes are at risk.

OPS09-BP07 Alert when operations anomalies are detected.

Commvault provides the **CommServe anomaly digest** alert which proactively generates notifications for workloads that may miss their SLA or have missed SLA. Anomalies are identified for workloads that may miss SLA and a host of other conditions observed from Commvault-established baselines.

Consider forwarding anomaly alerts to CloudWatch and publishing them as **custom metrics** for centralized visualization and alarming.

Related documents: OPS09-BP07 Alert when operations anomalies are detected.

OPS09-BP08 Validate the achievement of outcomes and the effectiveness of KPIs and metrics

Commvault provides a business-level **Executive Summary Presentation** that rolls up SLA success and failure and backup retention growth over the past 12 months. Use the Executive Summary Presentation and **Recovery Readiness** Reports to review the achievement of business recovery readiness across your organization.

Hybrid operational reporting including service availability submitted and resolved incidents, and compliance and governance can be built within **Amazon CloudWatch Dashboards** if sufficient log, event, and metrics data have been published to CloudWatch.

Related documents: OPS09-BP08 Validate the achievement of outcomes and the effectiveness of KPIs and metrics.

OPS10-BP01 Use processes for event, incident, and problem management.

Responding to planned and unplanned events impacting your data management platform is a key requirement to staying *Recovery Ready*. Commonly anticipated events (capacity exhaustion, failed backup/restore) should have procedural runbooks defined and tested in pre-production to allow a rapid, repeatable response.

Unanticipated events should have investigative playbooks to gather logs, metrics, and event history and analyze for the root cause. If an event occurs more than once, a runbook should be developed and communicated to operations personnel.

Using the data classification or business importance AWS tags attached to all workloads, events that impact mission-critical business workloads should have an automated escalated notification response (i.e., Slack, Teams, or SMS message to on-call resources).

Commvault **Workflows** can be used to automate event response, including enriching the event with additional metadata before pushing to Amazon CloudWatch, or executing further data gathering using **Amazon Systems Manager**.

Commvault recommends developing documented processes to respond to data management events, incidents, and repeated problems, before attempting automation. Use **docs.commvault.com**, **maintenance advantage**, and **knowledgebase** to develop processes.

Related documents: OPS10-BP01 Use a process for event, incident, and problem management.

OPS10-BP02 Have a process per alert.

Commvault provides **predefined alerts** for common events that you may encounter. **Custom alerts** may also be developed based on business needs. Alerts can include **Alert Tokens** to enrich and enable timely resolution of the alert condition.

Define a process or **runbook** per configured alert.

Define generic **playbooks** for collecting minimal diagnostics data for triage and how to resolve or escalate for assistance.

Ensure runbooks and playbooks are in a shared secure area that all team members can access.

Related documents: OPS10-BP02 Have a process per alert.

OPS10-BP03 Prioritize operational events based on business impact.

Commvault recommends using **tags** to indicate the data classification and/or business importance for all workloads. Use business classification to prioritize the order to process operational events. Existing operational procedures captured in **runbooks** and **playbooks** should clearly articulate how to prioritize issue resolution using business importance or business-value tags. You can use Commvault to restore tags as part of an AWS workload recovery or customize tags to reflect a different data classification for say a newly restored or cloned Amazon EC2 instance.

Related documents: OPS10-BP03 Prioritize operational events based on business impact.

OPS10-BP04 Define escalation paths.

Commvault will automatically **Escalate a Notification** that is not resolved during the initial notification period. Commvault **webhooks** can be used to push notifications directly to operational and management teams via Splunk and Teams (as an example).

Alternatively, pushing Commvault alerts into Amazon CloudWatch allows reusing centralized alerting and notification built on Amazon Simple Queue Service (SQS) and Amazon Simple Notification Service (SNS).

Ensure that operational processes include an escalation path for ensuring the timely resolution of incidents. Notifications can be configured to **automatically escalate** if remaining unresolved. Ensure escalation processes and business-facing service documentation clearly articulates the hours of support (24x7, 9x5) and escalation contacts.

Do not forget to track dependencies between workload sub-components. Consider using Commvault **Smart Server Groups**, and **AWS Resource Groups** to group workloads to better analyze and communicate the true organizational impact of escalations.

Related documents: OPS10-BP04 Define escalation paths.

OPS10-BP05 Enable push notifications.

Commvault **Alerts and Notifications** push notifications directly to users via email, SNMP, event viewer, command-line script, or RSS feed. Alerts can also use a **Webhook** to push messages to chat platforms like Microsoft Teams and Slack.

Alternatively, Commvault Alerts and Notifications can be published to Amazon CloudWatch as **custom metrics**, and then use **Amazon CloudWatch Alarms** to push into Amazon SNS, and Amazon SQS to align with your organization's event management practices.

Related documents: OPS10-BP05 Define a customer communication plan for outages.

OPS10-BP06 Communicate status through dashboards.

Commvault provides built-in **Dashboards** for common workloads and supports building **custom dashboards** with information available in the Commvault database.

Alternatively, **CloudWatch dashboards** can be used to aggregate workload and Commvault insights for logs, events, and metrics forwarded to CloudWatch.

Related documents: OPS10-BP06 Communicate status through dashboards.

OPS10-BP07 Automate responses to events.

Commvault **Alerts and Notifications** can be configured to run a **command-line** script or **workflow** to response to the event. Additionally, **CloudWatch Events** can be used to trigger an automated action and execute a script with **AWS Systems Manager Run Command**.

Related documents: OPS10-BP07 Automate responses to events.

Evolve

The one constant in all cloud and IT systems is change. To maintain consistent data management service levels, a cycle of continuous improvement is mandatory. This includes:

Learn, Share, and Improve

OPS11-BP01 Have a process for continuous improvement.

A culture of continual learning in a *data-driven approach* is key to being ready for unanticipated events that affect the business. A post-incident review (PIR) should be performed for any event that has a direct impact on business service, workload, or end-user (see **Post-incident analysis**).

Commvault recommends a process of continuous improvement is implemented to maintain SLAs, cost, and performance of your data management platform.

Related documents: OPS11-BP01 Have a process for continuous improvement.

OPS11-BP02 Perform post-incident analysis.

Post-incident reviews should include updates to procedural runbooks, investigative playbooks, and gathering the 'voice of the customer' feedback from affected teams and/or users. Learnings from the PIR should be written down and distributed via a Knowledge Management System (KMS) and may include updates to operational support processes, customer-facing processes, or knowledgebase articles.

Commvault recommends performing post-incident reviews to identify opportunities to update runbooks, playbooks, or **engage Commvault** to identify a best-practice resolution.

Related documents: OPS11-BP02 Perform post-incident analysis.

OPS11-BP03 Implement feedback loops.

Commvault recommends creating a capability for customers to submit feedback on data management services offered. Ensure feedback is reviewed and responded to promptly.

Related documents: OPS11-BP03 Implement feedback loops.

OPS11-BP04 Perform knowledge management.

Commvault recommends providing customer-facing and internal team-facing information on the offered data management services and how to operate commvault self-service features.

Consider placing information in a secure shareable service like:

- Amazon EFS
- Amazon FSx
- Amazon FSx for NetApp ONTAL
- Microsoft SharePoint Online

> 💡 **Pro-Tip**
>
> Commvault protects centralized shared content located in Amazon EFS, Amazon FSx for Windows, Amazon FSx for NetApp, and **Microsoft SharePoint Online**.

> 💡 **Pro-Tip**
>
> Consider where your critical operational processes are stored and any dependency on other systems (identity & access management, DNS, key management/CloudHSM services). In the event of a system-wide failure, your operational runbooks and playbooks must be accessible to operations personnel.

Related documents: OPS11-BP04 Perform knowledge management.

OPS11-BP05 Define drivers for improvement.

Data classification or business-value tags on affected workloads should be used to direct further investment in removing or optimizing data management hotspots. Learning insights and relevant impact on operational metrics and reporting should be communicated broadly to affected parties. Communication can occur formally via quarterly business reviews (QBRs) or informally in communities of interest (e.g., lunch and learn sessions).

Commvault recommends working with business stakeholders to understand business needs and priorities. Improvements range from cost savings, risk reduction, performance improvements, or adding new features.

Related documents: OPS11-BP05 Define drivers for improvement.

OPS11-BP06 Validate insights.

Review the improvement initiatives identified through data-driven insight (i.e., Amazon CloudWatch Insights) with business stakeholders to understand findings and better understand the impact.

Related documents: OPS11-BP06 Validate insights.

OPS11-BP07 Perform operations metrics reviews.

Commvault recommends reviewing SLA reports and Infrastructure-level reports with cross-functional teams to share lessons learned and identify improvement initiatives.

Related documents: OPS11-BP07 Perform operations metrics reviews.

OPS11-BP08 Document and share lessons learned.

Document and detail learnings from incidents and problem resolution so that the knowledge transitions from **tribal** knowledge to **explicit** knowledge.

Related documents: OPS11-BP08 Document and share lessons learned.

OPS11-BP09 Allocate time to make improvements.

Most importantly, dedicated time and resources must be available for continuous improvement. Potential improvements should be investigated in on-demand pre-production environments that leverage the elastic on-demand nature of Amazon EC2 to duplicate targeted use cases, and then terminate resources after testing is complete.

Commvault recommends time is reserved for team members to continually review, assess, and improve your data management platform.

Related documents: OPS11-BP09 Allocate time to make improvements.

Additional Resources

- **Operational Excellence Pillar**.
- **Monitor System and Commvault Resource Logs**.
- **Creating Incidents on ServiceNow**.
- **Commvault Terraform Module**.
- **AWS re:Invent 2021- Get insights from operational metrics at scale with CloudWatch Metrics Insights**.
- **How to Easily Setup Application Monitoring for Your AWS Workloads - AWS Online Tech Talks**.
- **Monitoring Performance with CloudWatch Dashboards - AWS Virtual Workshop**.
- **AWS re:Invent 2021 - Ready, set, operate: The AWS Cloud operations model**.

Security Pillar

The *security posture* of your Commvault data management platform is paramount to ensuring you can protect and recover your organization from unplanned data loss events. The **Security Pillar** guides the architecting, designing, and operating of your Commvault software to meet current AWS, Commvault, and industry recommendations and strategies for a secure cloud.

Security foundations

AWS account management and separation

SEC01-BP01 Separate workloads using accounts.

Commvault recommends account-level separation between production, development, and test environments. Additionally, Commvault backup data and compute instances should be segregated in a centralized **service account**, separate from protected workloads or accounts. Use **AWS Organizations** to centrally govern and enforce permission policies and boundaries. Use **Service Control Policies (SCPs)** to set and enforce permission guardrails for all organization accounts.

> ⚠ **Important**
>
> Ensure that Service Control Policies (SCPs) grant the required permissions to perform Commvault Backup & Recovery per Permissions. Failure to grant required IAM actions at both the IAM Role and SCP will result in data management activities failing.

> 💡 **Pro-Tip**
>
> Commvault utilizes tags to record, and cleanup resources created during backup, recovery, and replication activities. When using mandatory tagging SCPs, be sure to add Commvault tags (**tag syntax**) or Commvault will be unable to provide backup and restore services.

All changes to the Production environment must be validated in the Non-Production (Pre-Production) environment first. While Commvault does support cross-account restores, the ability for restores between Production and Non-Production logical boundaries must be tightly controlled to a small number of authorized users. Consider implementing **Commvault Business Logic Workflows** to provide an active approval process for any data restoration across environment boundaries. Consider enforcing **data masking** as an organizational guardrail to prevent accidental data leakage or exposure.

Related documents: SEC01-BP01 Separate workloads using accounts.

SEC01-BP02 Secure account root user and properties.

Commvault does not require or recommend the use of the AWS account 'root' user to perform backup & recovery of AWS workloads. Commvault publishes a set of least-privilege **AWS IAM policies** to ensure that a least-privilege approach to maintained for Commvault-configured identities. Commvault recommends that the AWS root user account has Multi-Factor Authentication (MFA) enabled per AWS best practices.

Additionally, Commvault creates an 'admin' user during initial setup, this user should be considered a **break glass account** and any usage should be alerted to Security Operations (SecOps) platforms for investigation and credential rotation. Commvault recommends that both the admin user and any administrate user or user group configured within Commvault and associated with the 'master' role, have **Two-Factor Authentication (2FA)** enabled within Commvault.

Do not use the AWS root user for any Commvault data management activities.

Related documents: SEC01-BP02 Secure account root user and properties.

Operating your workloads securely

SEC01-BP03 Identify and validate control objectives.

Create and curate a **threat model** that identifies core risks and compliance requirements for your workloads. Work with legal, governance risk & compliance to identify internal and regulatory requirements and **certifications** for your data management system. Identify the controls that will be applied to your Commvault data management platform.

Related documents: SEC01-BP03 Identify and validate control objectives.

SEC01-BP04 Keep up-to-date with security threats.

Commvault recommends subscribing to the **Common Vulnerabilities and Exposures List**. Review the Commvault **Security Vulnerability and Reporting** advisories for relevant threats and mitigations.

Related documents: SEC01-BP04 Keep up-to-date with security threats.

SEC01-BP05 Keep up-to-date with security recommendations.

Commvault recommends subscribing to Critical Alert Messages and Maintenance Release Alerts at **Maintenance Advantage**. Commvault will proactively send recommendations based on configured frequency. Commvault also publishes the AWS Cloud Architecture Guide (this document) annually with the latest security recommendations.

Related documents: SEC01-BP05 Keep up-to-date with security recommendations.

SEC01-BP06 Automate testing and validation of security controls in pipelines.

Commvault infrastructure provisioning may be automated using **AWS CloudFormation** and **Terraform**. Validate and test security configurations using **AWS Systems Manager** and report fleetwide compliance using AWS Config (i.e., Commvault host-based firewall configuration).

Related documents: SEC01-BP06 Automate testing and validation of security controls in pipelines.

SEC01-BP07 Identify and prioritize risks using a threat model.

Commvault recommends creating and maintaining a prioritized **threat model** that identifies internal and external actors, threats, and active controls or mitigations.

Commvault has protections against **insider threats** and external threats like **ransomware or malware infection**.

Related documents: SEC01-BP07 Identify threats and prioritize mitigations using a threat model.

SEC01-BP08 Evaluate and implement new security services and features regularly.

Commvault recommends regular review of **AWS What's New**, **AWS Security Blog**, and **Commvault What's New** for new security enhancements and controls to enable.

Related documents: SEC01-BP08 Evaluate and implement new security services and features regularly.

Identity and Access Management

Running your workloads on AWS requires that you provide your users and applications access to your AWS resources. Identities may be administrators and end-users (**human identities**) or AWS infrastructure or services (**machine identities**) that request and require access to data in your AWS accounts. To secure your Commvault data management platform in a multi-account multi-region deployment, the following factors must be considered:

Identity management

Commvault enforces strong **password complexity**, enforces **two-factor authentication** (2FA), and Single Sign On (SSO) via **SAML**, **OIDC**, and **Active Directory**.

Commvault uses Secure Token Service (STS) **AssumeRole** to request and use temporary credentials to perform data management activities, removing the requirement to rotate credentials.

Commvault integrates with **CyberArk Password Security Platform** to securely manage Commvault and application credentials.

Commvault enforces **strong passwords** when being used as the identity store. Additionally, Commvault supports multi-factor authentication (MFA) using a physical **FIDO2-compliant web authenticator** (such as Yubikey from Yubico) to authenticate to web-based administrative interfaces. Commvault also supports username and password-less authentication using **Common Access Card (CAC)** authentication to Commvault web-based administrative interfaces. Additionally, Commvault supports two-factor authentication (2FA) using a six-digit PIN provided by a **Time-based One-time Password (TOTP)** algorithm, as detailed in RFC 6238. Commvault supports the use of **Google Authenticator**, **Microsoft Authenticator**, and **Commvault Token Desktop applications** to generate authentication pins.

Interface	MFA Key	CAC	2FA PIN
Commvault Command Center™	●	●	●
Web Console	●	●	●
CommCell Console	●	●	●
Mobile apps (Edge, NOW)	●	●	●
Command line			●
Commvault **REST API**			●

Command Center, Web Console, and CommCell Console use SAML-based logins to leverage the MFA key and CAC sign-in.

Source: **Two-Factor Authentication for Your CommCell Environment**

When using Commvault with an external identity provider, configure authentication rules to enforce the use of multi-factor authentication (MFA) for all users.

Related documents: SEC02-BP01 Use strong sign-in mechanisms.

Commvault uses Secure Token Service (STS) **AssumeRole** to request and use **temporary credentials** to perform data management activities for AWS compute, cloud databases, and cloud storage.

Credentials use a **least privilege approach** that grants permission only for the data types and **use cases** required.

Permission boundaries may be created using **condition keys** that grant access only to resources with a policy-specified **AWS Resource Tag**.

Commvault uses temporary credentials for workforce users when using an external identity provider with SAML 2.0 or OpenID Connect (OICD). Workforce user credentials or authentication tokens will periodically expire and transparently re-authenticate.

Machine-based identities may utilize **AWS IAM Security Token Service (STS) AssumeRole** to request and refresh temporary credentials used to perform cross-account data management activities. Commvault supports STS:AssumeRole for accessing and protecting Amazon EC2, Amazon EBS, Amazon RDS, Amazon Redshift, Amazon DynamoDB, Amazon DocumentDB, and Amazon S3 workloads. Commvault supplies customer-managed IAM policies which are attached to Commvault Amazon EC2 Cloud Access Nodes (see **Instance Profiles**) to perform protection.

See **Amazon Web Services User Permissions for Backups and Restores** for the Commvault required IAM policies and permissions required to protect each AWS product. Commvault recommends utilizing IAM roles (with or without STS:AssumeRole) attached to Amazon EC2 infrastructure exclusively. Commvault software is built on the latest AWS SDK and utilizes Instance Metadata Service v2 (IMDSv2) to create, distribute, and rotate temporary security credentials granted via IAM roles attached to EC2 instances.

> ⓘ Note
>
> Commvault does not require IMDSv1 to be enabled or functioning to perform data management. IMDSv1 can be safely disabled without any impact on the Commvault software function.

Commvault uses IAM permissions to securely access EC2 instances using **AWS Systems Manager**, without the need for storing and transferring usernames and passwords.

Commvault supports but <u>does not recommend</u> the use of AWS IAM users or long-term access keys.

Related documents: SEC02-BP02 Use temporary credentials.

SEC02-BP03 Store and use secrets securely.

Commvault provides a secure encrypted **credential manager** for cloud, database, and storage credentials that require secrets such as passwords or access keys.

Additionally, Commvault can integrate with **CyberArk** for secure storage and management of secrets remotely.

Commvault does not integrate with AWS Secrets Manager at the time of writing.

Commvault stores any non-IAM-related credentials within the Commvault database inside the encrypted **Credential Manager** store. All credentials are encrypted using a FIPS-140-3 compliant ^{pending CMVP review} crypto library and with key rotation performed periodically by Commvault. Access to credentials in plaintext is not possible, the role-based access controls (RBAC) within Commvault dictate which users can utilize the persisted credentials.

All operations on stored credentials are journaled to the **Commvault Audit Trail**. Commvault recommends **installing the Amazon CloudWatch agent** and pushing audit log events to CloudWatch for centralized security incident & event management.

Related documents: SEC02-BP03 Store and use secrets securely.

SEC02-BP04 Rely on a centralized identity provider.

Workforce identities should be managed by a centralized identity provider so that creation, updates, and deletion activities occur from a single location and action.

Commvault supports centralized identity providers (IdPs) exposed via **OpenID Connect (OIDC)**, **SAML 2.0 (Security Assertion Markup Language 2.0)**, or **Active Directory**.

Commvault can utilize AWS Directory Services as the central identity store for Commvault users and group authentication and authorization.

Commvault can function as a single identity store or centralized identity provider (IdP) for all users and administrators that require access to Commvault. Alternatively, you can use any identity provider that supports the SAML 2.0 protocol to perform a web-based Single Sign On (SSO) to Commvault. Any SAML 2.0 iDP may be used, such as **AWS IAM Identity Center**, **AWS Directory Service**, **Okta**, **Auth0**, and **Ping Identity**.

After successful authentication, Commvault performs authentication using the username and optionally the user group(s) supplied in the identity provider response (see **Mapping SAML Attributes**). Commvault stores usernames, groups, and authorized permissions in its internal database to perform authorization. Commvault does <u>not</u> store user passwords for SAML-authenticated users.

Alternatively, Commvault can utilize **OpenID Connect (OIDC)**, **Active Directory**, and **Central Authentication Service (CAS)** to perform centralized authentication for workforce users.

Commvault Web Service	Single Sign-On (SSO) Support		
	SAML 2.0	OIDC	Active Directory
Commvault Command Center™ and **REST API**	●	●	●
Web Console (Tomcat) and **REST API** (remote)	●	●	●
CommCell Console	●	●	●
Command line	●	●	●
Mobile apps (**Edge**, **NOW**)	●	●	●
Commvault **REST API**	Local Commvault account authentication only.		

> ⚠ **Warning**
>
> SSO for Web Consoles on Linux computers is not supported.

Related documents: SEC02-BP04 Rely on a centralized identity provider.

SEC02-BP05 Audit and rotate credentials periodically

Commvault supports mandatory password aging for **users** and **user groups**. **Two-factor authentication (2FA)** can be enforced for human identities accessing Commvault, or just administrative users.

With **CyberArk integration**, you can synchronously rotate account passwords so that backups continue after credential rotation.

For AWS accounts and roles used by Commvault, use **instance profiles** attached to Amazon EC2 instances, which removes the requirement to rotate credentials.

User and User Group Permissions Report may be used to periodically audit the configured users, groups, and assigned permissions.

Commvault enforces **password age in days timers** and forces users to reset passwords that are stored in the Commvault Database (CSDB). When using a centralized identity store, the only user account(s) with passwords are

emergency administrative users. Commvault does <u>not</u> alert the user or enforce periodic rotation of AWS IAM access keys and secret keys. Consider the **Automatic rotation of IAM user access keys with AWS Organizations pattern** to automate access key and secret rotation in AWS. The **Commvault REST API** can be used to rotate the access key and secret keys stored in Commvault Credential Manager (e.g., **POST Change Hypervisor Client Credentials**) after rotation by AWS Organizations.

Related documents: SEC02-BP05 Audit and rotate credentials periodically.

SEC02-BP06 Leverage user groups and attributes

Commvault provides **user groups** to place users with similar security needs in common groups. Groups may be passed by the centralized identity provider as **attributes** to perform authorization via group membership for authenticated users.

Commvault software supports and recommends using **user groups** and **external user groups** (Active Directory groups) to group users and teams that require common data management permissions. External users may be assigned to user groups via user attributes that are returned by the centralized identity provider response.

User groups are associated with **roles** (or permissions) and one or many Commvault entities, this is referred to as a **security association** within Commvault software. User groups within Commvault also allow managing **user quotas**, granting access to **administrative consoles**, and granular **per-workload access rights**.

Commvault does not publish resource-based policies, permissions boundaries policies, attribute-based access control (ABAC), AWS Organizations service control policies (SCPs), or session policies but promotes their use in a multi-layer permissions management approach. Care should be taken when implementing AWS Organizations Service Control Policies (SCPs) to ensure the mandatory Commvault IAM permissions are maintained and not inadvertently denied.

See **Permissions Management** for a definition of the various methods of granting access to AWS resources.

Related documents: SEC02-BP06 Leverage user groups and attributes.

Permissions Management

SEC03-BP01 Define access requirements.

Commvault recommends using AWS Organizations **service control policies (SCPs)** to provision and enforce Commvault-required **IAM roles** that allow protection by the centralized Commvault service account.

Commvault software requires AWS IAM permissions to access AWS products for protection and recovery. Commvault publishes a list of **required IAM permissions by AWS product** and detailed information on **how each IAM permission is used**. Commvault identity-based policies are *customer-managed policies* and are the best practice for providing consistent, repeatable, secure data protection of your AWS resources.

> ☀ Pro-Tip
>
> **Commvault IAM policies** are published in the Commvault documentation and must be reviewed before performing a Commvault platform release to ensure permissions have not been added or removed.

Commvault recommends creating Managed policies (vs. inline policies) for each AWS product or service to protect and then attaching them to a Commvault Backup & Recovery role that is associated with each of your Amazon EC2-based Commvault Access Node instances. This approach allows granular management of permissions on a per-service basis. Roles should follow the principle of **least privilege** and any associate permissions for services that require protection.

Commvault software utilizes **Condition operators** to restrict the scope of granted permissions to resources that contain either Resource Access tags (**ec2:ResourceTag**) or AWS global condition context keys or **TagKeys** that are applied by prior Commvault data management activities.

Commvault recommends using **AWS Organizations – Tag policies** to ensure mandatory data protection resource tags are applied to all resources that require protection. Commvault IAM policies may then be enhanced to utilize condition operators that grant protection permissions only when specified resource tags are present on the resource.

Related documents: SEC03-BP01 Define access requirements.

SEC03-BP02 Grant least privilege access.

Commvault supports a least privilege approach to defining the IAM role required for protection. Policies need only implement the permissions for the **AWS services** and **use cases** required for protection.

Consider using **AWS resource tags** to implement permission boundaries on in-scope resources.

> ⚠ Important
>
> Commvault <u>does not support</u> removing permissions from Commvault-supplied IAM policies unless the entire use-case is being removed, per **Amazon Web Services Permission Usage** (e.g., all permissions for 'Agentless file recovery can be removed, individual permissions within Agentless file recovery cannot be removed).

Commvault uses the principle of **least privilege** when developing and publishing the required IAM policies required to protect each AWS resource. Commvault IAM policies detail the service actions, resources, and conditions that must be true to grant permission. Commvault recommends using AWS Resource Tags and Condition Filters to limit which AWS resources are accessible by Commvault software. Additionally, **Account Factory** can optionally filter or remove IAM policies for data types that will not be used in the target AWS account, through attribute-based access controls (ABAC) at the account level.

Commvault recommends auditing permissions usage with **IAM Access Analyzer** periodically to right-size IAM policies and service actions to only the permissions being actively used and observed via AWS CloudTrail logs. Caution should be taken before reducing permissions, as the removal of a critical service action may impact a future unplanned recovery event.

If in doubt, consult the **Amazon Web Services Permission Usage page** for details and Commvault use cases that require each permission. Additionally, ensure the scope of permission reduction is restricted only to the AWS account demonstrating reduced permissions usage, as a change to an organization-wide IAM policy could have a significant recoverability impact.

Commvault recommends the use of Security Token Service (STS) `AssumeRole` authentication and provides **instructions on configuring STS:AssumeRole** in multi-account protection environments. See the **Amazon Security Token Service (STS) AssumeRole Activation Guide** for additional details.

Related documents: SEC03-BP02 Grant least privilege access.

SEC03-BP03 Establish emergency access process.

Commvault recommends that a per-account alternate elevated permission set be available for the data management administrator to use for problem resolution.

Commvault has a built-in administrative account called 'admin' that provides authenticated access to administer the Commvault software in the event of externalized authentication failures. A process to grant temporary Commvault-local user accounts to IT administrative staff in the event of systemic authentication failures should be established.

This will allow administrators to ensure business protection SLAs are met while troubleshooting and resolution of authentication issues occur.

Related documents: SEC03-BP03 Establish emergency access process.

SEC03-BP04 Reduce permissions continuously.

Consider using AWS Identity and Access Management (IAM) Access Analyzer to identify IAM actions that are not being used and remove them from the granted permission set.

> ⚠ **Important Warning**
>
> Removing permissions that have not been used during the review period may impact future backup or restore attempts if that data type starts being used within the target AWS account.

Commvault recommends reviewing granted IAM permissions periodically to confirm privileges are still required. AWS IAM provides last-accessed timestamps to identify **unused users and roles**. More importantly, the individual service actions within a role may be audited for last accessed or last used timestamps to identify unused service actions. Reviews can be performed initially in the AWS console using *runbooks* but may also be implemented programmatically when the number of accounts justifies the investment in automation.

It should be noted that Commvault already restricts the permissions requested to the bare minimum required to protect and recover the workload. If service actions are identified as unused, this may indicate that a restore activity has not occurred recently, but permission is still required to permit timely self-service recovery when required.

Related documents: SEC03-BP04 Reduce permissions continuously.

SEC03-BP05 Define permission guardrails for your organization.

Commvault recommends using **service control policies (SCPs)** to ensure access to the required AWS services, resources, and actions required to perform Commvault data protection.

Commvault recommends using AWS Control Tower and AWS Organizations to segregate data generation and data protection accounts in your organization.

When creating your multi-account landing zone (see **What is a Landing Zone?**), consider placing Commvault backup & recovery infrastructure in a tightly controlled **Shared Services Account** (see **AWS Landing Zone**). The Shared Services Account is where your data management infrastructure (Commvault CommServe® instance, MediaAgents, and Access Nodes) will reside. The Shared Services Account is where your backup data will reside in one or more Amazon S3 buckets. Shared Services Accounts are segregated from the general organizational user (OU) accounts or workloads and provide a **data isolation** boundary that prevents backup data from being accidentally or maliciously deleted (see **AWS Control Tower - Use multiple AWS** accounts for more detail).

Commvault infrastructure should be isolated into an Infrastructure OU which contains your shared services and networking accounts. This further restricts user access to high-value data management infrastructure and the backup and archival data it is responsible for collecting and securely storing.

Commvault recommends using **service control policies (SCPs)** to enforce the application of mandatory AWS Resource Tags that both grant IAM permissions and can be used by Commvault to auto-discover and protect AWS resources. Additionally, **permission guardrails** should ensure Resource Tags and backup data (I.e., Amazon EBS snapshots, Amazon Machine Images (AMIs), AWS KMS keys) cannot be deleted by non-administrative users.

AWS Control Tower **Account Factory** should be used to provision and apply consistent AWS IAM policy across all Organizational Units (OUs). Commvault recommends that all AWS accounts are provisioned with automatic trust to

the shared services backup account, and per-service permissions granted to allow the protection of AWS resources within the account.

Related documents: SEC03-BP05 Define permission guardrails for your organization.

SEC03-BP06 Manage access based on lifecycle.

Commvault recommends that user access is updated when the user changes organization roles or leaves the organization.

Commvault supports integrating authentication and authorization with central identity providers including SAML 2.0, OIDC, and Active Directory. Additionally, Commvault role-based access can utilize **centralized user groups**, allowing centralized group management, that results in discrete user group authorization within the Commvault platform.

Related documents: SEC03-BP06 Manage access based on lifecycle.

SEC03-BP07 Analyze public and cross-account access.

Commvault does not require public access to perform cross-account cross-region data protection and management. Use IAM Access Analyzer to identify resources that have been shared with an external entity.

Commvault recommends that in multi-account environments the data management resources are centralized into a shared services account. Backup and restore activity then utilizes **AWS Security Token Service (STS) AssumeRole** to temporarily assume a role in each protected account. Cross-account access is granted by the child account implicitly trusting ec2.amazonaws.com infrastructure running in the shared services (admin) account (see **Configuring STS Role Authentication Using a Tenant Account**). IAM Access Analyzer should be used before establishing cross-account trust and periodically re-run to ensure cross-account trust relationships are still required.

Commvault does not require any components to be shared or made available publicly unless services are offered over the public internet (i.e., edge-based data protection services).

Related documents: SEC03-BP07 Analyze public and cross-account access.

SEC03-BP08 Share resources securely within your organization.

Commvault recommends segmenting your data management environments between production and non-production to prevent accidental or intentional sharing a sensitive data between environments of differing criticality. Commvault recommends configuring AWS Organizations and segregating resources into differing organizational units (OUs) with appropriate policies per sensitivity of data retained. Commvault recommends implementing a stringent data perimeter (using IAM policies, network policies) for your central backup account, as it holds the backup copy of all your protected member account workloads.

Commvault does not require or recommend sharing of data (Amazon S3 buckets, Per-service snapshots, Amazon EC2 AMIs) from your Commvault data management platform publicly. Commvault ensures that any Amazon S3 resources created have public sharing disabled.

Commvault supports AWS Identity & Access Management (IAM) including AWS Security Token Services (STS) AssumeRole for securely accessing, sharing, protecting, and recovering data across AWS accounts. Resources are accessed using temporary, limited privilege credentials, and securely copied to a central backup account. Commvault integrates with AWS KMS and encrypts data transfers in transit (on the write) and at rest (on disk).

Commvault orchestrates the **secure copying and sharing of Amazon EBS snapshots** between accounts within the region, cross-region, and cross-account. Commvault must have access to the KMS key used to encrypt the source snapshot to perform the copy and/or share. When Commvault copies snapshots outside of the region it looks for the cvlt-master, cvlt-ec2, or cvlt-rds AWS KMS key alias in the destination account to re-encrypt the snapshot (see

Configuring cross-account sharing of AWS encrypted snapshots). Secure snapshot sharing/copying is supported by Commvault for **Amazon EBS**, **Amazon Aurora**, **Amazon RDS**, and **Amazon Redshift**.

Related documents: SEC03-BP08 Share resources securely within your organization.

SEC03-BP09 Share resources securely with a third party.

Commvault does not recommend or endorse using or sharing long-term AWS IAM credentials (i.e., access keys, secret keys) that grant access to your AWS accounts or data. Commvault recommends using **external ID conditions** when granting access to Commvault or central backup account resources to a third-party.

Additionally, Commvault does not recommend or endorse sharing Commvault account credentials as a method to

access backup data. Commvault recommends federating authentication to a central identity provider using SAML, OAuth, or Active Directory. Commvault recommends ensuring that credentials are time-based or short-lived. One example of federated and centralized management is Commvault integration with **CyberArk Password Security Platform** which securely centralizes credential management and rotation.

Any **access tokens** issued by Commvault for automated or API-based access have a built-in expiry timer.

Related documents: SEC03-BP09 Share resources securely with a third party.

Detection

Maintaining the security posture of your Commvault data management platform requires continuous detection of changes and validation of Commvault, AWS, and industry best practices. Detection involves the following key approaches:

Configure

- **Configure service and application logging**
- **Analyze logs, findings, and metrics centrally**

Investigate

Once you have a centralized, robust, secure audit trail of events in your environment, you can perform investigation and remediation (if required) using the following practices:

- **Implement actionable security events**
  Commvault-specific **runbooks** and **playbooks** should be developed for the most frequently occurring events and issues. Runbooks should be kept brief and written in clear, concise language that can be followed by operations personnel without additional assistance.

  Runbooks should also consider whether the event is considered a *false positive* and influence enhancement of alert and event routing logic to reduce the burden on operations teams. One approach is to ensure that when the event is generated it is *enriched* with additional metadata now normally included in the alert. This additional metadata can be used by CloudWatch Event rules and/or the SecOps agent to triage the issue more effectively (e.g. if a repeated login event has occurred on Commvault, enriching the event with the source IP address where the failures originated with help identify if the event is occurring from an internal or external network).

- **Automate response to events**
  A best practice, scalable approach to automated response and remediation can be achieved by using

Amazon EventBridge. In multi-account landing zones, you should ideally have a centralized logging account and/or Amazon VPC. All detection events, logs, and metrics should ultimately result in an Amazon EventBridge alert firing and then an EventBridge rule will match the event and execute/forward the event to a **supported target**. An example might be a **Commvault file anomaly detection alert** that identifies potential malicious I/O activity on an Amazon EC2 instance, the EventBridge rule will respond by running an EC2 `StopInstances` API call (see **EventBridge targets**).

SEC04-BP01 Configure service and application logging.

Commvault recommends that **AWS CloudTrail**, and **Amazon CloudWatch Logs** are enabled for Commvault and protected workloads, at a minimum. The Operating system and Commvault **Log Files** should be forwarded to Amazon CloudWatch Logs.

Additional logs may be required to perform service-level performance and security troubleshooting like **Amazon VPC Flow Logs** (network transfers), and **AWS Config** (resource quotas configured per account).

The first step to any detection or forensic investigation is a robust set of auditable events and service and application logs and metrics. Commvault recommends using Account Factory to ensure that all AWS accounts provisioned into your AWS Organization have **AWS CloudTrail** event history and **AWS CloudWatch** service and third-party application logging is enabled by default. There are multiple approaches to the automated deployment of CloudWatch agents on Amazon EC2 infrastructure (see **CloudWatch agent installation approaches for Amazon EC2 and on-premises servers**), depending on your use of Amazon Systems Manager, AWS CloudFormation, and Amazon Machine Images (AMIs).

Consider using a Configuration Management tool like **AWS Config**, CFEngine, and Puppet to define configuration baselines for AWS services and Commvault software. Configuration Management tools (including AWS Config) can then monitor your environment for configuration drift and alert you on non-compliant configurations. AWS Config provides a large list of **managed rules** which can be used to ensure an AWS account or service is configured for protection, for example:

- Is **cloudtrail-enabled**?
- Is **cloud-trail-encryption-enabled**?
- Is **ec2-instance-managed-by-systems-manager** (required for file and folder recovery)?
- Is **ec2-instance-profile-attached** (required for data protection activities)?
- Are **required tags** applied (required in attribute-based access control configurations)?

When performing large-scale data movement for backup and recovery, **VPC Flow Logs** should be enabled on all VPCs that will be transferring data for visibility and traceability.

As your environment grows, active security monitoring, logging, and alerting via **Amazon GuardDuty** and **AWS Security Hub** is warranted. Commvault exposes its services via multiple web-based services, including an API endpoint for programmatic control. Active monitoring for malicious activity and anomalous network activity should be employed to provide visibility to events, allowing manual or automated action.

The data retention policy for your business or industry vertical will dictate how long your events, logs, and metrics must be retained for active use. Commvault recommends backing up and removing infrequently used logs to a Commvault-controlled Amazon S3 bucket with Object Lock enabled. This provides deletion protection for the logs while storing them in optimized compressed, deduplicated, and encrypted format at a reduced cost. S3 lifecycle rules can be used to automatically remove logs after a fixed period.

Related documents: SEC04-BP01 Configure service and application logging.

Commvault recommends that all AWS accounts send their Amazon CloudTrail and Amazon CloudWatch logs to a central account. Centralizing logs will allow holistic search and analysis using **Amazon CloudWatch Logs Insights** and **Amazon Athena**.

Commvault logs and metrics can be sent to a central logging account using the **Amazon CloudWatch agent**.

Commvault can forward its operating system and application logs to Amazon CloudWatch using the CloudWatch agent installed within Linux and Windows infrastructure. Additionally, Commvault has a rich **alert & notification** system that is capable of pushing alerts near-realtime to centralized security incident and event management (SIEM) systems using **SNMPv3**, **webhook** to a third-party system, pushing to **OS event logs**, or running a **command-line** or **workflow**.

Commvault can therefore push information and critical alerts into a centralized SIEM or SOAR system, or directly **raise a ServiceNow incident** for SecOps investigation and remediation.

For DevSecOps teams that use real-time collaborative chat platforms, alert notifications can call third-party webhooks to push alerts directly into chat platforms, for example, **Teams chat groups** and **Slack channels**.

Logic associated with whether a particular alert or event should be forwarded to an upstream system on the SIEM system may be implemented in Commvault command-line scripts, workflows, or centralized (recommended) by using **Amazon CloudWatch Event rules**. As your environment grows, your event rules may call a decision engine in Amazon Lambda to aggregate multiple data sources before determining where to route an event.

Related documents: SEC04-BP02 Analyze logs, findings, and metrics centrally.

Commvault recommends using Commvault **command-line notifications** to automate responses to common events.

High-risk events like the **File Activity Anomaly Alert** that indicate potential ransomware activity can automate responses like taking instances offline or locking backup aging.

Amazon EventBridge can also perform automated responses to common events using Amazon Systems Manager Run Command.

Related documents: SEC04-BP03 Automate response to events.

Commvault notifications can include **Alert Tokens** that provide critical context to allow timely action when received by SecOps teams (date/time, resources, failure message).

Related documents: SEC04-BP04 Implement actionable security events.

Infrastructure Protection

Infrastructure protection of your Amazon EC2 infrastructure running your Commvault data management platform is critical to protecting your business workload backup and archival data. Infrastructure protection involves securing your infrastructure to internal IT best practices, standards, and industry or regulatory security standards (e.g., **CISA Cloud Security Technical Reference Architecture (Version 2)**).

Your infrastructure security approach must consider computing infrastructure, network access, operating system hardening, and ongoing patch management. Secure authentication and data handling using encryption, multi-factor authentication, and data-at-rest encryption should be used. High-risk interfaces like web-based systems should leverage active monitoring and machine learning-enabled detection systems to provide application-level protections.

Availability of your Commvault data management platform will also be dictated by distributing the workload across multiple Regions, Availability Zones, AWS Local Zones, and AWS Outposts.

Consider the following approaches to infrastructure protection and apply protection where your business availability and security needs require additional protection:

Protecting Networks

When architecting protection for your Commvault data management platform, consider the **Zero Trust** approach to applying security at all layers of your workload. Zero trust is a security approach that assumes that risks and threats exist for all workload components, regardless of location. This model means that each infrastructure and application component will require discrete security controls and does not inherently trust any other component.

Your networking topology will include multiple segmented network boundaries (public, private, shared services), and the ability to control (accept, deny) and inspect traffic must be built into your network design from day one.

Consider the following network protection elements to build a robust, secure, operatable network.

SEC05-BP01 Create network layers.

Commvault data management resources must have a network path to the following AWS **service endpoints** and **external URLs** for patch download and remote support.

Commvault data management resources must be permitted to access and transfer data to/from protected workload VPCs. Commvault accesses protected networks via VPC peering or **AWS Transit Gateway** in large multi-region multi-account environments.

Commvault consists of multiple logical components (CommServe® instance, MediaAgents, and Access Nodes) that may exist as a single Amazon EC2 instance or multiple instances in scaled-out deployments. Commvault infrastructure should be owned by an AWS Shared Services Account within your multi-account multi-region **AWS Organizations Landing Zone** and deployed in a dedicated shared services (backup) subnet.

Commvault infrastructure does not require a public subnet or public IP address assigned to Commvault infrastructure in normal operation. In environments where Commvault is providing services accessible over the internet, Commvault **WebServer/WebConsole** components should be segregated into a public-facing subnet with additional controls before communication with internal Commvault CommServe®.

Commvault is granted access to AWS networks and workloads to protect, using either **Amazon VPC peering** or **AWS Transit Gateway** connections between your Commvault shared services VPC and per-account, per-region workload VPCs. Commvault recommends when the number of accounts, VPCs, and on-premises networks grows that AWS Transit Gateway be leveraged to provide a centralized hub for routing traffic between your AWS VPCs and other network locations.

AWS transit gateway provides additional network protections to inter and intra-region traffic by keeping traffic local to the AWS private network. AWS inter-region traffic is encrypted and is delivered without any single point of failure or bandwidth bottlenecks.

Related documents: SEC05-BP01 Create network layers.

SEC05-BP02 Control traffic at all layers

Commvault Access Nodes require access to **AWS API endpoints** to permit backup and recovery activities. The following region and global endpoints must be reachable by Commvault Access Nodes and MediaAgents.

Commvault recommends the use of **AWS PrivateLink** VPC endpoints to keep network traffic internal to your VPC (where supported) and to reduce the cost of performance backup and recovery services (see **Reduce Cost and Increase Security with Amazon VPC Endpoints**).

Commvault supports **interface** and **gateway** VPC endpoints, including **FIPS Endpoints** and AWS GovCloud **VPC Endpoints**.

Typically, the Access Node will read and write workload data by accessing the backup data in Amazon S3 via MediaAgent. When **Commvault Storage Accelerator** is installed on the Access Node, the Access Node also speaks with the Amazon S3 service directly.

Service Name / Service Endpoint	Endpoint type	AWS service protected	Access Node	MediaAgents
dynamodb.*region*.amazonaws.com	**Gateway**	Amazon DynamoDB.	●	●
com.amazonaws.*region*.ec2	**Interface**	Amazon VPC, Amazon EBS, Amazon EC2.	●	●
com.amazonaws.*region*.ebs	**Interface**	Amazon EBS, using EBS direct APIs.	●	●
glacier.*region*.amazonaws.com	Not supported	Amazon Glacier.	●	●
kms.*region*.amazonaws.com	**Interface**	*Used to access and utilize AWS KMS keys to decrypt and encrypt protected sources.*	●	●
rds.{region}.amazonaws.com	**Interface**	Amazon RDS (including Amazon Aurora).	●	●
redshift.{region}.amazonaws.com	**Interface**	Amazon Redshift.	●	●
s3.{region}.amazonaws.com **s3-outposts.{region}.amazonaws.com**	Gateway **Interface**	Amazon S3 (as source). Amazon S3 (as target).	●	●
ssm.{region}.amazonaws.com	**Interface**	*Used to perform in-place agentless file recovery to Amazon EC2 instances and Operating System detection.*	●	●
sts.*region*.amazonaws.com	**Interface**	*Used to obtain temporary credentials to perform backup and recovery activities. Also requires access to the global endpoint.*	●	●

Access to the following **Global Service Endpoints** is required by Commvault Access Nodes and MediaAgents. These endpoints will require internet access from Commvault (CommServe®, MediaAgent, Access Node) to the following endpoints. Access may be provided using Internet Gateway (IGW), NAT Gateway, or an HTTP_PROXY.

Service Endpoint	VPC Endpoint	AWS service protected	Access Node	MediaAgents
iam.amazonaws.com	Not supported by AWS Privatelink	*Used to perform authentication and authorization for AWS machine identities used by Commvault software.*	●	●
sts.amazonaws.com	Not supported by AWS Privatelink	*Used to obtain temporary credentials to perform backup and recovery activities. Also requires access to the regional endpoint.*	●	●
importexport.amazonaws.com	Not supported by AWS Privatelink	*Used to submit AWS Import/Export jobs for migrating on-premises VMs to Amazon EC2 instances.*	●	

ⓘ **Note**: Commvault does <u>not</u> require service endpoint communication to CloudHSM, CloudTrail, CloudWatch, EFS, EKS, FSx, Outposts, Snowball, Storage Gateway, Direct Connect, and Workspaces services but can protect/interact with these services.

Refer to **AWS Service Endpoints and Usage Information** for additional information.

> The use of VPC Endpoints is highly recommended as it reduces availability risks and bandwidth constraints on the VPC's link to the public Internet.

An Amazon S3 VPC Endpoint must first be defined by creating an Endpoint Policy within the AWS console, but there is no change to the FQDN hostname used to define the Cloud Library within Commvault. Instead, AWS will ensure that DNS queries for the hostname will resolve against the Amazon S3 VPC Endpoint, instead of the public address, and apply appropriate routing (provided the Endpoint Policy is successfully created).

For more information on VPC Endpoints, please refer to this AWS documentation: **VPC endpoints**.

Amazon VPC network ACLs and security groups should be used to enforce communication over **Commvault-published ports and protocol** requirements.

Commvault **network topologies** should be used to control the accepted communication between Commvault components.

Commvault CommServe® instance will require outgoing internet access to perform software downloads, dial-home telemetry & software support (see **External URLs for Commvault Features**). Outgoing access should be restricted to allow connections only to trusted external URLs via Amazon VPC **security groups** and **network ACLs**.

Communication to external URLs does not require direct access to the internet and may occur via an **Internet gateway**, **Egress-only Internet gateway**, or NAT gateway. Additionally, Commvault can direct outbound external communications from the CommServe® instance and Cloud Access Nodes via an authenticated **HTTP_PROXY** and or **Commvault Internet Gateway** for additional control and inspection point when crossing internal network boundaries.

Commvault requires public-facing web services when **edge-located** end-users access Commvault via the internet. This pattern occurs typically when providing **endpoint protection and file-sharing** services. In this pattern, the

Commvault web or presentation layer components reside in a **public subnet**, while the Commvault CommServe® application and database reside in a **private subnet**. Network ACLs should be used to narrow the ports and protocols permitted inbound and outbound from both public and Commvault shared services private subnets.

Commvault recommends using Amazon VPC endpoints to provide another layer of network control and permission management (refer to the **VPC endpoint policies** column). Commvault recommends using **AWS PrivateLink** endpoints to ensure communication to AWS service endpoints is kept internal to the AWS network and your VPC. Commvault software will identify when non-optimal network communication occurs and generate backup and restore job alerts identifying the misconfiguration.

Commvault recommends using **security group** rules that specify the source (inbound rules) as the *ID of a security group* from the same VPC or a peered VPC. This allows all Commvault protection infrastructure associated with the selected security group to speak securely with the instance. Commvault publishes the **port requirements for Commvault** which should be applied to Commvault infrastructure and protected infrastructure using both Amazon VPC network ACLs and security groups (see **Compare security groups and network ACLs** for guidance on which control fits your network design).

Commvault publishes a list of required **Service Endpoint Connectivity for Access Nodes** that Commvault makes API calls against. You should provision regional service endpoints to keep traffic internal to your VPC using **gateway endpoints** (Amazon S3) or interface endpoints (see **Supported Services**). Global endpoints will require an HTTP_PROXY or access via Internet Gateway (IGW) / Egress-only Internet Gateway from Commvault Access Nodes (Amazon EC2 instances with the Commvault Virtual Server Agent package installed).

Commvault can create additional network controls in the form of **dedicated backup networks** and **network topologies** which create authenticated **network routes** over secure encrypted tunnels between Amazon EC2 infrastructure with Commvault core software installed. Network traffic can be directed to known ports in the following topologies/patterns:

- one-way (client to server, server to client).
- two-way.
- **port-forwarded** through a network gateway that utilizes network-address translation (NAT)
- forwarded from the **network perimeter (DMZ)**.
- between **Network Zones**.

Cascaded network gateways may be used to create a software-based network control and inspection point for each subnet that communicates with Commvault shared services network(s).

Additionally, Commvault recommends protecting Amazon VPC configurations so that failed or unauthorized changes can be audited and remediations to security posture can be made. See **Auto-discovering, protecting, and recovering Amazon VPC resources** for the AWS IAM permissions required to collect and recover supported Amazon VPC resources. Commvault will recreate missing Amazon VPCs, Subnets, Elastic Network Interfaces, and Security groups during EC2 instance recovery, ensuring the security controls on the application are maintained.

Related documents: SEC05-BP02 Control traffic at all layers.

SEC05-BP03 Automate network protection.

Commvault services exposed via HTTP can be integrated with web-based traffic filtering and protection provided by **AWS WAF**.

Commvault anomaly detection alerts can run custom notifications that perform automated network protection, like applying a DENY ALL security group to segregate a workload.

Ensure that protection mechanisms are always active and effectively deliver a self-defending network. Ensure that both static network blocks (Network ACLs) and intelligent anomaly-based detection (AWS WAF) are applied at each network perimeter to continuously assess incoming traffic for threats and automate mitigation where feasible.

Commvault automates network protection by providing an application-level firewall for every Commvault instance running the Commvault core package. Commvault recommends and automatically configures Microsoft Windows and Linux-based firewalls to allow incoming connections on authenticated ports & protocols Commvault listens on. Commvault also protects and recovers missing network resources and controls (Amazon VPCs, Subnets, Security groups) as part of Amazon EC2 instance recovery to ensure restored instances are 'secure by default'.

Related documents: SEC05-BP03 Automate network protection.

SEC05-BP04 Implement inspection and protection.

Commvault can operate in environments with **Amazon GuardDuty** enabled and monitoring for malicious activity.

Consider periodically auditing accepted network flows into the centralized Commvault service account using **VPC Network Access Analyzer**.

For high-sensitivity data transfers, consider **VPC traffic mirroring** to perform out-of-band security monitoring and threat detection on traffic flows.

Inspection of traffic at all layers of your network should occur frequently to optimize and restrict network flows that are no longer used or permitted. Use **VPC Network Access Analyzer** and **VPC Flow Logs** to monitor as-configured and observed network transfers.

Commvault web-based services that expose services via the HTTP protocol may use an **AWS WAF** in front of incoming connections to allow inspection and control to **managed** or **custom** rules. AWS WAF is a web application firewall that monitors and controls HTTP(S) requests and then forwards them to your Application Load Balancer (ALB).

As your deployment of AWS WAF resources expands, consider using **AWS Firewall Manager** to centrally manage the definition and distribution of WAF rules across all your AWS accounts and VPCs. As the number VPCs grows in your multi-account multi-region landing zone, you can deploy **AWS Network Firewall** to perform stateful layer 3-7 intrusion detection and prevention for your VPCs.

Related documents: SEC05-BP04 Implement inspection and protection.

Protecting Compute

Commvault Backup & Recovery runs on x86 Linux-based and Microsoft Windows-based Amazon EC2 instances. Protecting your compute infrastructure is crucial to ensuring your business can continue protecting and recovering from unintended data loss events.

A defense-in-depth approach with zero trust principles should be applied to reduce the attack surface and monitor and mitigate vulnerabilities across Commvault instances and workloads.

Consider the following key capabilities when developing your compute protection strategy.

Related documents:

SEC06-BP01 Perform vulnerability management.

Implement continual scanning and patching for vulnerabilities in Commvault and protected workload infrastructure.

Consider **Amazon Inspector** to perform holistic vulnerability scanning across all the resources across your AWS Organization.

Commvault provides automated software update download and install (by schedule), to ensure a continuous improvement approach to vulnerability management.

Commvault recommends frequently scanning your Commvault CommServe® instance and related infrastructure (MediaAgents, Access Nodes) for vulnerabilities. Static software vulnerability scanners like **Amazon Inspector** should be used to periodically scan Amazon EC2 infrastructure for software vulnerabilities and network exposure risks.

Identified risks can be routed directly into **Amazon EventBridge** and **AWS Security Hub** to ensure visibility, audibility, and automated remediation or routing to appropriate SIEM/SOAR systems or incident management helpdesk.

Commvault infrastructure is available in pre-configured, pre-secured, and scanned Amazon Machine Images (AMIs) from the **AWS Marketplace**. Commvault AMIs deployed via AWS Marketplace utilize **AWS CloudFormation** secure-by-default templates to provision and secure your deployments from day 1. Once deployed with CloudFormation, you continue to manage and update your Commvault infrastructure using continuous deployment via **AWS CodePipeline** or equivalent continuous deployment tools. Commvault recommends automated provision (build), test, and release of Commvault infrastructure changes using infrastructure-as-code to reduce error and allow security scanning of planned changes with tools like **CloudFormation Guard**.

Automated Operating systems and Commvault application patching are required to maintain stable and secure data management operations. Commvault provides the **Install Windows Updates** workflow for pushing Microsoft Windows and Microsoft SQL Server patches to Commvault and protected workloads. Commvault application updates may be automatically downloaded and installed on a **defined schedule**. Commvault releases both **platform releases** (every six months), **long-term support platform releases** (once per year), and **maintenance releases** (monthly). Care should be taken to test all software updates in pre-production before deployment into Production. Platform releases require additional rigor in testing as platform releases cannot be uninstalled or reverted after installation.

Alternatively, you can use **AWS Systems Manager Patch Manager** to automate and orchestrate patch delivery to your Amazon EC2 infrastructure running Microsoft Windows and AWS-supported Linux variants. Commvault AMI images available in the AWS Marketplace have AWS Systems Manager pre-installed. Patch Manager uses *patch baselines*, which include automated rules for auto-approving patches within days of their release, in addition to lists of approved and rejected patches.

The patch manager can also scan your instances for compliance and generate and store compliance reports in Amazon S3 for compliance reporting.

Commvault application patches may be installed using **Commvault REST API** executing an existing pre-configured **software update schedule** or install updates in unattended mode using an **AWS Systems Manager Run Command** script. Using AWS Systems Manager will require a centralized software depot that AWS SSM can access when performing patch download and installation.

Commvault maintains a **centralized software cache** on the CommServe® instance.

Related documents: SEC06-BP01 Perform vulnerability management.

SEC06-BP02 Reduce attack surface.

Commvault recommends using **AWS Marketplace images** that have been patched and hardened to the latest release.

Commvault provides **CIS Hardening scripts** for further hardening the Commvault CommServe instance.

Commvault **auto-scaling access nodes** have a reduced attack surface as they are created without a configured Amazon EC2 key pair, preventing login.

Commvault recommends using the Commvault AMIs available via AWS Marketplace for a pre-hardened all-in-one CommServe® instance (Commvault Backup & Recovery BYOL) or combined MediaAgent Access Nodes (Commvault Cloud Access Node).

Commvault Cloud Access Nodes (via AWS Marketplace) are used to provide auto-scaling access nodes for Amazon EC2 backup and launch <u>without</u> a configured EC2 keypair, further reducing the attach surface as no interactive login is permitted.

Commvault provides **CIS Level 1 hardening scripts** that should be applied to your CommServe® instance to achieve CIS Level 1 benchmark protection for your Operating System and SQL server installation. Commvault provides prescriptive guidance on **Securing your CommServe® instance**, each recommendation should be reviewed for applicability to your environment. All hardening recommendations have been tested by Commvault and will not limit or impact protection activities.

Commvault uses third-party static code analysis tools (i.e., **Quay/Clair**), dependency checking tools, and third-party and internal penetration testing to ensure that Commvault software is protected against known and newly released common vulnerabilities (CVEs). Commvault promotes **reporting vulnerabilities** and provides transparency of **detected and mitigated vulnerabilities**.

Related documents: SEC06-BP02 Reduce attack surface.

SEC06-BP03 Implement managed services.

Commvault recommends application optimization by adopting fully-managed services over infrastructure-based solutions. Commvault can help migrate **Oracle** from on-premises or Amazon EC2 to Amazon RDS.

> ⓘ **Note**
>
> The Commvault-embedded Microsoft SQL Server database cannot be deployed on Amazon RDS.

Consider leveraging AWS fully-managed services to offload the operational burden and responsibility of building, operating, and maintaining compute and network services you require.

- Consider placing your backup & recovery operational runbooks and playbooks in highly-available managed storage services like Amazon EFS, Amazon FSx for Windows, or Amazon S3.
- Consider leveraging fully-managed networks and security services like AWS Transit Gateway, AWS WAF, and AWS Shield to simplify and secure the management of your distributed AWS network resources.

Additionally, Commvault **Enterprise Support**, **Residency Services**, and **Remote Managed Services** can help you use Commvault data management experts to monitor, optimize, and secure your Commvault data management platform.

Related documents: SEC06-BP03 Implement managed services.

SEC06-BP04 Automate compute protection.

Commvault recommends automating the deployment of secure Commvault infrastructure using **AWS CloudFormation** and Commvault-published Marketplace AMIs.

Commvault can automatically **download** application, OS, and Commvault HyperScale™ scale-out storage appliance patches and **install** updates on a schedule. **AWS Systems Manager Patch Manager** can be used to apply operating system patches.

Commvault patch management and vulnerability scanning may be automated utilizing Commvault Workflows or AWS Systems Manager Run Command. Additionally, Commvault provides a rich REST API, command line, and Python SDK for bespoke automation initiatives.

Related documents: SEC06-BP04 Automate compute protection.

SEC06-BP05 Enable people to perform actions at a distance.

Commvault removes console access for auto-scaled access nodes, used on **AWS Systems Manager Run Command** to automate management tasks.

Commvault CommServe and MediaAgent continue to require secure console access via SSH and/or RDP.

Commvault operation and maintenance can be automated and orchestrated *from a distance* using AWS Systems Manager **Run Command** and/or industry configuration management looks like Ansible, CFEngine, Chef, Puppet, and SaltStack using the Commvault **command line** and **Python SDK**.

Commvault infrastructure should be deployed using published AWS CloudFormation templates available in AWS Marketplace and changes made, deployed, and tested through existing continuous integration/continuous deployment systems for example **Amazon CodePipeline** in pre-production, before production.

Commvault software is built for lights-off management with day one definition of business policies in **server plans**, then automated workload discovery and protection occurs across accessible AWS accounts and regions by **AWS tags** attached to resources. Commvault uses A.I. and machine learning to continually tune data management activities to meet configured business policies or SLAs as your environment grows and shrinks.

Related documents: SEC06-BP05 Enable people to perform actions at a distance.

SEC06-BP06 Validate software integrity.

Commvault utilizes MD5 checksums to validate the integrity of software updates downloaded from the Commvault update store (cloud.commvault.com, downloadcenter.commvault.com).

Commvault performs automatic software download if network access to Commvault External URLs has been allowed. Automatic software downloads utilize MD5 checksums to ensure the software has not been modified.

Commvault also publishes software downloads and MD5 checksums via **store.commvault.com** and **Maintenance Advantage** support portal. Commvault does not cryptographically sign software packages with public and private keys.

Related documents: SEC06-BP06 Validate software integrity.

Data Protection

Data Protection involves understanding what data you have, how sensitive or valuable the data is, and how to protect data based on differing sensitivity levels. Data protection practices prevent and/or detect data mishandling, data leakage, and data loss risk. Complying with your business or industry regulatory obligations will require a stringent data protection strategy and mitigations. Consider the following when developing your data protection strategy:

Data Classification

Data classification requires understanding the criticality and sensitivity of a dataset, so you can apply appropriate data protection controls. It should be noted that the criticality of data will change over its lifetime from creation to eventual deletion. The sensitivity of data often does not change across its lifetime but should be validated with **Commvault Risk Analysis** or **Amazon Macie** periodically.

SEC07-BP01 Identify the data within your workload

Commvault provides **Commvault® Risk Analysis** to define, find, and manage critical or sensitive data in your backups and live systems. Once data is classified, you can apply appropriate data protection controls ranging from backup, migration, or deletion.

Classify your data to capture the different criticality and sensitivity levels present in your business. This includes understanding who creates the data, which business processes use the data, and any legal or compliance requirements on the handling of the data (i.e., compliance with your published privacy policy).

Sometimes business processes will attach metadata to documents to indicate whether personally identifiable information (PII) or intellectual property is present. Often, there is a large set of unstructured data that may not have a classification attached. **Commvault® Risk Analysis** can be used to identify high-value PII data in your organization for classification and appropriate data protection. Risk Analysis can also inspect your data landscape and identify infrequently used content that can be classified for archival. **Commvault® eDiscovery & Compliance** can be used to perform targeted discovery across structured data (i.e., emails) for legal and compliance requests.

Data Type	Sensitive Data Identification	Duplicate, aged, orphaned Data Identification	eDiscovery
File servers (Windows, Linux)	●	●	●
Endpoints (laptops, desktops)	●	●	●
Object storage	●	●	
Email messages (Microsoft Exchange)	●		●
OneDrive for Business	●		
SharePoint Online	●		

Source: **Data Source and Language Support in Commvault Risk Analysis**.

Related documents: SEC07-BP01 Identify the data within your workload.

SEC07-BP02 Define data protection controls.

Commvault recommends defining a set of data classifications and protection policies (**Plans**) based on data sensitivity and value. Apply AWS resource tags to enforce the protection policy applied during backup. You should ensure that AWS workloads (including tags) are protected, and recovery **runbooks**/**playbooks** include validation or customization of tags to ensure data protection controls are maintained for newly created resources.

Reflect the data classification of individual workloads and/or all data stored within specific AWS accounts using **AWS Resource Tags**. Commvault will use resource tags to auto-discover and protect workloads based on business policy. Based on the resource tag an appropriate backup frequency, storage location, storage encryption key, and additional backup copies can be enforced by **Commvault server plans**.

Commvault infrastructure will need to be granted access to each of the AWS KMS keys used to encrypt the data in each AWS account. Commvault will maintain encryption of data when copying data between accounts and regions. Commvault supports both single-region and **multi-region keys in AWS KMS**. When using single-region keys Commvault copies data outside of the scope of a key, and Commvault will re-encrypt the data with a destination-accessible key.

Commvault server plans will securely upload the backup data and backup copies to an appropriate encrypted Amazon S3 bucket with **Server Side Encryption with KMS keys stored in AWS Key Management Service (SSE-KMS)**, **Server Side Encryption with S3-Managed Keys (SSE-S3)**, or **Server-Side Encryption with Customer-Provided Keys (SSE-C)**, or **Dual-layer server-side encryption with keys stored in KMS (DSSE-KMS)**. Alternatively, **Commvault client-side software encryption** may be used to encrypt data before transfer to Amazon S3, client-side encryption will directly impact the client CPU and should be planned carefully.

Related documents: SEC07-BP02 Define data protection controls.

SEC07-BP03 Automate identification and classification.

Consider automating the application of data classification tags based on region, account, and data governance discovery. Consider an appropriate default classification. Commvault will automatically recover tags on supported AWS resources during the restore. Additionally, Commvault allows customization of tags during Amazon EC2 compute restores.

Commvault automates the identification and classification of data using **data classification plans**. Data classification plans perform the automated indexing of supported data sources in preparation to discover and classify infrequently accessed, duplicate, or personally identifiable data (PII). Commvault can **tag files** that match identification or classification filters. Commvault recommends using Commvault® Risk Analysis and Commvault® eDiscovery & Compliance to analyze and manage your structured and unstructured data for criticality and sensitivity.

Related documents: SEC07-BP03 Automate identification and classification.

SEC07-BP04 Define data lifecycle management.

Identify the types of data the organization generates, stores, and uses. Define a data lifecycle strategy for the protection, sharing/reuse, and eventual destruction of data types. Ensure legal counsel is consulted to ensure compliance with internal and regulatory requirements.

Commvault automates data lifecycle management from frequent access to infrequent access storage classes using **storage copies**.

Commvault Risk Analysis can be used to identify redundant, obsolete, or trivial data and then **archive it**.

Commvault software will protect workload data during the initial operational lifetime with frequent or infrequent access expected for backup copies. Commvault recommends using **Amazon S3 Intelligent-Tiering** with the **Archive Instant Archive Tier** to store active backup data with an unknown amount of archival content (see Commvault Combined Storage Tier below for predictable use archives).

> ⚠ Important Warning
>
> Based on how Commvault deduplication processes your data, the access patterns used to inform S3 Intelligent-Tiering when to transition objects *may* hold data in more frequent access tiers. Commvault recommends copying archives out to Commvault Combined Storage Tier to avoid this issue.

As data ages or is no longer frequently accessed but must be retained for legal or regulatory reasons, a **selective copy** or **synchronous copy** of targeted data should be made to a **Commvault Combined Storage Tier** bucket. Commvault combined storage is a combination of frequent/infrequent access and archival Amazon S3 storage classes.

> 💡 **Pro-Tip**
>
> Commvault does not recommend the use of Amazon S3 Glacier Flexible Retrieval, S3 Glacier Deep Archive, or S3 Intelligent-Tiering Archive Access and Deep Archive Access buckets directly. Archival of Commvault indexing data will delay data recall, use **Commvault Combined Storage Tier** to leverage ultra-low-cost S3 Glacier while simplifying restore operations in Commvault.

Consider which Commvault users and administrators require access to data copies throughout their lifecycle. Regulatory and compliance backup copies should have increased access controls to ensure that long-term retention data cannot be accidentally or intentionally deleted. Regulatory data and data used to recover critical workloads may require immutable copies, **Amazon S3 Object Lock** may be enabled to prevent the deletion of objects during the retention period.

Commvault® Risk Analysis can be used to identify unused, duplicate, and inappropriately secured files as part of best practice data management. Files may then be **archived** then automatically deleted after the request is reviewed and approved. Alternatively, files may be flagged to be **deleted** and then automatically deleted after the request is reviewed and approved.

Commvault provides multiple levels of access for data by providing another level of authenticated and authorized access to backup and archival data. Users with access to the primary backup location do not need to be granted access to backup or long-term archival data copies.

Related documents: SEC07-BP04 Define data lifecycle management.

Protecting Data At Rest

Commvault supports and enforces encryption of data at rest stored in all supported storage locations including **Amazon S3**, **Commvault HyperScale™ scale-out storage**, **Network Attached & Direct-Attached Storage**, and **tape** (including Amazon Tape Gateway).

SEC08-BP01 Implement secure key management.

Commvault utilizes AWS KMS to create, manage, and control the use of encryption across AWS services, including Commvault backup data. Commvault supports multi-region keys for replicating Amazon snapshots cross-region without requiring re-encryption at the destination.

Commvault can also use any **KMIP-compliant key manager** (including AWS CloudHSM) to create, access, and use software encryption keys for Commvault encrypted data.

Commvault software-based encryption uses **FIPS-140-3**-certified ^pending CMVP review cryptographic modules.

Commvault implements secure key management that includes storage, rotation, and access control to stored keys. Commvault integrates seamlessly with **AWS Key Management Service (AWS KMS)** and publishes IAM policies with required KMS service actions to facilitate backup and recovery activities.

> ⓘ **Note**
>
> Automatic key rotation must be disabled when using AWS KMS to perform key management for Commvault software encryption keys (see **Key Rotation Guidelines for AWS Key Management Service Server**).

Commvault supports the use of **AWS CloudHSM** and any third-party **Key Management Interoperability Protocol (KMIP)** compliant products to store and protect Commvault software encryption keys. Commvault cryptographic handling for backup data (at rest) and replicated (in-flight) is **certified FIPS-140-3 compliant** ^{pending CMVP review}

.

Commvault recommends enabling AWS CloudTrail on AWS KMS and regularly auditing the use of KMS keys. Using a tool like **Amazon CloudWatch Insights** you can identify anomalies in key usage and remove any access not required.

Related documents: SEC08-BP01 Implement secure key management.

SEC08-BP02 Enforce encryption at rest.

Commvault recommends and supports encryption at rest for all persistent storage written by Commvault into AWS Storage including file, block, and object storage. Storage snapshots created and replicated by Commvault inherit the encryption settings of the source workload or optionally are re-encrypted by Commvault when migrating between regions or AWS accounts.

Commvault supports and recommends the use of Amazon S3 server-side encryption with either Amazon S3-managed keys (**SSE-S3**), AWS KMS keys (**SSE-KMS**), customer-provided keys (**SSE-C**), or dual-layer SSE with KMS keys (**DSSE-KMS**).

Commvault recommends enabling **S3 Bucket Keys** to reduce requests from Amazon S3 to Amazon S3.

Commvault supports snapshot and streaming backup and recovery of encrypted AWS compute, database, and storage services across regions and AWS accounts.

Commvault can enforce encryption of data at rest by assuming the responsibility to perform the encryption before uploading to the final storage location, or by enforcing the use of a bucket with server-side encryption (SSE) enabled.

Commvault recommends pre-creating Amazon S3 buckets with the preferred encryption scheme (SSE-S3, SSE-KMS, SSE-C, DSSE-KMS) and ensuring that the Commvault backup & recovery role is granted access to keys to perform data management to the bucket. Commvault recommends using **S3 Bucket Keys** with SSE-KMS for a cost saving of up to 99% on AWS KMS requests.

Consult **Should I use an AWS KMS-managed key or a customer-managed KMS key to encrypt my objects on Amazon S3** to determine which encryption scheme is best for your needs.

Commvault can also force the writing of encrypted data to buckets that have had encryption disabled at the bucket level, see **Enabling Server-Side Encryption with Amazon S3 Managed Keys (SSE-S3)**. Note that after January 2023, all buckets have encrypted enabled by default.

Commvault recommends in regulated environments that require multiple levels of encryption (i.e., to meet National Security Agency CNSSP 15 for FIPS compliance and Data-at-Rest Capability Package (DAR CP) Version 5.0 guidance), that **Amazon S3 Dual-Layer Server-Side Encryption with Keys Stored in AWS Key Management Service (DSSE-KMS)** is used.

If data is written to an Amazon S3 bucket unencrypted and then **batch operations** are used to encrypt existing unencrypted objects in a Commvault cloud storage library, Commvault access to the object will continue unaffected, as long as Commvault is granted access to the relevant KMS keys.

When using Commvault with AWS Snow family to perform offline data migration, data will be encrypted while in transit using **Encryption in AWS Snowball**. Commvault can provide a third level of encryption by using FIPS-140-3 compliant ^{pending CMVP review} *software encryption* when writing to the Snow device.

Commvault supports protecting and creating snapshot-based backups with **encryption** enabled. For more information see:

- **Copying and Sharing Snapshots Within Amazon** (using the cvlt-master, cvlt-ec2, cvlt-rds CMK keys).
- **Enabling Cross-Account Copying of an Amazon RDS Snapshot Copy** (Amazon Aurora, Amazon RDS).
- **Creating a Snapshot Copy of Amazon EC2 in a Different Account** (Amazon EC2, Amazon EBS).
- **Enabling Server-Side Encryption with Amazon S3-Managed Keys (SSE-S3, SSE-KMS)** (Amazon S3).

Related documents: SEC08-BP02 Enforce encryption at rest.

SEC08-BP03 Automate data at rest protection

Commvault recommends using AWS Config and **AWS Managed Config Rules** be used to validate that the encryption is enabled by default on all persistent storage and services.

AWS Config Rules should be used to automatically **remediate** non-compliant resources (e.g., **Automatically encrypt existing and new EBS volumes**).

Commvault recommends using Amazon server-side encryption (SSE) over Commvault software encryption to avoid the CPU impact on data transfer operations.

Commvault provides **Client Encryption reporting** that can be used to automate the auditing and remediation of clients or data copies configured without encryption enabled. Remediation can be performed directly within the **Client Encryption – Edit encryption settings** function. Commvault has a rich REST API that can be used to automate the execution of repeatable actions, including enabling client and storage copy encryption.

Commvault infrastructure deployed from the Commvault Amazon Marketplace Images (AMIs) selects encryption by default. Commvault recommends using AWS Config and AWS Organizations Service Control Policies (SCPs) to ensure that Commvault infrastructure is built with encryption enabled on Amazon EBS volumes and related Amazon S3 buckets.

Related documents: SEC08-BP03 Automate data at rest protection.

SEC08-BP04 Enforce access control.

Commvault recommends segregating backup data into separate accounts per data classification level. Ensure only Commvault infrastructure and machine identities are **permitted** to access and delete content from Commvault Amazon S3 buckets.

Ensure that AWS KMS keys are shared to the Commvault service account, and apply **cvlt-master**, **cvlt-rds**, and **cvlt-ec2** key aliases to ensure re-encryption of snapshots can occur when restoring across key boundaries.

> ⓘ Note
>
> Commvault does not support Amazon S3 versioning, define an **Amazon S3 lifecycle policy** to **DeleteObjectVersions** if versioning is enabled.

Commvault recommends that **Amazon S3 bucket policies** and **Amazon S3 gateway endpoint** policies are used to prevent access to Commvault-controlled Amazon S3 buckets. For example, you can limit access to the S3 bucket from only a **trusted list of IP addresses**. Access to Commvault-controlled buckets should be permitted from

Commvault data management infrastructure, <u>only</u>. There is no requirement for Commvault users or Administrators to access Commvault Amazon S3 buckets directly, data is stored in a shared proprietary format.

> 💡 **Pro-Tip**
>
> Commvault does not support **Amazon S3 versioning in buckets**. Ensure that versioning is disabled on Commvault-controlled buckets or create S3 lifecycle policies to periodically delete all but the latest version of objects. Commvault recommends setting the `NoncurrentVersionExpiration` action to delete non-current versions after 2 days.

Enable **Amazon S3 Object Lock** for archival or backup copies that require additional protections provided by **write-once read-only (WORM)** copies. Consider which datasets require Object Lock protection as the use of WORM or immutable Amazon S3 buckets results in typically a 2.6x increase in data storage consumption for a typical backup schedule. You can also enable NetApp® SnapLock® Compliance and NetApp SnapLock Enterprise on Amazon FSx for NetApp ONTAP file-mode shares and exports – see **Configuring WORM Storage Mode on Amazon FSx for NetApp ONTAP**.

Related documents: SEC08-BP04 Enforce access control.

SEC08-BP05 Use mechanisms to keep people away from data.

Use Commvault Command Center™ to orchestrate the creation of volumes and restore data without requiring direct administrator access to sensitive data.

Commvault enables workload owners to enable **Privacy** which requires a passkey to perform restores effectively **preventing admin access to sensitive data**.

Commvault recommends limiting access to the Commvault cloud library or Amazon S3 buckets from Commvault *machine identities* <u>only</u>. Additionally, Commvault software includes **Data Privacy** controls to prevent administrators from viewing or downloading server backup data. Access to backup data requires a client-specific or tenant-specific passkey to authorize restores. Additionally, laptop users may enable Data Privacy and set a **personal passkey** to authorize restores and effectively prevent administrator access to laptop backup data.

Related documents: SEC08-BP05 Use mechanisms to keep people away from data.

Protecting Data In Transit

Workload data accessed by Commvault will be returned in an unencrypted format as AWS transparently decrypts the data for authorized identities. Data in transit protection ensures that data being transferred between the protected workload and Commvault, and between Commvault backup copies is kept protected on the network.

SEC09-BP01 Implement secure key and certificate management.

Commvault utilizes **Transport Layer Security (TLS) 1.2** or greater for all Commvault control plane communication and optionally data plane traffic. Commvault implements its own certificate authority (CA) and creates and self-manages the rotation of 2048-bit RSA-based **client certificates**.

Commvault utilizes TLS 1.3 (from Commvault Platform Release 2022e) for communication between Commvault entities and utilizes per-client certificates stored securely within the Commvault database. Commvault **authenticates** all network communication with these certificates, rotates client certificates every 6 months (**configurable**), and rotates the CommServe® instance CA certificate every 5 years (**configurable**).

Commvault implements its own **Certificate Authority (CA) service** which runs on the CommServe® instance. Client certificates are 2048-bit RSA keys and matching RSA private keys are stored on the clients in AES245-encrypted envelopes and never transmitted over the network. Ciphers used to generate the client's private keys are **configurable**.

Commvault practices secure key management for encryption keys persisted in the Commvault database. Commvault **key generation** and **key lifecycle** are securely managed and have been third-party audited for use in FedRAMP, PCI-DSS, and SOC Type II environments (see **Certifications and Compliance**).

Related documents: SEC09-BP01 Implement secure key and certificate management.

SEC09-BP02 Enforce encryption in transit

Commvault enforces TLS 1.2+ encryption for all communication with AWS HTTPS service endpoints. Commvault uses **TLS 1.3** for all data management communications and may optionally implement **secure tunnels**, **software-based VPNs**, and **authenticated gateways** matching business encryption requirements.

Commvault encrypts all authentication and control traffic using TLS 1.3 by default (from Commvault Platform Release 2022e) with the replaced **TLS-RSA-AES256-CBC-SHA cipher suite**. To encrypt the data payload in transit a **network topology** is created within Commvault that directs all data traffic through an encrypted tunnel. After successful authentication, HTTPS encapsulated data traffic is encrypted with TLS-RSA-AES256-CBC-SHA cipher suite. Clients without certificates or revoked certificates can be rejected when attempting to communicate with locked-down CommServe® instances.

Commvault encrypts all control and data plane communication with Cloud databases (i.e., Amazon Aurora, Amazon RDS) when performing application-based export or dump-based backups. Commvault utilizes TLS 1.1/1.2 to securely connect, backup, and recover PostgreSQL, MySQL (inc. MariaDB), Microsoft SQL Server, and Oracle databases.

Commvault recommends that any network communications between Commvault® modules routing over public internet space be encrypted to ensure data security. This is achieved using mutually authenticated secure sockets layer (MA-SSL) and optionally leveraging Commvault® firewall configurations (Two-Way and One-Way).

Commvault supports and recommends the use of **AWS PrivateLink** to keep communication between your VPC and on-premises locations secure. Using AWS PrivateLink, you can access AWS services and services in other AWS accounts as if they were inside your private network, even if the accounts have overlapping IP CIDRs. See **What is AWS PrivateLink?** for more information.

See **AWS services that integrate with AWS PrivateLink** to ensure you have secured all the AWS services you are using.

AWS PrivateLink keeps traffic on the Amazon backbone and helps maintain compliance with industry regulations like HIPAA and EU/US Privacy Shield.

Commvault provides pre-defined topologies to simplify and promote a *secure by default* approach to the handling of data in transit. Pre-defined topologies are provided for one-way direct communication, one-way communication forwarded through a network gateway, two-way direct connection, and cascading connections.

External certificate authorities (CAs) are not supported for Commvault per-client authentication certificates.

External certificate authorities for Commvault public-facing web services are supported. Consider Amazon Certificate Manager, and Amazon Private Certificate Manager for certificates security the end-user facing web services (Webconsole, Webservers, API endpoints)

Related documents: SEC09-BP02 Enforce encryption in transit.

Commvault **Anomaly Alerts** will identify clients that are offline and/or have file activity indicative of malware or ransomware for investigation.

Commvault recommends using augmenting Commvault monitoring with Amazon GuardDuty monitoring for **Amazon S3 Exfiltration**, and optionally Amazon VPC Flow Log alerts indicating transfers to suspicious accounts.

Failed client authentication attempts will be logged by Commvault software in the cvfwd.log and appear as the following log entry:

```
Error Code 9:90 Authentication failed for host [...]. Network password does not match.
```

Commvault recommends installing the **Amazon CloudWatch agent** on all Windows and Linux infrastructure and using the **Logs section** in the agent config file to collect parse and forward authentication failure to Amazon CloudWatch for automated action via **Amazon EventBridge** and/or alerts via **Amazon Simple Notification Service (SNS)**.

Additionally, **Amazon VPC Flow Logs** can be used to analyze, detect, and alarm abnormal connections, both successful and denied. Communication incoming to the Commvault known TCP/IP ports 8400-8403 should only occur from Commvault infrastructure.

Related documents: SEC09-BP03 Automate detection of unintended data access.

All network communications in Commvault utilize Transport Layer Security (TLS) 1.2 or greater. TLS 1.2 is used when communicating with AWS API endpoints. TLS 1.3 is used between Commvault components.

Commvault authenticates all network communications using the TLS 1.3 protocol with the replaced TLS-RSA-AES256-CBC-SHA cipher suite. Network authentication enforces that the client has a valid Commvault CommServe™ CA generated certificate. Client certificates are automatically renewed every 6 months (**configurable**). Commvault internal CA rotates or renews its certificate every 5 years (**configurable**). Client certificates can be revoked to prevent communication with Commvault infrastructure, for example when malware or ransomware-like activity is detected on the client.

All communication to **AWS service endpoints** uses HTTPS only.
All dial-home communication with **Commvault External URLs** uses HTTPS only.

Additionally, Commvault encrypts all control and data plane communication with Cloud databases (i.e., Amazon Aurora, Amazon RDS) when performing application-based export or dump-based backups. Commvault utilizes TLS 1.1/1.2 to securely connect, backup, and recover PostgreSQL, MySQL (inc. MariaDB), Microsoft SQL Server, and Oracle databases.

HTTPS Proxies

Commvault supports the use of authenticated HTTPS proxies to reach the internet.

Ensure that all HTTP(S) proxies used between your MediaAgents and Access Nodes to reach the AWS service endpoints enforce authentication. Be aware that HTTPS proxies may have a performance impact on backup/restore operations.

Where possible, Commvault® MediaAgent instances should be configured to have direct network access to the Amazon S3 service endpoint for optimal performance.

Related documents: SEC09-BP04 Authenticate network communications.

Incident Response

Security threats are continually evolving and even a robust security approach, detection capability, and data protection controls may not mitigate emerging threats. A mature incident response capability ensures that your Security Operations (SecOps) teams can quickly and efficiently respond to new incidents, reduce, or contain the incident, then perform a forensic investigation on the root cause to develop further controls to protect the business.

Consider the following approaches to developing and refining your incident response capabilities:

Educate

To respond effectively when a security incident occurs, your SecOps teams, security specialists, and event first responders will require continuous education. Commvault provides a broad range of educational offerings from self-paced web-based learning to classroom workloads (see **Education Advantage**).

Your data protection specialists should hone their data protection security skills via the **Commvault® Expert** course, and the **Commvault Workflow Fundamentals** course that assists in automating response and forensic data gathering.

Commvault recommends incident response **AWS gamedays** that allow SecOps and infrastructure/platform operations specialists to work collaboratively to discover, contain, and most importantly *recover from simulated security incidents*.

Prepare

Gameday learnings and operational procedures (runbooks, playbooks, automation) should be stored in a centralized secure knowledge management system that is highly available ensuring the accessibility of key processes during incidents. Knowledge management systems should be protected by Commvault Backup & Recovery and written to more than one location for resiliency from a region-wide event.

Additionally, access to the tools, workload applications, and Commvault Backup & Recovery software should be pre-provisioned, documented, and tested before the security incident occurs.

Related documents:

SEC10-BP01 Identify key personnel and external resources.

Commvault recommends maintaining **an incident response contacts list** with contact details for key internal resources, external support resources, and any third-party partner security SMEs.

Ensure the contact list is kept secure but available to resources that require access during a high-risk event.

Your runbooks and playbooks used during a security incident should include key personnel (technical) and stakeholders (application owners, senior leadership). Teams should utilize a centralized system of record for incidents that tracks the incident timeline and activities performed. Commvault can automatically create incidents to reduce response time, see **Creating Incidents on ServiceNow** for an example.

Your security response should consider using your third-party providers and partners to provide an external perspective on your security posture and response. Commvault has a broad range of **Alliance and Technology Partners**, **Global Service Integrators**, and **Managed Service Providers** with extensive cybersecurity experience. Use the **Commvault Find a Partner** and **AWS Partner Network** to locate a partner with the competency to meet your business and industry security needs.

Related documents: SEC10-BP01 Identify key personnel and external resources.

SEC10-BP02 Develop incident management plans.

Document and socialize a **Security Incident Plan** or **Playbook** that details the triage, response, and recovery steps for an incident.

Ensure that playbooks include detailed steps to execute, and the permissions or accounts required to perform the steps.

Incident response plans can start as *manual playbooks* and as hotspots are identified can utilize an automated approach using **Commvault Workflows**, **Business Logic Workflows**, or event-based automated response using **Amazon CloudWatch alarms and event-state changes**.

Incident management plans should ensure that the longer an incident is active, it is escalated to higher levels of management for visibility and action. Commvault alerts and notifications, which are used to generate service incidents can **automatically repeat and escalate in priority**, this allows different rule handling based on increasing criticality.

Incident response should consider what data needs capturing before any remediation. Use Commvault Backup & Recovery to safely capture system state in cloud-native AWS snapshots, and optionally streamed backup copies before deleting or recreating affected workloads. Commvault recommends using **AWS Resource Tags** to flag resources that are removed from Production or intentionally isolated from end-user access. Commvault will discover, protect, and recreate tags on protected AWS workloads, meaning any accidental replication or cloning of affected resources will also be automatically isolated by incident management processes.

Related documents: SEC10-BP02 Develop incident management plans.

SEC10-BP03 Prepare forensic capabilities.

Incident management will include both forensic backup to the affected environment, and recovery of the environment to a known good baseline before the event.

Ensure the Amazon CloudWatch Log agent is installed and forward workload OS and application logs to CloudWatch. Backups may be protected by writing backups to an immutable Amazon S3 bucket with **S3 Object Lock enabled**.

Commvault recommends using Commvault Backup & Recovery to collect a forensic *snapshot* of any affected workload before any automated remediation is applied. The forensic response should aim to capture the current state (including OS and application log files) and then determine if further analysis will occur on the 'live' instance or immediately terminate the workload and recover from the Commvault backup copy. Commvault cross-account protection and recovery allow the recreation of the affected workload in a dedicated security account with customized networking controls to reduce further infection risk.

Related documents: SEC10-BP03 Prepare forensic capabilities.

SEC10-BP05 Pre-provision access.

Ensure that **break glass in case of emergency** AWS accounts exist for security personnel to triage and remediate security incidents without delay.

Incident response teams must have access to affected workload backup copies using Commvault role-based access controls (RBAC). Commvault audits all login and logout activity in the **Audit Trail** log and Amazon CloudWatch can be used to collect, parse, and alert on the usage of high-privilege SecOps accounts and activities within Commvault.

Backup copies can be used to recover the entire affected workload to a known good state and to baseline the normal operation of the workload before the event.

Related documents: SEC10-BP05 Pre-provision access.

Ensure the security personnel have systems and tools pre-deployed to assist in triaging a security incident. This includes ensuring a robust **tagging strategy** is applied to all workloads, to help inform the urgency and response to a security incident.

Security Operations should have forensic tools and appropriate AWS IAM policies to perform forensic analysis from an AWS Organizations security services OU and set of trusted accounts. Commvault recommends an intelligent inspection and anomaly detection be employed to reduce the likelihood of false positives. Commvault provides several **pre-defined anomaly-based alerts** that can reduce the amount of event or alert-based notifications that often result in alert fatigue.

Commvault recommends pre-deploying **AWS Systems Manager Agent (SSM Agent)** so that automated actions can directly interact with the operating system of any affected instances.

Related documents: SEC10-BP06 Pre-deploy tools.

Simulate

Simulating real-world security incidents provides a safe space to learn and hone your security response skills before a real event occurs. Consider the following approaches:

Commvault recommends regular **game days** to validate security incident response and the operation of the Commvault data management plan during mock **incidents** and **events**. Ensure learnings are recorded for team enablement.

Game days are simulations or exercises to test established runbooks and playbooks used to respond to security events. AWS publishes **Incident Response Runbook templates** that should be augmented to include validating workload backup status, taking a workload incident backup, and recovering use backup copies.

Consider the intent and relative risk of your **Security Incident Response Simulation (SIRS)**, events testing security events and recoverability of your Commvault Backup & Recovery from **Disaster Recovery backups** should be executed safely in pre-production environments to limit the impact on the business.

Game days associated with Backup & Recovery should focus on:

- *Workload recoverability* to ensure business services can be quickly recovered.
- *Commvault infrastructure recovery* to ensure shared data protection resources like **CommServe® instance**, **MediaAgents**, and **Access Nodes** can be recovered quickly and efficiently before workload protection is affected

Commvault data integrity cannot be intentionally or unintentionally breached and ensure backup and archival data cannot be deleted or corrupted.

Related documents: SEC10-BP07 Run game days.

Iterate

As more real-world events and gamedays are executed you will gain a detailed understanding of the processes and tools that deliver the most value. Look to fully automate your incident response from initial containment, data gathering, eradication, and recovery.

While some environments may allow for manual recovery by authorized application owners (i.e., dev/test systems), you may instigate an automated **Disaster Recovery planned failover** or automated *recreate and recover from a backup* restore for unavailable critical production applications.

SEC10-BP04 Automate containment capability.

Consider an **automated response to identified anomalies or potential ransomware events** that immediately isolate a resource from the network, power down the instance, and potentially protect the resource before deletion.

Containing an incident involves preventing further infection by removing or restricting access. Commvault **File Activity Anomaly Report**, **Unusual File Activity Dashboard** identifies and alerts on file activity that indicates a potential infection.

Commvault **notification types** allow forwarding these events to the Windows event log (where Amazon CloudWatch can collect and forward them), sending SNMP traps, sending email alerts, or running a custom **command-line script** or **Commvault workflow**.

Commvault recommends that the first step in *evidence gathering* be an ad-hoc backup of the affected workload. Using a cloud-native AWS snapshot, the workload state can be captured before any changes are made.

If incident containment processes (see below) dictate the removal of all snapshots to avoid recovering an infected or compromised workload, Commvault can take a *backup copy* of the snapshot before removal. Application-integrated *database export* can also be performed to capture the entire database state before further recovery actions.

Incidents can be automatically contained by using Commvault or AWS CloudWatch and Amazon EventBridge to automate security group modifications limiting end-user access.

Related documents: SEC10-BP04 Automate containment capability.

Eradicate the incident
Commvault recommends after forensic activities are completed that the workload is deleted so that unintentional infection cannot reoccur. Additional network or AWS IAM controls should be implemented to reduce the attack surface and workload accessibility in the event of reinfection.

Recover from the incident
After the infected or breached workload has been deleted, it should be recovered from the last known good Commvault backup copy. Depending on the extent of the breach, the entire workload and dependent workloads or services may need to be recovered.

Post-incident debriefs
Once the workload is recovered and business services resume, a Post Incident Review (PIR) is recommended to review learnings and identify areas for improvement. Using your service desk or incident management system here is important to understand the *end-to-end timeline* of the event and identify areas that could benefit from automation to speed recovery.

Reviews may uncover increased business reliance on a workload that results in an architectural review and implementation of **Commvault Auto Recovery (formerly Disaster Recovery)** to provide pilot light/**warm site recovery** or full active/active **hot site** protection.

Additional Resources

- **Security Pillar: AWS Well-Architected**
- **NIST Special Publication – SP 800-61 Rev. S Computing Security Incident Handling Guide**

- **Commvault Ransomware Protection**
- **Amazon Web Services Permission Usage**
- **Commvault User Administration and Security**
- **Commvault Network Topologies**
- **Commvault Encrypting Backup Data**

Reliability Pillar

Designing Commvault Backup & Recovery data management platform for consistent reliability provides *recovery readiness* for your application workloads. Designing for reliability ensures you can operate and continually test your Commvault data management platform to ensure your business protection SLAs are always met.

Reliability can be increased through automated recovery, regular testing, scaling to meet increased data size, right-sizing to observed workload size, and managing change in a repeatable validated way. Keep the following key capabilities in mind when designing for reliability:

- **Resiliency** is the ability of a workload to automatically recover from infrastructure or service disruption. All elements of the Commvault data management platform may operate in resilient deployments that can automatically recover from partial or complete resource outages.

- **Availability** is a measure of the percentage of time a workload is **available for use**. Commvault may be deployed in configurations ranging from 99% availability (per year) to 99.95% availability, including periodic maintenance and disruptive software updates.

  - ○ **Calculating workload availability**
    Availability is measured as the time Commvault is **_available for use_** to perform backups or recovery, divided by the **_total time_** including both scheduled and unscheduled interruptions to service.

    $$Availibility = \frac{Available\ for\ use\ time}{Total\ time}$$

  - ○ **Measuring availability based on requests**
    You may also measure availability based on the number of successful vs. failed data management activities. This method is useful in reporting Service Level Agreement (SLA) adherence but is less useful when the Commvault system is unavailable, preventing measurement. Commvault performs this calculation based on the SLAs set on your protection plans, see **SLA (Service Level Agreement) and Strikes**.

  - ○ **Calculating availability with hard dependencies**
    Commvault is dependent on many other AWS and potentially on-premises services that affect its performance and ability to perform data management within business service levels. Where hard dependencies exist between systems the overall achieved system availability may be calculated as the product of each dependent system's availability.

    *For example, if we consider that **Commvault** infrastructure is dependent on the underlying **network** and the ability to resolve protected workload hostnames and IP addresses using **DNS**, and the ability for admins and users to login via **AWS Identity Center** to perform self-service data management:*

Availability $_{\text{Commvault}}$ x Availability $_{\text{network}}$ x Availability $_{\text{DNS}}$ x Availability $_{\text{Identity Center}}$

$$99.95\% \times 99.995\% \times 99.995 \times 99.9 = 99.84$$

> ⓘ **Note**
>
> Availability measures in the calculation above are for demonstration purposes only and do not reflect actual Commvault or AWS service availability.

Ensure you consider all services and their stated or measured availability when designing your Commvault data management platform resiliency.

- ○ **Calculating availability with redundant components**
  Commvault can be deployed with redundant components. For example, **MediaAgent grids** deliver resilience for data management, replication, and pruning activities. **Access Node groups** provide load-balancing and restart ability for backup and recovery activities.

  When multiple independent, redundant components are available and distributed across multiple Availability Zones (AZs) for resilience, theoretical availability is calculated as 100% minus the product of the component failure rates.

  Availability$_{\text{effective}}$ = Availability$_{\text{max}}$ − ((100% - Availability$_{\text{dependency}}$) x (100% - Availibility$_{\text{dependency}}$))

  $$99.9999\% = 100\% - (0.1\% \times 0.1\%)$$

  Care should be taken to consider where a component is truly redundant. If the business SLA for backup time or recovery time cannot be met in the degraded state then the total number of redundant components (MediaAgents, Access Nodes) must be increased to meet SLAs during the degraded state.

- ○ **Calculating dependency availability**
  Some services and third-party providers provide a **Mean Time Between Failure (MTBF)** and **Mean Time to Recover (MTTR)** instead of publishing guaranteed availability time. To estimate the availability of these components uses the following formula:

  $$Availability\ \textbf{estimate}\ = \frac{MTBF}{MTBF + MTTR}$$

  For example, if the MTBF is 180 days, and the MTTR is 2 hrs, the availability estimate is 99.95%.

- ○ **Costs for availability**
  Your Commvault data management platform will need to provide an availability level equal to or ideally better than your most critical business application.

  Ensure that a cross-functional risk management team assesses and approves availability requirements in your organization, as availability needs increase so too does the cost and complexity of the workload. Delivering high levels of availability requires investment in automating failover activities and automatically monitoring and scaling resources to meet business SLAs.

- **Disaster Recovery (DR) Objectives** are a measure of acceptable data loss (recovery point objective) and recovery time (recovery time objective) when a large-scale event affects your workload. Commvault may be deployed in multi-AZ multi-region deployments to provide Disaster Recovery coverage in-region and across the region.

It is also key to remember that your availability needs for your workloads will change over their lifetime and even within a given year or month (i.e., financial reporting applications are business critical at the end of the financial year).

Foundations

It is important to ensure that your workload has the resources and environmental dependencies to deliver reliable services to the business. Consider the following foundational environmental components when designing for reliability:

Manage Service Quotas and Constraints

AWS sets service quotas (also referred to as service limits) on components that affect the performance and reliability of shared services. Quotas are intended to prevent accidental over-provisioning of resources and may often be updated by logging a **quota increase request** to Amazon.

REL01-BP01 Aware of service quotas and constraints

Review the **AWS Service Quotas** and the relevant service quotas in the **Commvault Design Principles and Best Practices** to understand the impact of quotas on data management activities.

Commvault recommends using AWS Trusted Advisor to regularly perform service quota checks and proactively increase quotas that are reaching exhaustion. Commvault deploys its infrastructure on Amazon EBS gp3 volumes to allow provisioning of performance independent of volume capacity. Commvault recommends using **Amazon EC2 Optimizer** to regularly review Commvault EC2 instances and EBS volumes for performance right-sizing advice.

Service	Resource quotas and limits
Amazon Aurora	**Amazon Aurora service quotas**
Amazon DocumentDB	**Amazon DocumentDB service quotas**
DynamoDB	**DynamoDB service quotas**
Amazon EBS	**Amazon EBS service quotas**
Amazon EC2	**Amazon EC2 service quotas**
Amazon EFS	**Amazon EFS service quotas**
Amazon FSx	**Amazon FSx service quotas** (includes Lustre, ONTAP, OpenZFS, and Windows)
Amazon RDS	**Amazon RDS service quotas**
Amazon Redshift	**Amazon Redshift service quotas**
Amazon S3	**Amazon S3 service quotas**

Service	Resource quotas and limits
Amazon S3 Glacier	**S3 Glacier service quotas**
Amazon VPC	**Amazon VPC service quotas**
IAM	**IAM Service quotas**
AWS KMS	**AWS KMS Service quotas**
AWS STS	No published service quotas or limits.

Source: **Service endpoints and quotas**

Service limits range from a maximum number of provisioned items, and concurrent operations, to rate limits. Commvault provides documentation on the relevant service quota limits if relevant to backup and recovery performance, for example, **EBS Direct API Backups for Amazon Web Services** and `GetSnapshotBlock` requests per account.

Related documents: REL01-BP01 Aware of service quotas and constraints.

REL01-BP02 Manage service quotas across accounts and regions.

Review consumption of **AWS Service Quotas** across AWS accounts, regions, and environments (dev-test, production). Ensure that changes made in pre-production are reflected in production and vice-versa.

Commvault will log failures related to exhausting service quotas during backup and recovery operations.

Service quotas are tracked per account and are mostly AWS region-specific. To deliver consistent workload protection across accounts, regions, and production and non-production accounts you need to ensure quotas are identical for each protected workload.

Related documents: REL01-BP02 Manage service quotas across accounts and regions.

REL01-BP03 Accommodate fixed service quotas and constraints through architecture.

Review the **Commvault Design Principles and Best Practices** service quotas information for the quotas that impact data protection operations. Pay particular attention to resource quotas that do not permit tuning, like the maximum number of concurrent EBS snapshots creations per-account, per region.

Some service quotas cannot be increased on request. Carefully review service quotas relating to concurrent connections, baseline, and burstable bandwidth, and the maximum number of snapshots so that your workload protection can be designed to accommodate fixed service quotas. Quotas that cannot be modified will often require distributing the workload across multiple AWS accounts to avoid hitting the quota.

Commvault helps mitigate maximum snapshot limits by copying snapshots out to Commvault cloud storage and then optionally removing the snapshot. Commvault operations can be manually or automatically balanced across the backup window to accommodate fixed throughput quotas.

> **Pro-Tip**
>
> Commvault requires a **Change Block Tracking (CBT)** EBS snapshot to be held for each protected Amazon EBS volume between backups, to provide the ability to identify incremental changes from the last backup. These snapshots will have an **AWS Resource Tag** of "Name=CV_CBT_SNAP". If these snapshots are deleted, the Access Node must download all blocks in the volume, CRC the blocks, and compare them to previously protected blocks. This process is computationally intensive and increases backup times, increasing the EC2 runtime costs of providing backup.

Related documents: REL01-BP03 Accommodate fixed service quotas and constraints through architecture.

REL01-BP04 Monitor and manage quotas.

Commvault recommends using **AWS Service Quotas** to manage your quotas for 100+ AWS services. Capture the default and current quota settings for each account and region and update as quota increases are applied.

AWS Trusted Advisor may also be used to view current quotas across accounts.

Monitor resource quotas interactively in the **Service Quotas console** and alarm consumption thresholds via **Amazon CloudWatch Service Quota Alarms**. You can also trigger service quota alarms from **AWS Trusted Advisor**.

Related documents: REL01-BP04 Monitor and manage quotas.

REL01-BP05 Automate quota management.

Commvault recommends automating the monitoring of quota consumption with tools like **Quota Monitor for AWS**. Ensure that notifications are forwarded into your service desk system for tracking of requests and assignment of the owner to request a quota increase.

Amazon recommends implementing automation such as **Automating Service Limit Increases and Enterprise Support with AWS Control** to automatically provision new resources within appropriate resource quotas set. Commvault backup and restore jobs will alert on the inability to create resources based on limits, allowing an automated **AWS SDK** or **AWS command-line call** to the `service-quotas` service.

Related documents: REL01-BP05 Automate quota management.

REL01-BP06 Ensure that a sufficient gap exists between the current quotas and the maximum usage to accommodate failover

Service quota increases are not automated or immediate. Determine your critical service quotas that could affect data management, availability, or performance and request an increase to accommodate temporary peaks or buffers to avoid service impact.

Ensure that service quotas consider resources that may fail and 'hold' quota resources for some time before the timeout and automated cleanup. Quota alarm thresholds should be set with a calculated overhead to accommodate failed resources and quota reclamation delay.

Related documents: REL01-BP06 Ensure that a sufficient gap exists between the current quotas and the maximum usage to accommodate failover.

Plan your Network Topology

Commvault will need to protect all your business workloads regardless of locality in the AWS region, AWS Local Zones, AWS Wavelength, and AWS Outposts or on-premises facilities. Availability of sufficient network connectivity, IP address space, and Domain Name Resolution (DNS) availability is key to continued operation. Consider the following network best practices when designing for reliability:

REL02-BP01 Use highly available network connectivity for your workload public endpoints

Commvault can be deployed with public web services (endpoints) using an **Amazon Elastic Load Balancer** (Application Load Balancer) and configuring **High Availability for Web Console and Command Center** to ensure high-availability access to end-users. You can failover your public-facing DNS for Commvault web services using **Amazon Route 53 health checks and configure DNS failover** as part of automatic CommServe failover and failback.

Commvault recommends using a highly available DNS service like **AWS Route 53** to publish business-facing URLs for accessing the Commvault Command Center™ console. Commvault recommends **High Availability for Web Console and Command Center** by load-balancing HTTP requests via an **Amazon Application Load Balancer (ALB)** with sticky session management.

Additionally, Commvault access to AWS services should utilize **AWS PrivateLink** endpoints to receive highly-available, secure, scalable access to AWS services without having to traverse an Internet Gateway (IGW) or NAT Gateway (NGW).

Related documents: REL02-BP01 Use highly available network connectivity for your workload public endpoints.

REL02-BP02 Provision redundant connectivity between private networks in the cloud and on-premises environments

Commvault recommends redundant network connectivity between on-premises and the cloud if business SLAs demand. Solutions can include multiple **AWS Direct Connect** connections, or Direct Connect and a **Site-to-Site VPN tunnel** as a backup.

Be aware that AWS VPN over the internet is limited to up to 1.26 Gbps throughput per VPN tunnel but does not support Equal Cost Multi-Path (ECMP) for outbound traffic.

Speeds may be limited to less than 1 Gbps during failover and impact the ability to perform backup and recovery services.

Where you are connecting your on-premises environments to AWS using **AWS Direct Connect (DX)** or **VPN tunnels**, ensure that redundancy exists both within and across regions (if required). Consider whether your backup and recovery service levels can be met in normal operation and with one or multiple connections in a failed state. Backups will occur using compression and deduplication for data transfer and restores will occur in a fully hydrated state (non-deduplicated) and incur a larger impact on your shared network resources.

Consider using the **AWS Direct Connect Resiliency Toolkit** to ensure you have redundant connections with equivalent speeds. When using different network technologies for normal operation and failed-over state, validate that network speed is equivalent or meets minimum throughput requirements for business-critical application backup & recovery needs.

Related documents: REL02-BP02 Provision redundant connectivity between private networks in the cloud and on-premises environments.

REL02-BP03 Ensure IP subnet allocation accounts for expansion and availability

Commvault recommends pre-planning and sizing your VPCs to accommodate for growth, emerging regulatory requirements for network segmentation, and integration with existing VPCs or accounts.

Consider making VPCs as large as possible as VPC CIDR blocks cannot be changed after creation. Be aware that if **AWS Transit Gateway** will be used, VPC subnets **cannot overlap**.

Remember that the Commvault central backup account will contain all Commvault infrastructure CommServe, MediaAgents, and auto-scaled access nodes.

When planning your IP subnet allocation consider the need to pre-allocate or reserve IP range for Commvault Backup and Recovery infrastructure. Typically, Commvault infrastructure will all reside in a single central backup account in either a single region or multi-region.

Commvault consists of a single **all-in-one configuration** on day one.

Expansion occurs into new regions using **MediaAgent grids** that start with a single host and may expand into a highly available grid of four (4) hosts capable of managing PB of backup data.

Commvault will dynamically provision ephemeral Access Nodes in the central shared backup account in each region and availability zone that Commvault is performing backup and recovery. A single Access Node can perform protection for 30TB of front-end TB (FETB) as measured by the workload being protected.

Access Nodes can protect all availability zones (AZs) in a region but should ideally be matched to the workload they are protecting to avoid paying network egress fees, so network planning should consider the total number of concurrent data management activities across all AZs in a region.

See the **Architectural Sizing** section for more details.

Related documents: REL02-BP03 Ensure IP subnet allocation accounts for expansion and availability.

REL02-BP04 Prefer hub-and-spoke topologies over many-to-many mesh

Commvault recommends (but does not require) **hub-and-spoke** topologies in multi-VPC deployments to simplify network interconnection and management.

VPC Peering, AWS Direct Connect, or VPN meshes may be created in very small environments but will quickly become complex to manage.

When you have a small number of Amazon VPCs to interconnect you may use **VPC peering**. As the number of VPCs, Amazon Direct Connect connections, and VPNs grows you should look to adopt a hub-and-spoke model like those supported by **AWS Transit Gateway**. The intent in adopting a hub-and-spoke model is simplified management.

Related documents: REL02-BP04 Prefer hub-and-spoke topologies over many-to-many mesh.

REL02-BP05 Enforce non-overlapping private IP address ranges in all private address spaces where they are connected

Commvault recommends pre-planning CIDR use across your VPCs and Subnets to allow routing between subnets in the future. Use service API operations to monitor subnet usage and consumption and migrate subnets to non-overlapping CIDR ranges if identified.

When using VPNs to connect networks with private addressing (**RFC1918 addressing**) the address space must not overlap. Implement an IP Address Management (IPAM) system to ensure IP address space uniqueness is considered and maintained as you expand.

Related documents: REL02-BP05 Enforce non-overlapping private IP address ranges in all private address spaces where they are connected

Workload Architecture

Consider the following best practices when designing your Commvault data management platform deployed in AWS:

Design Your Workload Service Architecture

When designing highly scalable and reliable workloads, care should be taken to design the workload so that it can integrate and scale with the other services within your organization. Commvault is typically deployed as an all-in-one configuration on day one and could be considered a monolithic application (consisting of a presentation, application,

and database layers in one instance) Commvault is a set of services (modules) with well-defined APIs for integration. Commvault is therefore a *service-orientated architecture (SOA)* with multiple software components (or **roles**) that perform independent data management tasks.

REL03-BP01 Choose how to segment your workload.

Commvault recommends deploying an <u>all-in-one</u> monolithic architecture on day one, with all components on a **single Amazon EC2 instance**. When the data management workload exhausts one or more resources (CPU, RAM, or network) scale data management roles horizontally onto a 1-4 node MediaAgent Grid with **Cloud Storage Pool**.

Commvault recommends scaling horizontally before considering vertical scaling, which is typically more expensive in the cloud.

Consider scaling and segregating task-based workloads onto dedicated resources that remain **powered down** when not actively required (i.e., search index nodes, eDiscovery nodes)

Commvault is a service-orientated architecture (SOA) that consists of multiple services involved in performing data management for your workloads. These services are deployed together on day one but can be segregated and scaled for performance and availability as your data management needs grow.

Related documents: REL03-BP01 Choose how to segment your workload.

REL03-BP02 Build services focused on specific business domains and functionality.

Commvault recommends <u>consolidating</u> services until business demand requires additional compute, storage, or networking resource. Consider a *data classification / business-criticality* approach to building segregated business domains. This will allow scaling services and resources based on data value, while less critical workloads will experience queueing if the adequate resource is not available.

You can create logically separated business domains by creating separate MediaAgent grids and/or Access Node groups that are assigned to specific business functions or workloads.

Commvault data management services include, but are not limited to:

- Configuration & Job Management.
- Reporting.
- Automation & Orchestration (Workflows).
- Hypervisor Protection.
- Database Protection.
- File & Folder Protection.
- Indexing.
- Replication.
- Disaster Recovery.

These services align with specific data management tasks and business objectives, you can centralize or decentralize each of these services. For example, the *Indexing* service is used to classify data, identify PII data, and respond to internal and external discovery requests.

You may choose to implement select services as *always online* (i.e., Configuration & Job Management, Reporting)**,** and some services as *always offline* until needed (i.e., Indexing, Replication).

This approach allows a specific reliability design for *data intelligence* that reflects the value to the business. By comparison, the *database protection* service may be considered business-critical due to the criticality of contained data, requiring additional reliability controls.

Related documents: REL03-BP02 Build services focused on specific business domains and functionality.

REL03-BP03 Provide service contracts per API

Commvault publishes its REST API at **api.commvault.com** and provides an interactive **<API explorer>** within Commvault Command Center™ for service integration teams to use for API syntax, versions, testing, and development. These APIs can be considered *service contracts* for consuming a specific Commvault role or service.

Service contracts are a documented agreement between teams on the service, the service machine-readable API definition, and any service levels provided or enforced (i.e., rate limits). Commvault keeps its REST API endpoints stable and versioned with each major platform release. This provides integration teams time to adjust to API enhancements. Commvault maintains **api.commvault.com** as the definition of the most current API version. A **Swagger-based API explorer** is available within the Commvault WebConsole for developers to interrogate and test the available API endpoints.

Related documents: REL03-BP03 Provide service contracts per API.

Design Interactions in a Distributed System to Prevent Failures

Commvault is designed to be deployed as a distributed data management system separated by local and wide-area networks. Latency or interruption to these networks will prevent communication of backup job success and failure but will not affect data management jobs from completing within the localized fault domain (i.e., an AWS Region).

REL04-BP01 Identify which kind of distributed system is required.

Commvault is a distributed data management system that can be considered a hard real-time system or an offline system depending on the activity. User interface requests within Commvault Command Center™ make requests to the AWS API and wait for responses before refreshing the interface. Alternatively, backup, recovery, and replication operations make requests to the AWS API and may wait minutes for a response confirming request completion or failure.

Commvault has multiple components that can be distributed but some components require hard real-time responses (synchronous communication), and some only require soft real-time responses (asynchronous communication).

Typically, components that are utilized during a backup activity should be co-located and operate with the assumption of low-latency real-time response, an example includes *deduplication hash lookup and updates for backup activities*. Deduplication lookups should occur with low latency and deduplicating MediaAgents should reside in the same region as the protected workload.

Alternatively, **CommServe LiveSync** which performs Commvault internal database backup, replicate, and recovery can be deployed in a many-to-one configuration. In this configuration, Commvault will expect soft real-time responses but will tolerate interruptions in the network that delay communication.

Related documents: REL04-BP01 Identify which kind of distributed system is required.

REL04-BP02 Implement loosely coupled dependencies.

Commvault implements loosely coupling interaction with the AWS API via a pool of Access Nodes. This is referred to as **Synchronous Loose coupling**. Once a backup, recovery, or replication event is scheduled to a coordinator access node, it distributes the request to a worker node. If the request fails, the coordinator will reschedule the request to a healthy worker node. This is referred to as **Asynchronous Loose coupling**.

Commvault supports loose coupling web-based services such as the Commvault Command Center™, WebConsole, and REST API endpoints. An Elastic Load Balancer (ELB) can be used to direct requests to healthy EC2 instances to mitigate failure in a single web component or availability zone.

Likewise, Commvault MediaAgents and Access Node components implement pools of highly available resources that receive incoming client requests and direct them to the most available, most performant instance.

Conversely, workload protection and recovery are *tightly coupled* in that each API request to the AWS service and/or workload must be acknowledged before the next recovery action is attempted. See the Reference Architectures section for details on how to deploy Commvault components across multiple Availability Zones (AZs) and/or Regions to improve resiliency and agility.

Related documents: REL04-BP02 Implement loosely coupled dependencies.

REL04-BP03 Do constant work.

Commvault implements a loosely coupled distributed data management system. MediaAgents perform data persistence to Amazon S3, and Access Nodes collect, optimize and transfer data to MediaAgents.

Isolation of an individual failing component is easily visible and does not interrupt the active data management jobs.

Related documents: REL04-BP03 Do constant work.

REL04-BP04 Make all responses idempotent.

Commvault API responses are not idempotent and do not return an idempotency token to prevent erroneous processing of requests multiple times. Commvault will wait for the HTTP responses and completion message before issuing an error indicating the failed action.

Commvault does not deliver services with **idempotent APIs**, APIs must be called only once to perform an action or change. If you are performing an integration that may call Commvault APIs multiple times, consider consulting the **Commvault Developer REST API Documentation** for idempotency token availability, and contact **dev-api@commvault.com** for further assistance.

Related documents: REL04-BP04 Make all responses idempotent.

Design Interactions in a Distributed System to Mitigate or Withstand Failures

REL05-BP01 Implement graceful degradation to transform applicable hard dependencies into soft dependencies.

Commvault backup and recovery actions have built-in **retry algorithms** that will re-attempt an operation multiple times before marking it failed, including tuning the **time between retries**. A data management job can be interrupted and restarted for backup **data protection operations** and **data recovery operations**.

See the **Job Restartability FAQ** for additional details.

Compute backup failure modes will gracefully select an alternative method to perform the activity rather than retry. For example, if an Amazon EC2 backup identifies that the appropriate Amazon EBS direct API permissions are unavailable, the backup will revert to a Commvault HotAdd transport type.

Commvault has a maximum amount of time a backup or restore job will continue to attempt retries, after which it will mark the job as *Completed with Warnings* or *Completed with Errors*.

Related documents: REL05-BP01 Implement graceful degradation to transform applicable hard dependencies into soft dependencies.

Commvault software performs internal **throttling of API requests to AWS** service endpoints as part of backup and restore activities. Commvault software tracks request rates against available **AWS service quotas** and `rate limit exceeded` errors will invoke a backoff algorithm before reattempting the failed action.

Additionally, Commvault has a central event engine called the *Job Manager* which is responsible for scheduling work based on configured **priorities** to available resources.

If insufficient Commvault resources are available, the Job manager will queue the request for data management until a resource is made available. Commvault does allow **throttling individual clients** based on **active jobs, data thresholds, and log data** thresholds. Commvault software will also **network throttle** individual clients, client groups, and MediaAgents based on schedule rules and/or measured Kbps network transfer rates. Commvault can **Throttle Throughput to Amazon S3 Storage** for specific backups, clients, or MediaAgents, including in specific **scheduled time windows**.

Related documents: REL05-BP02 Throttle requests.

REL05-BP03 Control and limit retry calls.

Commvault implements a non-**exponential backoff** when `rate limit exceeded` errors are experienced during backup and restore operations. Commvault will retry API operations multiple times (configurable) before marking an API call as failed so that a retry loop does not occur. Retries employ a custom backoff algorithm (non-**exponential backoff**) and backoff algorithms do not implement or introduce **jitter** into retry timing.

Commvault implements a *Total Job Running Time* (with restarts) and *Maximum Number of Retries* to ensure that job retries do not continue indefinitely and consume resources without ever completing. Once the maximum number of retries and/or total job running time is reached, the job is terminated, and any consumed resources are reclaimed.

Related documents: REL05-BP03 Control and limit retry calls.

REL05-BP04 Fail fast and limit queues.

Commvault will **fail fast** after a request is unable to be completed due to network interruption or multiple rate limit exceeded error. Failure will free up (clean up) resources created up to the failure, and then restart the operation if directed by the coordinator node.

Commvault software (when acting as the server side of a transaction) will always accept new requests for data management. If Commvault has insufficient resources to execute the job immediately, it submits the request to the *Job Manager* component to schedule the job when resources become available.

If Commvault has insufficient resources to respond to an API request for service, the client request will timeout and the client should retry the request after a backoff period. Commvault will fail fast when insufficient resources are available, it will fail and clean up resources created by the worker node and notify the coordinator.

Commvault promotes the use of queues as data management events are typically time sensitive and must be completed in the order they are received or following configured prioritization.

Related documents: REL05-BP04 Fail fast and limit queues.

REL05-BP05 Set client timeouts.

Commvault has built-in client-side timeouts when making requests for service against AWS service endpoints. Commvault sets these timeouts automatically but allows modification for specific access nodes, protection jobs, or AWS accounts.

The following are the more commonly tuned timeouts

Additional setting	Purpose
AMIImportTaskTimeoutInHour	Timeout for AWS VM Import/Export job in hours.
AMITimeoutInSec	Timeout for creation of Amazon Machine Image (AMI) during Amazon EC2 backup job (in seconds).
nAWSSocketTimeoutInSeconds	Timeout for AWS VM Import/Export job upload (PUT) requests (in seconds).
nAmazonValidateInstanceStatusTaskTimeOutInMin	Timeout for the total amount of time (in minutes) to wait for an Amazon EC2 instance status checks to return OK.
nAmazonValidateInstanceStatusPollingSec	Timeout or frequency of running an Amazon EC2 health check during recovery or replication jobs.
nAmazonVolumeTaskTimeOutInMin	Time to wait (in minutes) before marking an Amazon EBS volume create, modify, attach, or detach as failed.
nAmazonTaskTimeOutInMin	Time to wait (in minutes) before marking an Amazon EC2, EBS instance, or snapshot create, modify, attach, or detach as failed.
nAmazonSnapshotTaskTimeOutInMin	Time to wait (in minutes) before marking an Amazon EBS direct API snapshot creation and/or completion as failed.

Timeouts are applied to an Access Node (Amazon EC2 instance running Virtual Server Agent package) and apply to all jobs performed by that Access Node.

Related documents: REL05-BP05 Set client timeouts.

REL05-BP06 Make services stateless where possible.

Commvault is a *stateful* application that is responsible for capturing the state of protected applications and data. Commvault provides a stateless **RESTful API** that does not require the client to persist state but may require multiple API calls to obtain unique identifiers for created objects (i.e., `/create_backupclient`, `/get_backupclient_id`, `/executebackup/{backupclient_id}`).

Commvault is presented via multiple web-based services which require **sticky session management** to ensure the session state is maintained for the client. Commvault stores state in persistence engines relevant to the data management activity (i.e., SQL Server, MongoDB, C-Tree, Redis, and OrientDB databases).

Related documents: REL05-BP06 Make services stateless where possible.

REL05-BP07 Implement emergency levers.

Commvault provides **Activity Control** as an emergency lever to disable all jobs from all levels of a Commvault environment – CommCell, Server Groups, Clients, Agents or Instances, or Subclients. Alternatively, Commvault services may be forcibly shut down to prevent further access or interaction with Commvault services.

Commvault recommends implementing **Commvault Auto Recovery (formerly Disaster Recovery)** as an emergency lever for your protected AWS Amazon EC2 (**Virtual machine**), Amazon EFS and Amazon FSx (**File system**), Amazon RDS (**Database**), and Amazon S3 (**Object storage**) workloads. Commvault Auto Recovery provides an emergency-level response by failing over your application to an alternate Region or Availability Zone when issues are encountered. Failover and failback are automated after the emergency level is 'pulled' in the Commvault Command Center™.

Commvault can also be deployed in a highly-available configuration that allows **Disaster Recovery across Availability Zones and/or Regions**. CommServe LiveSync will maintain a replicated copy of your Commvault CommServe® instance and configuration in one or more failover locations for an emergency lever failover of your data management platform.

Care should be taken when performing an emergency lever failover of your data management platform to ensure that sufficient data protection resources (MediaAgents, Access Nodes) and cloud libraries (Amazon S3 buckets) are available in the new location to resume data protection. If recoverability is expected in the failed over location, backup copies from the primary location will need to be replicated and available in the DR location.

Remember that your emergency lever environment is doing LESS than normal production. You will be writing backups local to the Availability Zone or Region only. You will need to assess your emergency-level performance, capacity, and ability to meet business SLAs periodically. Consider whether you are required to deliver the same Service Levels in an emergency state, you may need to invest in automation to provision or scale-up resources only during emergency failover.

Related documents: REL05-BP07 Implement emergency levers.

Change Management

Managing change in your Commvault data management platform is key to being available to respond to recovery requests when required by the business. Change may involve planned software updates or configuration activities or unplanned resource spikes, security incidents, or the rollout of new features and functions.

Consider the following best practices in change management:

Monitor Workload Resources

Commvault provisions Amazon CloudWatch instance monitoring when deploying **Commvault via AWS Marketplace**. Commvault configures the following metrics and associated Amazon SNS email notifications to the Commvault administrator:

- `Logical % Free Space`
- `StatusCheckFailed_Instance`
- `StatusCheckFailed_System`

Commvault has built-in **Log Files**, **Resource metrics,** and **pre-defined alerts and notifications** to monitor and report the health of your Commvault data management platform. Logs may be forwarded to Amazon CloudWatch by installing the Amazon CloudWatch agent and configuring collection from Commvault log files stored in `<Commvault_software_install_path>/Log Files` folder.

Commvault monitors some key performance indicators (KPI) for adherence to Commvault best practices and alerts when best-practice thresholds are breached (i.e., **DDB query & insert time**). Commvault OS and Resource metrics are collected every 15 mins and written to the `<Commvault_software_install_path>/Log Files/ResourceMonitor` directory on your Windows, Unix/Linux Commvault client. Commvault does not monitor or alarm

ResourceMonitor metrics. Commvault recommends ingesting ResourceMonitor metrics into Amazon CloudWatch and using **Amazon CloudWatch Log Insights** to identify anomalous behavior. Ingesting ResourceMonitor metrics into CloudWatch will allow the baselining of normal behavior and alerting performance impacts before customers report an issue.

Commvault recommends using **Amazon CloudWatch** to monitor the availability, reachability, and performance of your Commvault data platform. Using CloudWatch decouples the monitoring and alerting of your workload from the workload itself, ensuring you are still notified if Commvault infrastructure becomes unavailable.

Monitoring of individual backup and recovery operations occurs within the **Commvault Job Monitor** and/or **Private Metrics Reports**. Commvault job metrics relating to backup, restore, and replication time is not able to be exported from Commvault CommServe Database (CSDB) for forwarding to Amazon CloudWatch. Commvault's pre-defined **CommServe Anomaly Alert** will detect and alert on job-related anomalies based on the number of run operations and the time taken to complete those operations. Commvault anomaly alerts can be forwarded to Amazon CloudWatch using a Commvault **command line** or **Workflow** notification type called the **aws cloudwatch** command-line sub-command.

Consider the following four (4) monitoring phases used within Amazon when designing your Commvault data management platform monitoring:

- **Generation – monitor all components for the workload**
  Commvault recommends utilizing AWS built-in monitoring tools to observe all components of your Commvault data management platform, including:
  - **Amazon EC2** published metrics for CPU, RAM, and disk usage.
  - **Amazon CloudWatch** OS metrics for CPU utilization, Memory utilization, voluntary and involuntary context switches, Disk IOPS, and Memory faults.
  - **Amazon VPC** flow logs for observed and achieved network throughput between components in normal and abnormal conditions.
  - **AWS Health Dashboard** for a personalized view of the performance and availability of the AWS services your Commvault data platform is dependent upon.

  Commvault user experience is dictated by the reachability and performance of web-based services like the Commvault Command Center™ console. Commvault recommends using **Amazon CloudWatch Synthetics** to create **canaries** that perform basic activities to validate reachability.

- **Aggregation – Define and calculate metrics**
  Commvault infrastructure and job log files, which may be collected using the Amazon CloudWatch agent, include the start and end times of all data management activities. Use Amazon CloudWatch filters to calculate elapsed time for backups, restores, and replication activities. The aggregated logs along with filtered output allow the creation of custom metrics and alarms when calculated results fall outside of acceptable thresholds. See **Creating metrics from log events using filters** for more information.

  Commvault Private Metrics provides aggregation of Commvault-specific performance and availability metrics. In environments where more than one CommServe® instance is orchestrating backup and recovery, Private Metrics aggregates reporting across all Commvault CommServe® instances. Commvault provides a **Reports store** where additional reports may be **downloaded**, including **Trending & Predication Reports** focused on long-term monitoring. Reports can be emailed to interested parties but cannot be forwarded in machine-readable form to Amazon CloudWatch for metrics tracking.

- **Real-time processing and alarming – Send notifications and automate responses**
  Commvault recommends aggregating your alert and notification handling within the Amazon Simple Notification Service (SNS) across all AWS accounts. Commvault can push logs, metrics, and events directly into **Amazon CloudWatch log streams/log groups**. Amazon CloudWatch can then forward critical alerts to Amazon SNS, which then forwards to any supported **Amazon SNS destinations** including email, SMS, external webhooks, and Amazon Lambda for custom handling or archiving.

  Commvault can send near real-time alerts on job completion, which can be forwarded into Amazon CloudWatch for aggregation, and analysis and then forwarded to Amazon SNS for notifications or automated remediation or action. Commvault **pre-defined alerts** include:
  - Backup job failed
  - Restore job failed
  - Aux copy job failed (replication)
  - Data aging job failed
  - The deduplication DataBase (DDB) reconstruction job failed
  - Disaster recovery job failed

- **Storage and Analytics**
  Commvault recommends using Private Metrics built-in reporting and Trending & Prediction Reports available in the Commvault Reports store to perform long-term analytics. Commvault does not currently provide a pre-defined method for Commvault backup and restore job metrics to be uploaded to Amazon CloudWatch. Once available in Amazon CloudWatch, you can use **Amazon CloudWatch Log Insights** to perform analytics on operational measures like availability, elapsed job time, data written, etc.

  Commvault recommends leveraging Amazon S3 Intelligent-Tiering for Amazon CloudWatch logs that are held and used for up to 90 days. Commvault recommends using frequent, infrequent-access, and active archive tiers <u>only</u>. If data must be kept for very long periods, consider using **Commvault Combined Storage Tier** for data placed within Amazon S3 Glacier Flexible Retrieval and Amazon S3 Glacier Deep Archive. Commvault Combined Storage Tier places minimal (1% of data stored) Commvault indexing data on frequent access tiers, allowing optimized retrieval using the **Cloud Storage Archive Recall** workflow.

REL06-BP01 Monitor all components for the workload (Generation)

Commvault recommends installing Amazon CloudWatch logs agent on all Commvault components to forward application Log Files to CloudWatch.

CPU, network I/O, and disk I/O metrics are recorded by Commvault in the **ResourceMonitor** folder and should be forwarded to CloudWatch for metrics extraction.

Focus on backup and restore throughput (GB/hr.) metrics that exceed an established known good threshold, and alert threshold breaches to your service desk for investigation.

Related documents: REL06-BP01 Monitor all components for the workload (Generation).

REL06-BP02 Define and calculate metrics (Aggregation)

Commvault recommends defining metric filters for specific terms and patterns in log data that should be tracked and alarmed (e.g., `rate limit exceeded`, `ERROR`, `Unauthorized`).

Related documents: REL06-BP02 Define and calculate metrics (Aggregation).

REL06-BP03 Send notifications (Real-time processing and alarming)

Commvault recommends creating real-time **alarms** that trigger when monitored metrics exceed established baselines.

Consider forwarding alarms to Amazon Simple Notification Services (SNS) so that alarms can be received by multiple subscribers and optionally archived for historical analysis per the business data retention policy.

Related documents: REL06-BP03 Send notifications (Real-time processing and alarming).

REL06-BP04 Automate responses (Real-time processing and alarming)

Commvault recommends automating responses to easily resolvable events using **EventBridge Rules**. Responses may start small, like generating an incident in your service desk for investigation, or may resolve the condition effectively by clearing the alarm.

For example, a CloudWatch log alert that indicates disk space has been exceeded on a DDB volume could trigger an increase in the EBS volume size, then a subsequent AWS Systems Manager Run Command could expand the operating system file system.

Related documents: REL06-BP04 Automate responses (Real-time processing and alarming).

REL06-BP05 Analytics

Commvault recommends collecting log files and metrics from your data management platform and protected workloads using **Amazon CloudWatch Logs Agent**. Commvault recommends **exporting log data to Amazon S3** for long-term retention and protection using **Amazon S3** backup.

You may then use Amazon CloudWatch Logs Insights to search and analyze the log data, or Amazon Athena to query S3 exports. Tools like New Relic, Splunk, and Logstash can also analyze Amazon CloudWatch logs.

Holding collected logs and metrics per your data retention policy will allow greater insight into broader trends across your data landscape.

Related documents: REL06-BP05 Analytics.

REL06-BP06 Conduct reviews regularly.

Commvault recommends regular reviews of monitoring practices and the identification of areas where insufficient metrics existed to resolve an incident.

Maintain multiple dashboards of key performance indicators (KPIs), one focused on business metrics (SLA, missed/met clients, SLA trends) and technical metrics (infrastructure availability, load, average throughput).

Commvault provides an aggregated business and technical view in **Commvault Command Center™ dashboards**, and the ability to drill down to targeted **SLA**, **SLA health,** and **infrastructure-level** reports (including **custom dashboards**).

Also, consider using **Amazon CloudWatch Dashboards** to build aggregated business and technical dashboards integrated with your wider application observability strategy.

Related documents: REL06-BP06 Conduct reviews regularly.

REL06-BP07 Monitor end-to-end tracing of requests through your system

Commvault centralizes end-to-end tracing of requests through its distributed architecture within the Commvault Command Center™ interface.

Authorized users can use the view logs job feature which aggregates all logs from each distributed component involved in the activity.

When forwarding logs to Amazon CloudWatch, end-to-end tracing may be performed by matching the **<job_id>** that prefixes all log entries across all component logs.

Related documents: REL06-BP07 Monitor end-to-end tracing of requests through your system

Design Your Workload to Adapt to Changes in Demand

Commvault is made up of multiple components that may be scaled to adapt to changes in demand. On day one you will most likely deploy all components on a single Amazon EC2 instance, Commvault refers to this as an **all-in-one configuration**.

As your data to protect and manage grows, you will separate your MediaAgent and Access Node resources, as these resources are responsible for data handling (backup, recovery, replication). Consider the following best practices when designing your Commvault data platform for changes in demand:

REL07-BP01 Use automation when obtaining or scaling resources

Commvault utilizes automation to **automatically scale** the required number of access nodes for backup or backup copy operations to meet the business RPO.

Commvault performs large parallelized GET/HEAD and PUT/POST/DELETE to Amazon S3 by default, Commvault uses a proprietary data layout and does require the creation of multiple prefixes to parallelize requests.

Commvault can utilize Elastic Load Balancing (ELB) to **distribute HTTP requests** in environments with large self-service user-base like desktop/laptop backup services.

Commvault does not implement or utilize **AWS auto-scaling** to respond to the demand for more resources. Commvault will automatically provision Graviton-based Amazon EC2 instances (**Access Nodes**) for backup activity, then terminate these instances when no longer used.

Commvault CommServe® instance and MediaAgent multi-node grids require active monitoring of the protected resources (instances, databases, laptops, etc.) and managed data volume (front-end TB, back-end TB) and then manually scaling Amazon EC2 infrastructure (**Change the instance type**) to meet business service levels for backup or restore time.

Commvault does not require CPU, RAM, and disk space matching the published T-Shirt sizing in Commvault documentation (**CommServe**, **MediaAgent**, **Access Nodes**). Commvault recommends customers measure and monitor key performance indicators like web-service responsiveness and elapsed backup and restore time, then scale resources accordingly. **Amazon Compute Optimizer** can provide indications of over-provisioned and under-provisioned Amazon EC2 and Amazon EBS resources.

> ⦿ Pro-Tip
>
> Commvault utilizes Amazon S3 to store backup data and **does not require** the creation of multiple buckets or prefixes (per bucket) to scale performance. Amazon S3 automatically monitors I/O across Commvault prefixes and will partition requests in the backend to optimize for GET/PUT activity in real-time.

Creating additional prefixes within Commvault buckets will not result in increased parallelization of reads or writes. Commvault increases the parallelization of backup and recovery operations based on the number of concurrent data management activities running concurrently.

Related documents: REL07-BP01 Use automation when obtaining or scaling resources.

REL07-BP02 Obtain resources upon detection of impairment to a workload

Commvault records the achieved backup and recovery throughput in historical job information and uses this information to auto-tune job scheduling and auto-scaled resources to meet business RPO. Commvault performs auto-scaling at the **start of a backup** job and will not scale additional resources if one or more auto-scaled resources fail or perform poorly. Unfinished jobs will be redistributed to the remaining healthy worker nodes.

Commvault does not automatically scale resources upon detection of impairment to a workload. Commvault resources (MediaAgent, Access Nodes) may be pre-deployed in highly available configurations that allow continued operation during the failure of one or more nodes. Commvault does not recommend pre-provisioning Amazon EC2 resource to proactively handle the additional load when one or more nodes is impaired.

Amazon CloudWatch can be used to perform monitoring of **Amazon EC2 instance status checks** for Commvault MediaAgent and Access Node resources and then automatically notify the administrator or create a service incident to manually deploy and register a new Commvault component when impairment is detected.

Related documents: REL07-BP02 Obtain resources upon detection of impairment to a workload.

REL07-BP03 Obtain resources upon detection that more resources are needed for a workload

Commvault performs auto-scaling at the **start of a backup** job and will not scale additional resources if backup throughput is not meeting the expected transfer rate. Backup throughput (GB/hr.) historical information is used to inform the scaling of the required number of Amazon EC2 resources on the next backup for the VM group.

Commvault considers this implementation a custom implementation of **ML-enabled predictive scaling** with a **scheduled scaling event** at backup initiation.

Commvault will automatically scale Access Nodes during backup activities to meet configured Recovery Point Objective (RPO) for the total data volume to be protected. Commvault cannot automatically scale resources during restoration activities. Commvault recommends monitoring key performance indicators like restore type, and replication time and proactively increasing Amazon EC2 instance size before impacting business service levels.

Related documents: REL07-BP03 Obtain resources upon detection that more resources are needed for a workload

REL07-BP04 Load test your workload

Commvault auto-scaling is automatically learning the achieved throughput of selected Amazon EC2 instance sizes during backup activities, effectively continually recording load test results in Production.

Commvault recommends using game days to temporarily provision cloned workloads for load-testing changes to your Commvault data management environment. Commvault recommends using Commvault's out-of-place restore capability to create on-demand production environment replicas to load-test and experiment with differing Amazon EC2 auto-scaling instance sizes.

Consult the **Performance Efficiency Pillar** for Commvault-performed test results on the recommended Amazon C7g family.

Testing should be repeated on each Commvault component in your environment to better understand the current load and impact of vertical and horizontal scaling. Commvault recommends continually right-sizing your Production environment based on recommendations from AWS Compute Optimizer. Load testing can be used to ensure that right-sized configurations will continue to meet business SLAs without impacting production workload protection, at a minimal cost.

Related documents: REL07-BP04 Load test your workload.

Implement Change

All changes to your Commvault data management platform should be controlled and reversible. Use the following best practices when implementing change to your platform:

REL08-BP01 Use runbooks for standard activities such as deployment

Commvault recommends developing standard **runbooks** as manual procedures for performing common tasks like restarting Commvault, validating operational and performant state, installing software patches, and reverting software patches. After successful execution and validation via gamedays, frequently used runbooks should be automated as code to optimize change execution and avoid human error.

Commvault recommends using **AWS CloudFormation** to perform all new deployments from Commvault-supplied **AWS Marketplace products**.

Commvault **does not support the rollback** of platform releases or maintenance releases, so a manual process for the software deployment/update, performing functional/non-functionality tests, and then a go-or-no-go decision and a subsequent **revert of Amazon EC2 instance from AMI** is recommended. **Ensuring rollback safety during deployments** is critical to ensuring that changes do not affect your ability to protect and recover your workloads, always evaluate all changes in pre-production environments before production rollout.

Operational Readiness Reviews (ORRs)

Operational readiness reviews are a final checklist to execute before transitioning a change into production. ORRs can include a review of test breadth, the ability to monitor and observe key performance indicators for the feature or function, and validation that the change allows SLAs to be met. ORRs are also a useful tool in ensuring your personnel is trained and aware of how to operate your Commvault data platform.

Related documents: REL08-BP01 Use runbooks for standard activities such as deployment.

REL08-BP02 Integrate functional testing as part of your deployment

Commvault recommends maintaining a list of mandatory functional tests that must pass whenever deploying or updating software. Functional tests may include validation that software services running, that's services are accessible using **check readiness**, and backup and recovery success without error.

> 💡 **Pro-Tip**
>
> Consider backup and restore time a mandatory **functional** test if a business SLA exists to achieve these tasks in an established period. Often backup time is considered an optional or nice-to-have **non-functional** requirement. Validate that business recovery times are achievable with each newly deployed change.

Functional tests must be implemented whenever automating a runbook. Functional tests may include:

- Validation that Commvault software starts and is accessible using check readiness.
- Validation that a Commvault backup and recovery is successful.
- Validation that a Commvault data management activity maintains consistent throughput.

Note that while historically a 'backup time' or 'restore time' may be considered a performance and efficiency test, the ability to complete data protection within a known time constraint is crucial to meeting business Recovery Point Objectives (RPOs), and is, therefore, a crucial functional test for all deployments.

This testing must be performed in a pre-production environment before the deployment of a change into production.

Related documents: REL08-BP02 Integrate functional testing as part of your deployment.

Commvault recommends that resiliency testing is performed regularly using the principles and practices of **chaos engineering**. Commvault components should be tested first in pre-production and then also in production via production-focused **game days**. Commvault recommends focusing on the following key components and resilience functions:

- Corruption or loss of the CommServe and **DR failover**.
- Corruption or loss of a DDB and **recovery**.
- Corruption or loss of a MediaAgent and **recovery**.
- Corruption or loss of an Access Node.

> ⚠ **Warning**
>
> Be aware that resiliency and corruption testing on your production deduplication databases (DDBs) may result in the **DDB being sealed**, resulting in increased backup and network data transfer on the next backup.

Related documents: REL08-BP03 Integrate resiliency testing as part of your deployment.

Commvault is a system of record for business applications and data and is managed largely as **mutable infrastructure** which is maintained in place by applying updates, security patches, and configuration changes to the production environment.

Commvault updates to CommServe, MediaAgents, and Access Nodes may be performed using the **immutable server paradigm** which applies updates by provisioning new infrastructure and migrating system-state. Commvault provides AMIs for an all-in-one CommServe instance and combined MediaAgent + Access Node deployments in the AWS Marketplace. Commvault can perform orchestrated **repave** of Commvault CommServe, MediaAgents, and Access Nodes using the built-in Workflow Engine (WFE). Contact your Commvault sales representative to discuss this solution.

Commvault data management directs workload data via an Access Node group, which then persists backup data via a MediaAgent Grid. **Canary deployments** should be used to validate changes to **Auto-scaling Access Node instance size and family** without impacting all protected workloads. Canary deployments will hard-code specific workloads to utilize the new access node group via a **custom hypervisor configuration**.

Commvault can implement a **Blue/green deployments** model to logically separate changes or deployments still under test. Commvault can automate the repave/refresh CommServes, MediaAgents, and Access Nodes for the purposes of software currency (patching) by decoupling and retaining persistent data and configuration and replacing the underlying OS drive. Speak to your Commvault sales representative about how Commvault can help you automate infrastructure refresh using Infrastructure as Code tools like AWS CloudFormation and HashiCorp Terraform.

 Related documents: REL08-BP04 Deploy using immutable infrastructure.

Commvault recommends using AWS CloudFormation with Commvault-supplied **AWS Marketplace AMIs** to automate infrastructure deployment. Commvault provides automated **software update download**, and automated **software installation** via targeted schedules.

Alternatively, Commvault installs and updates may be automated using **unattended installations** orchestrated via **AWS Systems Manager Run Command**.

Commvault can automate the download of software updates to the CommServe® instance and may optionally automate the deployment of updates to clients. Updates must occur in **phased deployment** with CommServe® instance, then MediaAgents, then Access Nodes and clients.

Commvault does not provide automation to execute software updates for the CommServe® instance, however, **command-line execution** is supported and possible using **AWS Systems Manager Run Command**.

Commvault can automate a push deployment of software updates to MediaAgents, Access Nodes, and clients using command-line, REST API, and/or Workflows. Additionally, Commvault configuration changes can be automated using the command line, developer SDK, REST API, and/or Workflows (see **Developer Tools**).

Feature flags (also known as feature toggles)

Commvault has a broad set of feature flags or advanced configuration options (referred to as **Additional Settings**) that can turn features and settings OFF/ON. These settings may be used to create *canary deployments* with specialized or isolated configurations for validation before a wider rollout. Commvault provides an **additional settings database** for querying and identifying additional settings related to AWS workload protection and configuration.

Fault Isolated zone deployment

Amazon will not touch or update more than one Availability Zone (AZ) at a time to ensure multi-AZ deployments maintain high levels of availability. Commvault recommends deploying CommServe Server®, MediaAgent Grids, and long-running Access Node groups across zones where in-region high availability is required.

Related documents: REL08-BP05 Deploy changes with automation.

Failure Management

Commvault Backup & Recovery provides your last line of defense in workload failure management. Commvault is used to recover your workloads when intentional or unintentional data loss occurs. Consider the following best practices, including your responsibility to **Back Up Data**, when designing for failure:

Back Up Data

Commvault and Amazon guidance is clear.

> Back up data, applications, and configuration to meet requirements for recovery time objectives (RTO) and recovery point objectives (RPO) **AWS Well-Architected – Reliability Pillar**

Consider the following data sources and data handling best practices for your backups:

REL09-BP01 Identify and back up all data that needs to be backed up, or reproduce the data from sources

Commvault recommends regular review and audit of the **AWS services** used by your workloads. Valuable workload data is often generated and stored across AWS **compute**, **database**, **volumes**, **shared file systems**, and **object storage**.

Consider that the logs and metrics recorded for your applications using AWS CloudTrail and Amazon CloudWatch cannot be reproduced if lost, potentially impacting security incident response.

Commvault recommends classifying data based on criticality to recovery. Apply classification as an **AWS Resource Tag** on workload resources, to ensure **Plans** will protect the data per business policy (RPO).

Commvault recommends using **Commvault Backup & Recovery** as your only backup solution to protect **AWS** and **edge-based** data sources in your business using native AWS snapshots and streamed backup copies.

Commvault recommends reviewing all your workloads to discover which AWS services are being utilized and require protection. Consult the **Commvault protection of AWS services** section to gain an understanding of the types of data that each AWS product/service generates and that Commvault protects.

Before excluding data from backup processes, consider the impact on the business of not having that data available. Additionally, for data that can be reproduced, consider, and estimate the cost to recreate the data, as the cost of storing a low-cost Amazon S3 backup copy may be far lower than the AWS service cost to regenerate the data.

Commvault recommends storing all primary backups on **Amazon S3 Standard Infrequent-Access (S3 Standard-IA)**. Data that requires long-term or regulatory retention should be auxiliary copied to a **Commvault Combined storage tier** bucket.

Consider the workload Recovery Time Objective (RTO) when selecting a backup storage location, storing operational backups in **Amazon S3 Glacier Flexible Retrieval** or **Amazon S3 Glacier Deep Archive** will reduce storage costs but can increase restore costs and restore time. Consider using S3 Batch Operations to increase the throughput available to restores from S3 Glacier at no additional recall (see **faster restores for S3 Glacier**).

On-premises backups can be written to on-premises Amazon S3 storage provided by **AWS Outposts**, scale-out storage appliances like **Commvault HyperScale™**, or written to Amazon S3 in the Region **directly** or via **AWS Storage Gateway**. Care should be taken when placing your primary on-premises backups in the Region as all recoveries will count towards egress charges to the internet.

> ⓘ Note
>
> Data transferred out to the internet for the first 100GB per month, aggregated across all AWS Services and Regions (except China and GovCloud) **Amazon S3 - Pricing**.

Data generated by other services and persisted to **Amazon S3**, **Amazon Redshift**, and other data storage services should also be considered. Consider backup for your human-readable and machine-readable infrastructure-as-code automation like AWS CloudFormation templates, by backing up your **GitHub**, Amazon S3, **Amazon EFS**, and **Amazon FSx** repositories.

Related documents: REL09-BP01 Identify and back up all data that needs to be backed up, or reproduce the data from sources.

REL09-BP02 Secure and encrypt backups

Commvault recommends, protects, and recovers AWS data sources that are encrypted using AWS-managed keys (**aws/s3**, **aws/ebs**, **aws/rds**, **aws/elasticfilesystem**, **aws/fsx**) and customer-managed keys in AWS KMS,

Note: Commvault does not support or utilize **FIPS endpoints**.

Commvault replicates encrypted **EBS** and **RDS** snapshots cross-account and cross-region, using **multi-region keys** or by re-encrypting content using pre-configured Commvault KMS key aliases (`cvlt-master`, `cvlt-ec2`, `cvlt-rds`).

Commvault writes data to Amazon S3 using server-side encryption using Amazon S3-managed keys (**SSE-S3**), KMS keys (**SSE-KMS**), customer-provided encryption keys (**SSE-C**), or using dual-layer server-side encryption with KMS keys (**DSSE-KMS**).

Alternatively, Commvault provides **software-based client-side encryption** using FIPS-140-3 certified cryptographic module ^{pending CMVP review}.

Commvault protects encrypted AWS workloads (i.e., encrypted Amazon EC2, Amazon EBS, and Amazon RDS) and maintains encryption of data in transit and at rest using Commvault, AWS KMS, or customer-managed keys (including via Amazon CloudHSM). Encryption can occur client-side or server-side, with server-side recommended to reduce CPU impact on protected workloads.

Access to backups is enforced by strict authentication to centralized identity stores using SAML 2.0, OIDC, or Active Directory. Authorization is granted with the least privilege to specific users, user groups, and protected workloads via role-based access. All activity on the Commvault data platform both successful and unsuccessful is written to the immutable Audit Trail.

Related documents: REL09-BP02 Secure and encrypt backups.

REL09-BP03 Perform data backup automatically.

Commvault recommends and performs backups automatically per the business-defined recovery point objectives (RPOs). The business backup policy is configured in **Plans** and automatically scheduled using A.I. and machine learning that adjusts to changing data volume and network performance to meet configured RPOs. Commvault data backup automation is provided for all supported **AWS data sources** and **edge-based** data types.

Commvault centralized **server plans** define the **frequency**, **retention period**, **lifecycle**, **backup copy destinations**, and **resources to protect**.

Commvault automatically discovers and protects workloads using **AWS Resource Tags**, Region Availability Zone, and many other metadata-based rules. As tags change, Commvault adjusts to writing to appropriate regions and storage classes. Commvault maintains AWS resource tags when replicating data across Regions and accounts, meaning tag-based managed practices follow the workload.

Commvault provides data lifecycle management for multiple AWS services (see below) by orchestrating the creation, sharing, and copying of snapshots within and across AWS accounts and regions. Commvault will re-encrypt data if copying a snapshot to a new Region, in data sovereignty-governed environments. Commvault also supports multi-region keys to avoid re-encrypting data when replicating snapshots across regions.

Commvault provides automated AWS snapshot-based protection for the following AWS products in the AWS Region or edge-based locations like AWS Local Zones and AWS Outposts:

- Amazon EC2 instances (including Amazon EBS volumes, including Amazon VPC resources)
- Amazon RDS instances (including Amazon Aurora serverless instances, including Amazon RDS custom)
- Amazon Redshift
- Amazon DocumentDB

Additionally, Commvault provides automated application-consistent backup for traditional applications running on Amazon EC2 infrastructure using IntelliSnap® snapshot management. **Commvault IntelliSnap for AWS** provides snapshot-based protection for a broad array of traditional databases, NoSQL databases, SAP HANA, and common file systems.

Backup automation also extends to the periodic backup, replication, and recovery of **Virtual Machines**, **Databases**, **File-systems**, **Object Storage**, and **Hadoop** data for Disaster Recovery within the AWS Region and from on-premises to the AWS Region.

Commvault Command Center™ provides an *equivalent API* button on most actions which allows the extraction of the Commvault REST API endpoint and JSON/XML payload required to execute the action. This capability can be used to automate common recovery tests and track success/failure and performance over time.

Related documents: REL09-BP03 Perform data backup automatically.

Backups are only useful if they can recover your workloads within the business-required Recovery Time Objective (RTO) and with the agreed Recovery Point Objective (RPO). Commvault measures and reports **Actual RPO (RPA)** and **Actual RTO (RTA)** to help verify the integrity of your recovery readiness.

Commvault recommends performing regular restore testing to an isolated test-only Amazon Virtual Private Cloud (VPC) location to ensure business recovery time expectations can be met. Notify or publish restore results for application owners and affected stakeholders.

Ensure that recovery testing includes long-term retention or archive data stored in **Amazon S3 Glacier storage classes** or ideally **Commvault Combined Storage Tier**. Ensure testing includes testing of S3 Batch Operations automated restores, see *faster restores for Amazon S3 Glacier* for more details.

Commvault Auto Recovery automates disaster recovery testing by providing **unplanned failover** and **test failover** for replicated Amazon EC2 instances.

Related documents: REL09-BP04 Perform periodic recovery of the data to verify backup integrity and processes.

Use Fault Isolation to Protect Your Workload

Commvault recommends deploying your Commvault data management platform in more than one isolated fault domain to limit the impact of infrastructure or network failures. Consider the following best practices when deploying your Commvault Backup & Recovery system for fault tolerance:

Commvault recommends and supports deploying CommServe, MediaAgents, and Access Nodes distributed across **Availability Zones (AZs)** (including **AWS Local Zones**), and **AWS Regions** (including GovCloud) for resilience. Commvault software may be built to rely on redundant components and automatically fail away from unhealthy resources. Distributing components across multiple locations allows Commvault to automatically recover from failure in an individual location.

Hybrid deployments spanning the AWS Region and on-premises infrastructure (i.e., **AWS Outposts**) are also possible.

Availability Zones (AZs)

AWS Regions of multiple Availability Zones (AZs) that are independent of one another, separated by sixty miles (100 kilometers). Availability Zones provide high-bandwidth low-latency private network connectivity with sub-millisecond roundtrip time latency. Commvault may be deployed across AZs to provide resilience from isolated failures like fire, floods, and tornados.
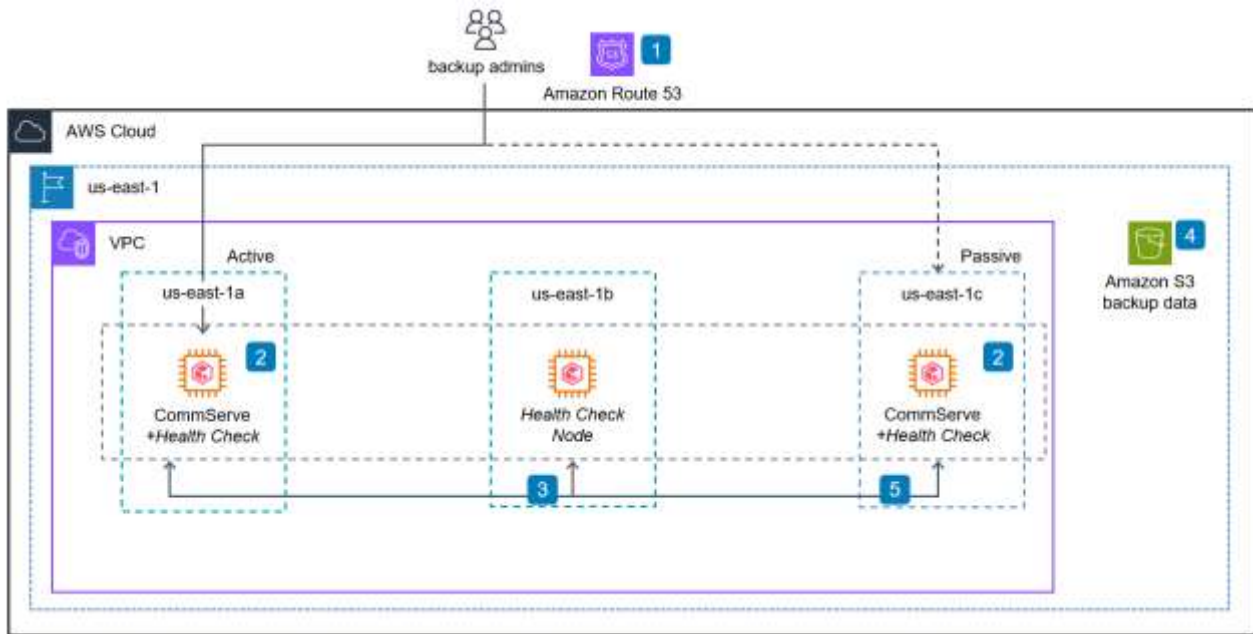
Commvault may be deployed across multiple AZs in an **active-standby** configuration:

1. **Amazon Route 53 DNS health checks** to **Amazon Route 53 health checks on private resources** monitor the health and performance of Commvault and route away from unhealthy resources.
2. **Commvault CommServe® LiveSync** maintains a single primary (us-east-1a) and a single replicated read replica (us-east-1c).
3. Commvault application-level health checks run across three AZs (us-east-1a, us-east-1b, us-east-1b) and perform **automated failover** and protection from split-brain scenarios.
4. Backup data resides within Amazon S3 within the region which is designed to deliver 99.999999999% (11 9's) of **data durability** of objects over a given year.
5. Commvault **replicates hard state** (configuration, job history, certificates, identities) periodically to the standby system

Additionally, if expanding into a new Region, Commvault MediaAgent grids and/or Access Node groups may be deployed across Availability Zones (AZs) for the resilience of backup and recovery operations:
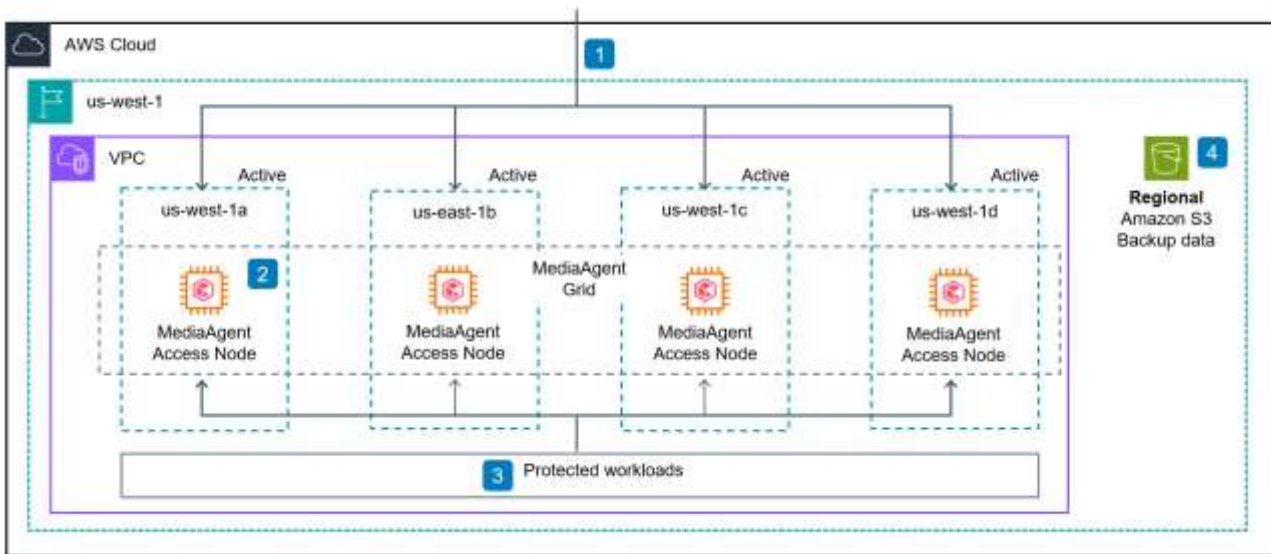
In a regional expansion, deploy an **active/active pattern** that provides compute, network, and block-based storage resilience:

1. **The Control plane** is remote with backup & recovery operations orchestrated from the **Commvault CommServe®** instance.
2. **MediaAgent grids** (min. 1, max. 4) provide AZ-independent regional backup, recovery, replication, and ongoing data management (data aging, pruning).
3. **Protected workloads** are accessed via AZ-specific VPC endpoints (gateway, interface).
4. **Backup data** is stored in the region within an Amazon S3 bucket, which is replicated across multiple AZs by the Amazon S3 storage service.

A failure of a single MediaAgent/Access Node will result in the data management activity being automated retried and rescheduled to another available compute resource.

AWS Local Zones

AWS Local Zones (**What are AWS Local Zones?**) provide compute, network, and storage services that can host Commvault active/active MediaAgent grids to perform backup and recovery of AWS Local Zones workloads. See **Architect with AWS Local Zones to meet data residency requirements and deploy low latency applications** for more details on the AWS services supported by AWS Local Zones.

> ⓘ Note
>
> Amazon S3 is not currently offered within AWS Local Zones, so backup data would need to transfer back to the Amazon S3 services in the parent region.

AWS Regions

A multi-region **active/standby** approach can be used to provide a Commvault data platform resilient to region-wide outages. This pattern is deployed as a *disaster recovery* solution where one-off large-scale events that affect an entire region occur. In this pattern, the Commvault *control plane* and *data plane* are failed over to an alternate region and must continue to provide data management for all protected and accessible regions.
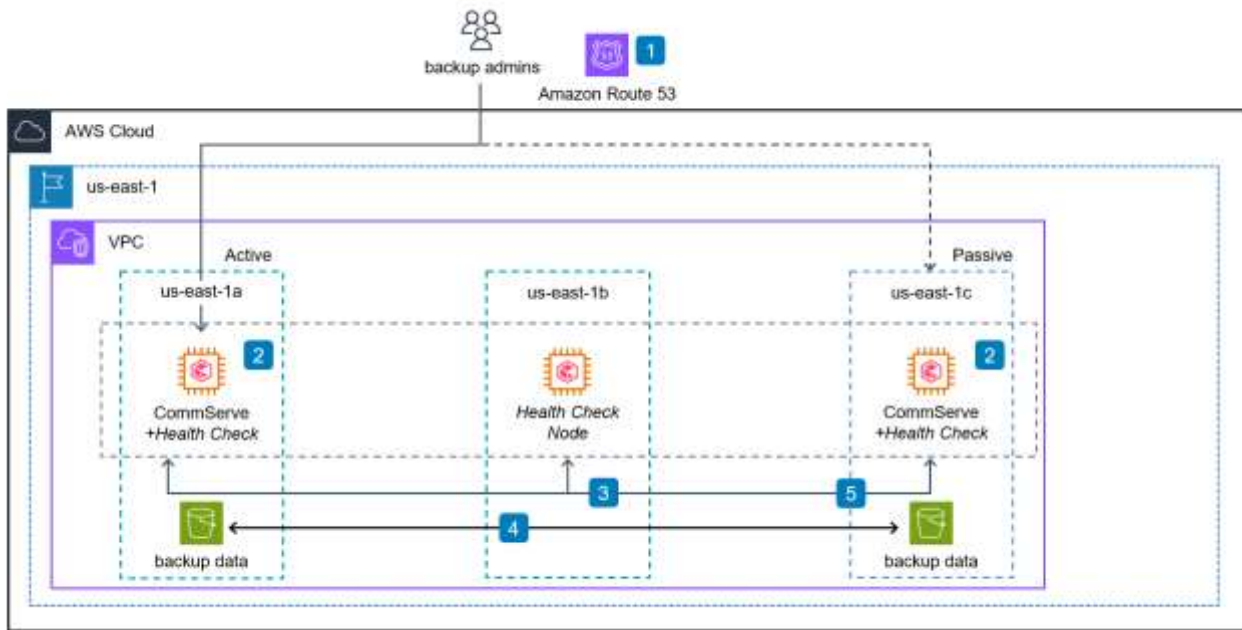
In this multi-region pattern:

1. **Amazon Route 53 DNS health checks** monitor the health and performance of Commvault and route away from unhealthy resources (region).
2. **Commvault CommServe® LiveSync** maintains a single primary (us-east-1) and a single replicated read replica (us-west-1).
3. Commvault application-level health checks run across three regions (us-east-1, us-east-2, us-west-1) and perform **automated failover** and protection from split-brain scenarios.
4. Backup data resides within Amazon S3 within the primary and secondary "DR" region which is designed to deliver 99.999999999% (11 9's) of **data durability** of objects over a given year.
5. Health checks automate failover to passive CommServe instance in same Region, alternate Availability Zone.

Commvault **replicates hard state** (configuration, job history, certificates, identities) periodically to the standby system

Related documents: REL10-BP01 Deploy the workload to multiple locations.

REL10-BP02 Select the appropriate locations for your multi-location deployment

CommServes should be deployed in **active:passive architectures** across AZs or Regions, depending on the scope of the failure scenario being mitigated. Use **Route53 health checks on private sources** to automatically failover CommServe DNS after an automated or manual CommServe failover.

MediaAgents should be deployed in multi-node active-active grids that distribute nodes across AZs, and provide multiple **paths** to shared **Amazon S3 storage**. Do <u>not</u> distribute MediaAgent nodes across regions as this could lead to unintended network egress costs during normal operations. Do <u>not</u> write directly to a remote region Amazon S3 storage service, as this creates cross-regional dependencies.

Access Nodes should be deployed in multi-node active-active groups that distribute nodes across AZs, Access Nodes may be co-located with MediaAgents.

Commvault does not recommend distributing MediaAgent or Access Nodes groups across regions, instead, operate each region as a self-contained fault domain complete with localized MediaAgents grids and Amazon S3 **storage copies**.

Commvault recommends deploying across multiple AZs within a single region unless multi-region disaster recovery is required. Multi-AZ deployment has the benefit of low single-digit roundtrip time latency between AZs allowing Commvault MediaAgents, Access Nodes groups, and Index Server pools to be distributed across physically isolated zones for resilience.

Commvault MediaAgent grids, Access Node groups, and Index Server pools may be deployed in redundant configurations within a single AZ or across multiple AZs, depending on the *blast radius* being considered.

Commvault recommends that business-critical data be replicated to another independent location (Region) to protect from regional Amazon S3 events.

Related documents: REL10-BP02 Select the appropriate locations for your multi-location deployment.

REL10-BP03 Automate recovery for components constrained to a single location

Commvault implements the ability for protected workloads to be recovered to their original data processing location (AZ, AWS Region). Commvault components may be rebuilt by redeploying the affected component from **AWS Marketplace AMIs**, then restoring the system state (**CommServe**, **MediaAgent**).

Commvault does not recommend automated recovery of Commvault components without significant testing, as Commvault may hold the only copy of your data required for organization-wide recovery.

Commvault has automatic failover or job restart when a failure occurs in multi-node single-AZ MediaAgent grid, Access Node group, or Index Server pools. Ideally, recovery should utilize Commvault Backup & Recovery *Full Instance* recovery from the last known good backup.

Accelerated infrastructure rebuild can utilize deployment of new instances from the AWS Marketplace, then subsequent hard state recovery (configuration) on a per-resource basis:

- **MediaAgent Recovery**
- **Access Node Recovery**
- **Index Server Recovery**

Automation of recovery for data management infrastructure requires careful testing to ensure that backup data, configuration state, and data integrity are validated as part of automated rebuild processes.

Related documents: REL10-BP03 Automate recovery for components constrained to a single location.

REL10-BP04 Use bulkhead architectures to limit scope of impact

Commvault writes data to **Cloud Network Storage Pools** using a **bulkhead architecture**. Each pool is deployed with up to four nodes or cells that are responsible for managing the deduplication of data and writing of unique data to Amazon S3.

Commvault uses a hash or partition key to determine which cell should manage the writing of the data to Amazon S3. In the event of a cell outage, data handling responsibility is redirected to the remaining cells, potentially writing duplicate data to the storage pool.

Commvault can also distribute workload across multiple CommServes, MediaAgent, or Access Node 'cells' for resilience, but balancing must occur configured manually.

Commvault software is a distributed data management platform that relies on *bulkheads* or *cell-based architecture* to scale data management horizontally to protect tens of thousands of workloads from a single platform.

Bulkhead pattern with Commvault Backup and Recovery

This reference architecture describes how Commvault Backup and Recovery software can be implemented with independent bulkheads (cells) that allow continued data management service, even when portions of the environment fail.

Commvault scales in independent **regional partitions** that consist of one or more independent **cells** responsible for performing data management *services.* Segmenting your primary Commvault region (CommServe) to utilize independent data movement cells can provide increased resilience at added infrastructure cost.

Commvault recommends consolidating Data Movers (MediaAgents), and Access Nodes within a **regional partition** to reduce Amazon EC2 infrastructure runtime cost. Where a specific service requires additional resilience you may deploy dedicated Cloud Access Nodes (MediaAgents, Access Nodes).

Commvault distributes data management workloads to Access Nodes and MediaAgents using custom scheduling and *data sharding*. Both MediaAgents and Access Nodes may be deployed in redundant groups, in the event of a partition (EC2 instance, DDB volume) becoming unavailable, its workload is redirected to the remaining healthy resources.

> ⚠ **Warning**
>
> When a Deduplication DataBase (DDB) partition goes offline, backup activities will continue but any deduplicated blocks owned by the unhealthy DDB (*shard)* will be hashed and written to an alternate DDB partition. This will result in reduced data duplication savings until the DDB partition is repaired and brought back online.

Related documents: REL10-BP04 Use bulkhead architectures to limit scope of impact.

Design your Workload to Withstand Component Failures

To deliver resilience to your Commvault data management services, consider the following best practices to avoid component failure impact:

- o **Monitor all components of the workload to detect failures**
  Commvault self-monitors connectivity to all registered clients and reports connectivity failures.

- **Failover to healthy resources**
  Commvault software can failover the CommServe, MediaAgents, and Access Nodes to healthy resources (if pre-configured by the failure event).

- **Automate healing on all layers**
  Commvault software contains indexes, deduplication databases, and configuration databases that contain crucial software state to recover the Commvault instance. Commvault does not promote an automated replace-and-restart approach to repairing unhealthy Commvault infrastructure due to this critical state information.

  Commvault recommends using Amazon CloudWatch alarms to automatically **reboot instances** or **recover instances** that are non-responsive. Environments that require minimal interruption, **static stability**, and over-provisioned MediaAgent and Access Node resources provide an alternative to remediate-on-failure runbooks.

- **Rely on the data plane and not the control plane during recovery**
  Commvault requires control plane availability to effectively failover to healthy services or remediate configuration to exclude unhealthy services (i.e., update MediaAgents used for a Cloud Library). Active data plane activities (backup, restore, replication) will fail if the control plane (CommServe®️ instance) becomes unavailable.
  Commvault recommends that **CommServe Livesync** and **CommServe DR backups** are kept up-to-date to allow for automated or runbook-driven recovery of the control plane to another availability zone or Region.

- **Use static stability to prevent bimodal behavior**
  Static stability recommends an approach where the performance and service levels achieved during a failure mode are identical to normal operation. This is often achieved by over-provisioning resources to ensure that failure modes continue to operate with no discernable difference.

  Commvault recommends reviewing your business requirements for backup and recovery to inform the cost and resource demands for your Commvault data management platform. Your *data classification* of business workloads will assist in identifying which workloads justify static stability patterns and investment.

  Commvault cannot use EC2 auto-scaling to provision-on-failure, however, fault detection and Amazon EventBridge could be used to scale up minimal pre-provisioned resource pool to match production instance sizes during failure.

- **Send notifications when events impact availability**
  Commvault recommends monitoring Commvault Amazon EC2 instances using Amazon CloudWatch **System status checks** and **Instance status checks**. Status checks will identify and should be configured to alert the administrator via Amazon SNS. Events detected include loss of network, loss of power, exhausted memory, corrupted file system, or incompatible kernel.

Commvault automatically monitors critical infrastructure and sends alerts when **MediaAgents, Storage, or web-based endpoints** go offline. Additionally, Commvault software implements an interactive administrator-initiated application health check with the **check readiness** probe.

REL11-BP01 Monitor all components of the workload to detect failures

Commvault recommends monitoring all Commvault instances using default **Amazon CloudWatch Instance metrics** (including CPUUtilization, NetworkIn, NetworkOut, and **exceeded network bandwidth**).

Forward Commvault application logs to Amazon CloudWatch Logs to proactively identify errors and failures as they occur and send alarms to operations teams. The default collection interval is typically acceptable.

Create alarms initially on **static thresholds** for common CloudWatch Metrics, then enhance with **anomaly detection** as past metric data grows.

Commvault post-execution scripts can be used to extract and publish **Custom Metrics** to Amazon CloudWatch for Key Performance Indicators (KPIs) like backup or restore throughput or calculated actual recovery time (RTA).

Visualize business and technical metrics in **Commvault dashboards** and/or **CloudWatch dashboards**.

Commvault Backup & Recovery BYOL (**all-in-one**) instances deployed via the AWS Marketplace automatically create and enable automated Amazon CloudWatch availability monitoring, remediation, and email-based alerts to the administrator. Being notified of a component being inaccessible is the first step to a timely resolution.

Commvault recommends implementing availability and performance monitoring of all Commvault Amazon EC2 infrastructure. Key **operating system (OS) performance indicators** (KPIs) that should be monitored include CPU load, memory consumption, Disk, and network I/O utilization. Resource exhaustion and contention at the operating system level should also be monitored by looking at OS metrics like voluntary and involuntary context switching and the total number of page faults.

Often *key performance indicators (KPIs)* are impacted when the infrastructure KPIs are unhealthy. Monitoring job success/failure and **SLA** achieved for data management activities will allow recovery processes to focus on items impacting the business. Commvault will automatically redirect, or restart workloads impacted by a MediaAgent or Access Node failure, manual runbook-based remediation can then occur.

Related documents: REL11-BP01 Monitor all components of the workload to detect failures.

REL11-BP02 Fail over to healthy resources

Commvault components deployed across multiple locations will automatically retry and redirect to healthy resources.

CommServe **automatic failover** will failover between active and passive CommServe instances. Amazon 53 health checks should be used to failover customer-facing DNS as part of automatic failover.

Cloud Network Storage Pools will redirect deduplication and data access requests to remaining MediaAgents.

Access Groups will re-schedule interrupted jobs to remaining healthy worker nodes.

Restores for backups written and copied to multiple locations will automatically request remote copies using **copy precedence** if the primary copy is unavailable.

Commvault **regional partitions** consist of **independent cells** that provide targeted data management services. The following failover capabilities should be used with multi-AZ and multi-regional deployment to automatically failover actions to healthy resources:

- Regional failover
    - CommServe® Amazon EC2 instances may be failed over between AZs or Regions using **Commvault CommServe LiveSync for High Availability Disaster Recovery.**

- When failover across regions occurs, it is recommended that a **regional data copy** exist, along with regional MediaAgent cells to provide continued recoverability of business-critical workloads. Commvault will not automatically select a secondary copy for recovery, the backup administrator must select the copy during recovery.
- Zonal failover
    - MediaAgents operate in multi-node grids that contain one to four Amazon EC2 instances operating as a single unit across multiple AZs. Commvault software will automatically redirect to healthy MediaAgents in the event of a failure.
    - Access Nodes operate in **multi-node groups** that contain multiple Amazon EC2 instances operating as a single unit across multiple AZs. Commvault software will restart impacted data management activities on healthy Access Nodes in the event of a failure.

Access Nodes have a single coordinator that coordinates and schedules activities with the worker nodes. If the single master (coordinator) fails, active data management activities will be marked failed and the CommServe will **restart the job**, selecting a new healthy coordinator.

Related documents: REL11-BP02 Fail over to healthy resources.

REL11-BP03 Automate healing on all layers

Commvault recommends using **Amazon EC2 Automatic Recovery** for Commvault CommServe, MediaAgent, and Access Node instances.

MediaAgent and Access Node reboots will result in a temporary interruption to service, with restarts or retries resuming services once the instance becomes available.

All instance metadata (including IP addresses) and persistent storage located on attached Amazon EBS volumes are unaffected during an automated recovery.

Backup activities leveraging auto-scaling access nodes will automatically restart failed backup operations and dispatch backups to alternate (healthy) Amazon EC2 instances if a failure occurs.

Related documents: REL11-BP03 Automate healing on all layers.

REL11-BP04 Rely on the data plane and not the control plane during recovery

Commvault recovery of protected workloads relies on both the control plane (provisioning of new resources) and data plane (restoring data into newly created resources) in full instance recovery.

Granular recovery modes do not require the recreation of existing resources or dependence on the control plane during recovery (**EC2 Agentless File Recovery**, **RDS dump-based recovery**, **EFS** and **FSx for Windows** file recovery, **FSxN** file recovery, **S3** object recovery).

Related documents: REL11-BP04 Rely on the data plane and not the control plane during recovery.

REL11-BP05 Use static stability to prevent bimodal behavior

Commvault recommends using a **static stability approach** to MediaAgent and Access Node sizing when performance in normal and failure modes must be identical.

Static stability results in increased compute runtime cost and wastage and will result in AWS Compute Optimizer recommendations consistently reporting resources as **over-provisioned**.

Commvault recommends cost-optimizing static stability deployments by restricting them to mission-critical dedicated resources only.

Related documents: REL11-BP05 Use static stability to prevent bimodal behavior.

Commvault recommends that **CloudWatch Alarms based on Static Thresholds** are enabled to inform operations teams of conditions that may invoke automated healing or directing workload away from an existing resource (i.e., CPU load that prevents scheduling to a specific worker, automated recovery or reboot of EC2 instances).

Alarms are intended to result in operations team action to discover the root cause for the condition and potentially put automated resolutions (EventBridge event rules) for future events.

Commvault recommends regular data management **game days** that evaluate procedures and incident response capabilities for responding to events in pre-production and production environments. Game days let you evaluate and hone documented procedures in runbooks and playbooks, CloudWatch dashboards, alarms, and automation and team competency. Game days also provide an ability to perform large-scale performance or load testing to evaluate your recovery readiness to respond to a large-scale event.

> 💡 **Pro-Tip**
>
> Consider running mass-recovery testing for ransomware or malware events by **adding multiple read-only MediaAgents** to an existing Cloud Storage Library.

Related documents: REL11-BP06 Send notifications when events impact availability.

Commvault software may be architected to meet varying availability targets and uptime service level agreements (SLAs). Commvault recommends reviewing the SLAs for the applications and business functions being protected by your Commvault Backup & Recovery system and aiming to exceed those availability and uptime targets.

Commvault is a centralized control-plane for data management across your multi-account landing zone, as such each component in the Commvault Backup & Recovery system must be protected from failure. Commvault is a simple three-tier architecture with the following components and availability mitigations:

- **Commvault CommServe®** is the centralized control plane for all backup, recovery, and replication activities. The CommServe is protected by a **Disaster Recovery (DR) backup** to a remote region or highly durable Amazon S3. Optionally, **CommServe LiveSync** may be used to maintain a one-to-many asynchronous replica of the CommServe state across Availability Zones or Regions. Only a single CommServe® may be active at any point in time, replica CommServes are read-only.
- **Commvault MediaAgents** are data movers the receive backup data and write it to highly durable Amazon S3, read and replicate backup data across Regions, or read and restore data. MediaAgents are regional and are responsible for Amazon S3 backup or archival data within a single Region. MediaAgents may be deployed in **fault-tolerant grids** across Availability Zones (AZs) to auto-resume backup, replication, and recovery activity in the event of one or more AZs and/or MAs failing. MediaAgents require specific **recovery processes** as they are responsible for maintaining the location of backup data and any copies.
- **Commvault Access Nodes** are Regional and/or Zonal workload-based data movers that interact directly with a protected AWS service and workload to both take backups and restore backups. Access Nodes may be deployed in fault-tolerant groups across AZs to auto-resume backup or recovery activity in the event of one or more AZs or Access Nodes failing.

CommServe resources may be scaled vertically. MediaAgents may be scaled horizontally (up to **four nodes** in a grid), and Access Nodes may be scaled horizontally with no node limit per group. Commvault recommends using **Automatic scaling for Access Nodes** to allow Commvault to assess and auto-scale the required number of access nodes at backup time. Commvault terminates the Access Nodes once the data management activity has been completed.

Commvault is heavily dependent on AWS IAM, AWS KMS, AWS STS, AWS PrivateLink, AWS Direct Connect, AWS Client VPN, AWS Site-to-Site VPN, and Domain Name System (DNS) to perform forward and reverse name resolution and communication between Commvault components.

Performing software upgrades on Commvault components can take as little as minutes or as long as an hour depending on the scope of the upgrade. Environments requiring very high uptime (i.e., 99.99, 99.999) can perform staggered upgrades using the **passive CommServe instance(s**).99.99%, 99.999% availability) can perform staggered upgrades using the **passive CommServe instance(s)**.

See **Reliability - REL11-BP07 Architect your product to meet availability targets and uptime service level agreements (SLAs)**.

Related documents: **REL11-BP07 Architect your product to meet availability targets and uptime service level agreements (SLAs)**.

Test Reliability

The only accepted method to validate the reliability of your Commvault data platform is to test frequently. Testing needs to confirm that *regional partitions* and *independent cells* provide fault isolation as expected and that automated failover delivers business-expected performance and service levels. Consider the following best practices when designing your test plans.

REL12-BP01 Use playbooks to investigate failures

Commvault recommends developing **playbooks** that provide prescriptive steps to investigate and resolve specific failure scenarios.

Playbooks should include the minimum IAM role or permissions required to perform investigation, commands to run, and expected (normal, abnormal) output if known.

Commvault recommends implementing discovery portions of playbooks as code. Commvault notifications allow the execution of **command-line** scripts and **workflows**. Amazon CloudWatch alarms also run automated playbooks as code, for example, **Use Amazon EventBridge rules to run AWS Systems Manager automation in response to CloudWatch alarms**.

Document the process to debug and investigate issues in your Commvault data management environment in *playbooks*. Playbooks are checklists that support, and engineering resources use to understand and pinpoint failures. As you gain more experience with Commvault you will define a set of playbooks, some targeted at specific workloads or tasks and others providing the generic detective skills to investigate and solve any non-routine failure.

> 🔆 **Pro-Tip**
>
> Store your playbooks in a highly-available, secure, and accessible location away from your Commvault backup and recovery system.

Related documents: **REL12-BP01 Use playbooks to investigate failures**.

REL12-BP02 Perform post-incident analysis

Commvault recommends creating a standard post-incident review (PIR) process and/or checklist.

Reviews should consider whether existing metrics and alarms captured the incident and remediate if key information was not available during the investigation.

Reviews should also consider updating existing playbooks and runbooks to accommodate the newly discovered failure modes and resolutions.

Every customer-facing event is an opportunity to improve your data management services. Review incidents to identify root causes, remove or reduce the likelihood of reoccurrence, and define playbooks to improve future response time. Communicate and record learnings with team playbooks and runbooks and inform affected parties of learnings. It may be useful to understand the root cause of events that drive *recovery events* to affect and inform improved application architecture.

Consider taking learnings and publishing recommended best practices, standards, and design guardrails to improve your organizational capability to build reliable workloads.

Related documents: REL12-BP02 Perform post-incident analysis.

REL12-BP03 Test functional requirements

Commvault performs functional unit testing as part of product development, testing, and release management.

When testing functional capabilities of a new feature or software update, Commvault recommends testing functionality using canary deployments that direct workloads under test to dedicated test-only infrastructure (i.e., upgrade an access node or MediaAgent to test infrastructure for a single workload).

Functional testing performs unit and integration testing to ensure that required features and functions are operating as expected. As your workloads, data types, and Commvault software change, you must continually re-validate that critical functional capabilities still function.

Due to the diverse nature of modern application workloads and data types, testing should be automated and use either synthetic testing (i.e. "**canary testing**") or temporarily spin-up clones of production workloads for tests.

Related documents: REL12-BP03 Test functional requirements.

REL12-BP04 Test scaling and performance requirements

Commvault recommends performing large-scale load testing with a representative set of workload types and data volume. You can use your production workload backups to restore and replicate a production-like isolated Amazon VPC to provide scale performance testing. Once testing is complete, you can destroy the scale testing VPC without any impact on production workloads.

Operating a reliable Commvault data management platform means delivering consistent backup and recovery performance as workload data grows. Scale testing involves temporarily scaling up protected workloads and Commvault protection resources to ensure that performance demands can be met and do not exhaust service quotas.

Scale testing should be performed within the production environment, if testing occurs in an isolated pre-production environment any learnings must be applied to production and re-validated. Scale testing must consider all elements responsible for handling workload data, including MediaAgents, Access Nodes, Index Servers, and the CommServe instance.

Related documents: REL12-BP04 Test scaling and performance requirements.

REL12-BP05 Test resiliency using chaos engineering

Commvault recommends that lab-based and production game day testing involve the use of **chaos engineering** principles or **AWS Fault Injection Simulator (AWS FIS)** to inject failures into CommServe, MediaAgents, and Access Nodes.

Consider using a mixture of failure modes that affect entire instances, or individual sub-components like the corruption of a DDB (see **Injecting Chaos to Amazon EC2 using AWS Systems Manager**).

> ⚠ **Warning**
>
> DDB corruption as part of testing may result in redundant data being rewritten by future backup activity.
> Consider DDB fault injection occurring in pre-production environments.

Chaos engineering is the discipline of injecting failures regularly into pre-production and production environments. Ensure that validated manual runbooks and playbooks for recovery are available before performing chaos engineering. Chaos engineering can be automated to routinely simulate failures and validate reliable operation. Enable automated component failure only for scenarios where you have designed resiliency (i.e., MediaAgent failover, Access Node failover, CommServe failover).

Amazon provides **Amazon Fault Injection Simulator (FIS)**, and **The Chaos Toolkit** to jumpstart chaos engineering. Testing should include the impact of the loss of key components in your Commvault data management platform including by not limited to;- Amazon EC2 instances, Amazon EBS volumes, and Amazon S3 buckets.

Related documents: REL12-BP05 Test resiliency using chaos engineering.

REL12-BP06 Conduct game days regularly.

Promote regular **game days** to exercise your team and developed processes (runbooks, playbooks) for timely recovery of Commvault data management services. Game days provide the opportunity to validate the operational process and team access, and remedy issues before a real failure event.

Related documents: REL12-BP06 Conduct game days regularly.

Plan for Disaster Recovery (DR)

Commvault Backup & Recovery is a crucial component in your Disaster Recovery (DR) strategy for your business workloads. Likewise, your Commvault data management system needs to be built to recover from zonal and regional failures.

Considering your *application workloads*, use data classification to identify the criticality and sensitivity of your workloads and whether they require **Commvault Auto Recovery (formerly Disaster Recovery)** protection. Commvault provides DR replication and recovery for **Virtual Machines** (Amazon EC2 instances), **Databases** (Amazon RDS), **File systems** (Amazon EBS, Amazon EFS, Amazon FSx for Windows), and **Object Storage** (Amazon S3).

Commvault provides multiple Recovery Point Objectives (RPO) and Recovery Time Objectives (RTO) for workloads protected using **Commvault Auto Recovery**.

Disaster Recovery pattern	RPO/RTO	RPO/RTO
Commvault Backup & Recovery	Hours	Less critical applications Data staged in Amazon S3 Create Amazon EC2 instances after a disaster
Commvault Auto Recovery - Pilot light (warm site)	Tens of minutes	Tier 3 business-critical applications Data live in Amazon EBS snapshots Create Amazon EC2 instances after a disaster

Commvault Auto Recovery - Warm standby (hot site)	Minutes	Tier 2 business-critical application Compute instance created and shutdown Power-on Amazon EC2 instance after a disaster
Continuous data replication	Sub-minute	Tier 1 mission-critical apps An application running in active/active mode Commvault replicating database and file-system changes between instances

Remember that your Commvault Backup and Recovery system needs to be resilient to disaster-level events to provide recovery services during a large-scale event. Commvault CommServe® instances may be separated by wide-area network distances (i.e., regional separation), and data handling infrastructure (MediaAgents, Access Nodes) must be deployed within a Region (i.e., zonal separation).

Discrete data copies must exist for all backup copies that will be required during a large-scale disaster. Typically, fewer recovery points are required in a disaster, only the *latest* copy of a workload is needed to resume service in the remote location. This allows differing **data aging** rules for data intended for disaster recovery.

Consider the following best practices when planning for disaster recovery:

REL13-BP01 Define recovery objectives for downtime and data loss

Commvault recommends working with business stakeholders for a protected workload to establish the **Disaster Recovery Point Objective** (acceptable data loss during a DR event) and **Disaster Recovery Time Objective** (acceptable recovery time after a disaster).

Ensure that specialized DRPO and DRTO classifications are assigned to workloads as **AWS Resource Tags**. You will maintain those recovery objectives (or recovery classifications) as Commvault discovers, protects, replications, and recovers your workloads with tags intact.

Document the recovery approach for these applications in your Disaster Recovery Plan, and test objectives are achievable.

Consider the *recovery time objective (RTO)* or acceptable delay between interruption and resumption of service, and the *recovery point objective (RPO),* or acceptable data loss between recovery point and interruption event for both protected workloads and your Commvault data management platform.

> 💡 **Pro-Tip**
>
> You may have two sets of objectives your *Operational Recovery Objectives* (ORTO, ORPO) which refer to recovery in the original region, and your Disaster Recovery Objectives (DRTO, DRPO). Be sure to consider your business requirements in both scenarios.

Your Commvault data management platform will require a Recovery Time Objective (RTO) equivalent to or better than your most critical workload RTO.

Related documents: REL13-BP01 Define recovery objectives for downtime and data loss.

Review the required Recovery Point Objective (RTO) and Recovery Time Objective (RPO) for your mission-critical workloads during a disaster event and select a **DR recovery strategy**. You will have multiple tiers of criticality in your data classification policy, which dictate a differing recovery approach and the overall cost and complexity.

Commvault supports multiple DR recovery strategies including Backup and Recovery (RPO in hours), **Warm site replication** (RPO in 10s of mins), **Hot site** replication (RPO in mins), and Active-Active (RPO in mins, **database**, **file systems**, **object**).

Commvault provides multiple workload recovery strategies (detailed above) that deliver recovery time objectives from 24 hours down to under a minute. Your protected workloads will have differing *data classifications* which will dictate the recovery strategy required. Commvault can utilize AWS Resource Tags on your infrastructure to create or inform **Replication groups** for differentiated recovery time by application importance.

Your Commvault data management platform will require a recovery strategy that can be executed and resume data management services in time to recover your most stringent workload. For example, if your most critical application has a recovery time of 6 hours, but a recovery point objective of 2 hours you need to ensure that your Commvault system can failover and resume workload replication in less than 2 hours.

Commvault disaster recovery can be implemented as:

- **Backup and restore** (RPO in hours, RTO in hours) by writing Commvault DR backups to the remote region and provisioning, installing, then **recovering** your Commvault CommServe® database DR backups.
- **Pilot light** (RPO in minutes, RTO in hours) by pre-provisioning a Commvault CommServe® with software pre-installed and using **DR backups** to configure after a disaster.
- **Warm standby** (RPO in minutes, RTO in minutes to hours) by using **Commvault CommServe LiveSync for High Availability** to replicate the CommServe Database (CSDB) regularly to the remote region to scaled-down or production-grade a replica. This option supports **automated failover** to reduce failover time. Scaled-down infrastructure must be capable of processing data replication from the production site promptly.

Commvault does not support **Multi-region active-active** disaster strategies for the Commvault data management platform.

Related documents: REL13-BP02 Use defined recovery strategies to meet the recovery objectives.

REL13-BP03 Test disaster recovery implementation to validate the implementation

Regularly test both operational recovery (day-to-day to original AZ or AWS Region) and disaster recovery (to alternate AZ or AWS Region).

Be sure to record the overall end-to-end recovery time and review it against business requirements. Consider involving the application and internal stakeholders in testing to ensure actual disaster recovery events are executed with minimal cross-team delay.

AWS provides on-demand, elastic compute, network, and storage resources to allow regular testing of your disaster recovery scenarios. Regularly test *protected workload DR* and *Commvault platform DR* to ensure that business RTOs and RPOs can be met. Failover testing should be executed in production to ensure the testing is valid and benefits the production workload reliability.

Related documents: REL13-BP03 Test disaster recovery implementation to validate the implementation.

Commvault recommends using AWS Config to record and compare regional configuration (i.e., AWS Service Quotas), and **remediate**, between production and disaster recovery regions and accounts.

Configuration drift for protected workloads should be minimal if using a periodic replication approach to DR.

Where a recovery from a backup is used for on-demand disaster recovery, ensure application owners maintain a configuration register to re-apply configuration changes made after the backup used to recover the workload.

Your DR strategy will only be successful if your day-to-day operational practices record all change and ensure it is replicated in DR instances. Ensure that infrastructure, application, and data management configurations are identical for primary and DR regions. This includes maintaining consistent **service quotas** across regions, network bandwidth, and instance-level OS and application configuration. Consider implementing automated configuration drift detection using tools like **AWS Config**.

Related documents: REL13-BP04 Manage configuration drift at the DR site or Region.

Commvault recommends using **Commvault Backup & Recovery** and **Commvault Auto Recovery (formerly Disaster Recovery)** to provide automated recovery of workloads to AWS Cloud.

Commvault has the **broadest industry support** for cloud, containers, SaaS, and traditional workloads and supports recovery to Amazon EC2 and fully-managed services during a recovery event from AWS or in edge-based locations.

Commvault automated recovery orchestrates multiple complex steps including provisioning of new compute, storage, and network, and then restoring system state backup. The use of Commvault reduces human error while allowing self-service recovery without detailed AWS service knowledge.

Commvault facilitates application owner self-recovery by providing 24x7 secure self-service access to the Commvault Command Center. Additionally, **Commvault Auto Recovery (formerly Disaster Recovery)** provides automated disaster recovery for AWS and on-premises workloads including virtual machines, databases, file systems, object storage, and Hadoop clusters. **Cross-vendor disaster recovery** is possible from on-premises virtualization platforms directly into native Amazon EC2 instances. Commvault software automates planned failovers, unplanned failovers, failback, and failover testing.

Commvault Auto Recovery is automated using Commvault application health checks across a minimum of three (3) **monitoring nodes**. Commvault will fail away from unhealthy resources and **Route53 DNS failover** will automatically detect the failover and redirect incoming connections to the new active resource.

Related documents: REL13-BP05 Automate recovery.

Additional Resources

- **Reliability Pillar: AWS Well-Architected**
- **AWS Service Quotas**
- **Monitor your instances using CloudWatch**
- The Amazon Builders' Library: **Implementing health checks**

- **Implement automatic drift remediation for AWS CloudFormation using Amazon CloudWatch and AWS Lambda**
- **AWS re:Invent 2020: Optimizing protection for AWS service workloads at petabyte scale & beyond**
- **AWS Marketplace - Commvault**
- **Commvault Backup & Recovery for AWS**
- **Commvault Auto Recovery (formerly Disaster Recovery)**

Performance Efficiency Pillar

A key success measure of your Commvault data management platform will be its ability to meet your business *recovery point objectives (RPOs)* and *recovery time objectives (RTOs)* as demand changes. Additionally, Commvault infrastructure must be utilized efficiently to avoid uncontrolled costs and the complexity of ongoing data management operations.

Consider the following key areas to design and maintain performance efficiency:

- **Democratize advanced technologies: Make advanced technology implementation easier for your team**
  Consider leveraging AWS-managed services over the development and maintenance of bespoke IT administrative processes (i.e., leveraging AWS WAF to secure Commvault web services). Use Commvault Backup & Recovery to automate snapshot management, cross-region, and cross-account snapshot transfers, and data archival. Avoid the use of scripted data management processes, they can lead to orphaned, uncontrolled, and unsecured resources

  Commvault integrates seamlessly with environments managed by **Amazon Managed Services (AMS)**, using a mixture of snapshot and streaming backup and recovery methods. Commvault software can be used to protect *AWS Managed Services Accelerate* and *AWS Managed Services Advanced* environments.

- **Go global in minutes**
  Gain on-demand 'site-less' Disaster Recovery for your Commvault data management platform and *protected workloads* by leveraging AWS global infrastructure and **AWS Regions**.

- **Use serverless architectures**
  Leverage serverless resources that remove the need to provision and maintain operating systems and applications. Commvault provides on-demand workload-optimized backup using **auto-scaling access nodes**. Additionally, utilize serverless storage products like Amazon S3, Amazon EFS, and Amazon FSx for highly available operational runbook and playbook storage.

- **Experiment more often**
  Integrate continual curiosity into your operational teams by promoting testing of new AWS services, instances, storage classes, and configurations to optimize data management.

- **Consider mechanical sympathy**
  Match the AWS service and technology class and service levels you need to deliver. Commvault technology selection must match the service levels required for the given data and region.

Performance efficiency is an ongoing process that needs to be reviewed regularly and refreshed based on workload changes and data criticality changes.

Selection

AWS resources (compute, storage, database, and network) come in varying types and configurations. An optimal selection of resources requires an understanding of your needs within a given Region and Availability Zone. Consider the following AWS resource selection best practices:

Performance architecture selection

Use a data-driven approach to assessing and selecting the optimal configuration of AWS resources for your needs. Commvault provides the AWS Cloud Architecture Guide (this document) with Reference Architectures that have been benchmarked and proven in Commvault, partner, and customer environments.

Consider the following when assessing your performance needs:

PERF01-BP01 Understand the available services and resources

Stay aware of recent innovations and announcements in AWS by subscribing to the **What's New with AWS?** updates.

Commvault is dependent particularly on computing, storage, and networking services and will announce new supportability in the **Newsletter for New Features in Commvault Platform Release 2022E**.

Commvault also updates the **Cloud Architecture Guide** (this document) for each long-term support release with recommended services to use with Commvault software.

Related documents: PERF01-BP01 Understand the available services and resources.

PERF01-BP02 Define a process for architectural choices

Ensure that architectural review of new technologies includes validation of technology to maintain consistent *recovery point objectives (RPOs)* and *recovery time objectives (RTOs)* for your protected workloads.

Commvault publishes the least cost and best price-performance recommendation for day-one and day-two CommServe, MediaAgents, and Access Nodes instances (see **Sizing Guidelines**). Architectural choice of components should consider the business-required RPO/RTOs, networking streamed performance per instance, and cost per instance.

Commvault recommends using the Cloud Architecture Guide, Reference Architectures, and Design Principles and Best Practices to architect your Commvault data management platform in AWS. Continually reference your architecture through **experimentation**, benchmarking, and new business demand.

Consult your Commvault Sales Engineer (SE) and FAST Solution Architects, and leverage **Commvault Technology Consulting** to help tune your deployment to current established best practices for performance efficiency.

Consider describing your performance requirements as **Amazon CloudWatch Metrics** that can be baselined and then reviewed post-implementation. Agree on these metrics with key stakeholders before moving to implementation.

Related documents: PERF01-BP02 Define a process for architectural choices.

PERF01-BP03 Factor cost requirements into decisions.

New technology selection must always assess the benefit against any increased cost to the business. Use your *workload data classification* to determine if an increased cost is in alignment with the delivery of agreed business RPO/RTOs (per workload data classification).

Commvault recommends optimizing your Commvault data management platform to reduce costs. Review the **Cost Optimization Pillar** for recommendations on leveraging appropriate instance families to optimize for the least cost.

Always deploy with the least cost resource that meets the functional and non-functional requirements of the workload. Regularly re-assess and right-size resources based on recommendations from **AWS Compute Optimizer**.

Remember that the value of the resource investment must match or exceed to value it provides to the business, consider the *business classification* of the workload being protected and the *data classification* of data being protected to help inform architectural decisions.

Refer to the **Well-Architected – Cost Optimization Pillar** for additional guidance.

Related documents: PERF01-BP03 Factor cost requirements into decisions.

PERF01-BP04 Use policies or reference architectures.

Utilize Commvault Reference Architectures to guide the alignment of your Commvault data management platform to existing internal policies and standards. Be aware that Commvault's data management architecture and design may drive the development of new policies.

Refer to the **Reference Architectures** section for additional details.

Related documents: PERF01-BP04 Use policies or reference architectures.

PERF01-BP05 Use guidance from your cloud provider or an appropriate partner.

Use the latest available **Commvault AWS Cloud Architecture Guide** (this document) to design for optimal performance of your Commvault data management platform. Utilize **Commvault Technology Consulting**, **AWS Professional Services,** and **Commvault** or **AWS partners** to assist in architectural guidance and implementation.

Refer to the **Design Principles and Best Practices** section for current guidance.

Commvault Technology Consult may be purchased via the **AWS Marketplace**.

Related documents: PERF01-BP05 Use guidance from your cloud provider or an appropriate partner.

PERF01-BP06 Benchmark existing workloads

Commvault recommends performing performance benchmarking during platform development and testing before the production release. Benchmarks must be performed to establish a known good performance baseline for the platform.

Understand how your workload performs today by baselining performance from a business perspective (i.e., RPO, RTO) and infrastructure perspective (CPU load, Free RAM, Disk IOPS, Network throughput). Ensure that any testing exceeds any burstable guarantees for a service (for example, **Understanding Burst vs. Baseline Performance with Amazon RDS and GP2**).

Commvault recommends performing real-user monitoring or testing with real production data to accurately simulate the performance that will be experienced in daily operations.

Commvault can perform data masking on **Oracle**, **SQL Server**, and **Salesforce** if masking is required to use production-like datasets.

Related documents: PERF01-BP06 Benchmark existing workloads.

PERF01-BP07 Load test your workload.

Commvault recommends testing proposed performance enhancements with load-testing or proofs-of-concept that perform production-scale testing. Use AWS CloudFormation to provide test environments on-demand from the **AWS Marketplace**. Use Commvault to perform out-of-place recovery of a representative set of production workloads and test and measure how a workload architecture improves performance under load.

Deploy Commvault Backup & Recovery cells with differing resource types and sizes to gauge improvement from baselines established during benchmark testing. Testing must include the performance impact (i.e., how fast

something completes) but also how efficiently It is performed (i.e., how much under-utilized resource existed during the test).

Consider testing with low-cost resources like **Amazon EC2 Spot Instances**, be aware though Spot Instances <u>cannot</u> be used by Commvault in a production environment. due to the impact on data protection operations when a resource is reclaimed. Consider testing in Regions where test instance prices are lower.

Terminate resources when testing is complete.

Related documents: PERF01-BP07 Load test your workload.

Compute architecture selection

Commvault provides **sizing recommendations** that identify optimal Amazon EC2 instance types and sizes for Commvault CommServe® instance, MediaAgents, and Access Node infrastructure. There are multiple computing technologies available, consider the following when assessing options:

PERF02-BP01 Evaluate the available compute options

Commvault provides prescriptive guidance on the compute options that deliver optimal performance while reducing cost and lowering operational complexity. Evaluate the **benchmark performance results** to select your instance.

Commvault details the recommended day-one **seed deployments** and day-two+ **scale-out deployments** for CommServe and MediaAgents.

Instances

Commvault provides prescriptive guidance on the selection of supported **Amazon EC2 instance types and sizes**. Commvault has analyzed the demands of Commvault software in lab testing and production deployments and selected both <u>least-cost</u> and <u>best-price performance</u> instance recommendations. **Commvault Marketplace AMI-based products** offer only the Commvault currently recommended instance types and sizes.

Commvault recommends current-generation, **nitro-based**, **EBS-optimized** instances. Commvault recommends the use of General Purpose 3 (gp3) SSD volumes exclusively to allow independent tuning of capacity and performance.

> ⓘ Note
>
> Commvault does not require or recommend hardware-accelerated instances with **GPU acceleration**, **FPGAs**, or **AWS Inferentia** chipsets. Commvault does not require or recommend **enhanced networking**, **enhanced storage**, or **local instance storage**-enabled instances.

Commvault Seed CommServe® instances

The Commvault control plane (CommServe) and data plane (MediaAgents, Access Nodes) can be combined into an all-in-one deployment for initial day-one seed deployments. Commvault software requires the use of Memory optimized R5a (least cost) or R6i/R6a (best price/performance). As protected data volume grows, vertically scale for additional resources.

Commvault recommends starting planning with the following configurations, instances are listed in priority order.

Commvault all-in-one CommServe® instance	Instance Details
Least cost	r5a.xlarge

Best price performance	r6i.xlarge r6a.xlarge
Access Node: Indicative max. protected workloads (snapshot)	Two thousand protected workloads
MediaAgent: Indicative max. protected client data (streamed TB)	60
Snapshot-only vs. Snapshot and Streaming backup ratio	80:20
Vertical scaling supported	Yes
Horizontal scaling supported	Yes

Derived from **On-premises Hardware Specifications for All-in-One CommServe Server**.

Assumes 30TB (VM/Files), and 30TB (Database) with 30 daily, 4 weekly, 6 months, and 1 yearly backup

Access Node and MediaAgent sizing are guidelines for planning and differ per data type and retention

See *Cost-effective resources* for a comparison of all-in-one CommServe® instance costs on differing Amazon EC2 instance types.

CommServe® instance Selection Guidance

Commvault recommends planning for one virtual CPU (1 vCPU) for every five hundred Amazon workloads protected using native snapshot protection.

Commvault recommends using **memory-optimized** Instance types with a memory-to-CPU ratio of no less than 8:1 (e.g., r6a.large instance size with 16GiB RAM, 2 vCPU = 8:1),

Commvault recommends vertical scaling when CPU or RAM is exhausted on the all-in-one CommServe, but sufficient network bandwidth is still available.

Commvault recommends horizontal scaling when the network bandwidth is exhausted on the all-in-one CommServe. MediaAgent grids must consist of identical nodes, so backup, restore, and replication activity should be moved onto a separate scale-out MediaAgent grid. Migration of data management workload to a MediaAgent grid may result in an opportunity to right-size the all-in-one CommServe.

Commvault MediaAgent Grids

Commvault data management and data restore capabilities may be deployed on resilient MediaAgent grids that provide load-balancing and failover capability if a node becomes unavailable.

Commvault recommends starting planning with the following configurations, instances are listed in priority order.

Commvault MediaAgent + Access Node	Instance Details
Least cost (and best performance)	c7g.xlarge [arm64]
Best performance	c6i.large [x86_64] c6a.large [x86_64]
Access Node: Indicative max. protected workloads (snapshot)	2000 per node

MediaAgent: Indicative max. protected client data (streamed TB)	60 per node
Max. Number of Nodes in MediaAgent Grid	1,2,3.4
Vertical scaling supported	Yes
Horizontal scaling supported	Yes

Derived from **On-premises Hardware Specifications for MediaAgents (deduplication mode)**

Includes Commvault MediaAgent, Virtual Server Agent, CloudApps, and IntelliSnap roles.

Assumes 30TB (VM/Files), and 30TB (Database) with 30 daily, 4 weekly, 6 months, and 1 yearly backup

Access Node and MediaAgent sizing are guidelines for planning and differ per data type and retention

MediaAgent Selection Guidance

See **CommServe® instance Selection Guidance** for details on sizing the Access Node component.

Commvault recommends planning for one virtual CPU (1 vCPU) and 6GiB RAM for every 100TB of stored data after deduplication and compression.

Commvault recommends planning with **memory-optimized** Instance types with a Memory to CPU ratio of no less than 2:1 (e.g., c7g.xlarge with 8GiB RAM, 4 vCPU = 2:1),

Commvault Auto-scaling Access Nodes

Commvault automatically provisions access nodes to handle the streaming of backup data from AWS services to Commvault-controlled cloud storage. Auto-scaled access nodes adjust the number of provisioned resources dependent on data volume and SLA for workloads.

Automatic Access Nodes are provisioned into the destination region and availability zone of the workloads to be protected. The matching availability zone must be configured in the auto-scaling **VM provisioning settings**. During a backup activity, Commvault will select any available Access Node within the same region as the workload to be protected. Set **bEnableHostDispatch**=1 to prefer the selection of an Access Node from the same availability zone as the workload., and fallback to access nodes within the same region.

> ⓘ Note
>
> Zonal Access Node selection only occurs for backup activities. Restore activities will pick any available access node in the region.

Commvault Access Node (Snapshot + Streaming)	Instance Details
Least cost (and best performance)	c6g.xlarge [arm64]
Best performance	c7g.xlarge [arm64]
	c6i.xlarge [x86_64]
	c6a.xlarge [x86_64]
Indicative max. protected workloads (snapshot)	2000
Indicative max. protected workloads (streamed client-side TB)	60

Max. Number of Nodes in MediaAgent Grid	1, 2, 3, 4
Vertical scaling supported	Yes
Horizontal scaling supported	Yes

Derived from **IntelliSnap and Streaming System Requirements**.

Includes Commvault MediaAgent, Virtual Server Agent, CloudApps, and IntelliSnap roles.

Assumes 30TB (VM/Files), and 30TB (Database) with 30 daily, 4 weekly, 6 months, and 1 yearly backup
Access Node and MediaAgent sizing are guidelines for planning and differ per data type and retention

Commvault Snapshot Only Access Node / Cloud Controller

Commvault performs snapshot-based protection of your AWS workloads by orchestrating snapshot creation, sharing, and copying across regions and accounts. The compute needs for a snapshot-only instance are significantly reduced as no data other than index information needs to be transferred to Commvault storage. These nodes are used in environments where the Commvault infrastructure is on-premises and expansion to AWS introduces a need to orchestrate snapshots (from on-premises).

Features	Commvault Access Node (Snapshot-only)
Least cost	t4g.small [arm64]
Best price performance	t3a.small [arm64] t3.small [x86_64]
Indicative max. protected workloads (snapshot)	2000
Indicative max. protected workloads (streamed client-side TB)	Not supported
Max. Number of Nodes in MediaAgent Grid	1, 2, 3, 4
Vertical scaling supported	Yes
Horizontal scaling supported	Yes

Use **AWS Compute Optimizer** to identify over-provisioned or under-provisioned resources (CPU, RAM, network, EBS bandwidth) and right-size accordingly. Commvault recommends leveraging enhanced infrastructure metrics (see **Pricing**) to use metrics for up to three months vs. the default of 14 days. Temporary increases in load driven by cyclic business activities or one-time data migration can require a **temporary increase** to the instance type and size, then reduce size when complete.

> ⓘ Note
>
> Commvault does not support the use of burstable compute instance types (T4g, T3, T3a, T2) for **streaming backups**, use only for **IntelliSnap® snapshot backup** and **MediaAgent Cloud Controllers**. This is due to the sustained CPU load that streaming backups require.

Compute needs will change continually through a workload lifetime, use Amazon CloudWatch to collect key OS metrics (CPU, RAM, network, disk I/O) and review periodically for tuning of your Commvault control plane and data plane resources.

Containers

Commvault software cannot be deployed as a containerized application at this time. Commvault is a **Linux-based** or **Windows-based** control plane with one or many Linux and Windows data-plane nodes. Commvault is capable of protecting **Amazon EKS** workloads in the Region, AWS Outposts, and on-premises via Amazon EKS-Distro support.

Commvault offers the **Metallic.io** Backup as a Service (BaaS) SaaS solution where there is no requirement to perform data management application or server management. The metallic control plane operates outside of AWS with **customer-managed data plane** nodes within your AWS account(s).

Functions

Commvault software cannot be deployed as a Function or in Function-as-a-Service (FaaS) deployment patterns (see Containers above).

Related documents: PERF02-BP01 Evaluate the available compute options.

PERF02-BP02 Understand the available compute configuration options

Commvault has performed extensive testing across instance families, sizes, architectures, I/O optimization, bursting impacts, and protection methods to identify both the least cost and best price-performance instance for each stateful Commvault component.

Commvault recommends the use of current generation **Compute-optimized**, AWS Graviton-based instances for MediaAgent and Access Nodes responsible for data movement, running Amazon Linux 2.

Commvault also supports and recommends the use of **Memory-optimized**, **General-purpose**, and **Burstable Instances** (in isolated use cases).

See the **Design Principles and Best Practices – Sizing Guidelines** for current recommendations.

Commvault recommends rightsizing deployed instances based on **AWS Compute Optimizer** recommendations. Be sure to account for monthly, six-monthly, and yearly backup activities when rightsizing instances.

Understanding snapshot protection performance

Commvault software orchestrates the creation, copying, replication, recovery, and deletion of AWS native snapshots for Amazon EBS, Amazon RDS, Amazon Redshift, and Amazon DocumentDB resources. Compute instances responsible for snapshot management require minimal CPU and RAM (e.g., T4g.small instance size) to perform snapshot lifecycle management.

Commvault recommends scaling snapshot-only access nodes horizontally, subject to the maximum number of concurrent operations with a single AWS account and region (see **Well-architected Reliability – Workload Architecture** for additional details).

Additional performance for snapshot-only protection may be achieved by:

- Distributing backups across multiple distinct VM groups for concurrency
- Creating VM groups per AWS account and regions for concurrency
- Increasing the **number of data readers** used on the VM group (default=5)

> 💡 **Pro-Tip**
>
> If a VM group is configured with readers = 10 and a backup for fifty (50) Amazon EC2 instances with 2 disks each is initiated, Commvault software will attempt to create 50 Amazon Machine Images (AMIs) and 100 Amazon EBS snapshots. No more than ten (10) instances will be processed at any one time. Commvault will wait for AMI creation and subsequent Amazon EBS snapshot creation, before scheduling protection of instance 11.

This approach allows Commvault to perform more concurrent snapshot operations and fully consume the available instance resources.

At the time of writing, Commvault can create up to 3,600 snapshots per day, account, and Region.

This is based on the following lab-based observations, your environment may differ:

- An average Amazon EBS snapshot creation time is 1-2 minutes.
- An hourly average per AWS account and region is 30 snapshots.
- The daily average per AWS account and region is 3,600 snapshots.
- Maximum snapshot protection is dependent on running five concurrent snapshots per storage type (io1, io2, gp2, gp3, standard).

It is important to be aware that the **Amazon EBS service quotas** limit the number of concurrent snapshots per EBS disk type (gp2, gp3, io1, io2, standard, sc1, st1).

Understanding streaming protection performance

Commvault performs network streaming backups for Amazon EC2, Amazon EKS, Amazon RDS, Amazon DynamoDB, Amazon EFS, Amazon FSx for Windows, Amazon FSx for NetApp, Amazon S3, and any Amazon EC2 instance with traditional applications installed.

When performing a streaming backup, Commvault is dependent on the network bandwidth available to the Amazon EC2 instance. As network usage is typically very peaky, Commvault recommends using instance sizes that provide baseline and burst network performance (see **Amazon EC2 instance network bandwidth**) to save cost. To avoid using an instance with base and burst network performance, select an instance size of `4xlarge` or larger.

Commvault recommends using source-side deduplication and compression on the client or Access Node to reduce the volume of data that is transferred to the MediaAgent for storage in Amazon S3. Reducing the volume of data transferred via deduplication and compression conserves valuable network credits ensuring more value for your Amazon EC2 investment.

The performance of streaming backup will be affected by the workload instance **network bandwidth**, **EBS bandwidth**, and the Access Node **network bandwidth**.

Commvault recommends using **Amazon C6/C7 family** instances running Amazon Linux 2 exclusively for Access Nodes and MediaAgents performing streaming data transfers (see **Seed MediaAgent**, **Scale-out MediaAgent** sizing).

Commvault recommends that VPC endpoints for **Amazon S3** and **Amazon EBS** service endpoints are created to maximize streaming performance.

Related documents: PERF02-BP02 Understand the available compute configuration options.

PERF02-BP03 Collect compute-related metrics

Commvault collects and visualizes **Infrastructure Load** metrics for all Commvault components (CommServe, MediaAgent, Access Nodes).

Commvault recommends centralizing and aggregating Log Files, Metrics, and Events in **Amazon CloudWatch** for troubleshooting performance events across Commvault and application workload resources.

Commvault CommServes deployed from the AWS Marketplace come configured with **Amazon CloudWatch agent** and alarms pre-configured.

Commvault recommends tracking **EC2 default metrics**, **EC2 memory and disk metrics** (particularly voluntary and involuntary context switches), and **enhanced network adapter (ENA) metrics**.

Related documents: PERF02-BP03 Collect compute-related metrics.

PERF02-BP04 Determine the required configuration by right-sizing

Commvault has analyzed its software requirements against available instance types and sizes and recommends memory-optimized instances for the least cost and compute-optimized instances for the best price performance.

Commvault recommends current-generation instances which offer the best price-performance mixture of CPU, memory, and network in most cases. The use of previous-generation instances is possible and recommended for cost optimization and to meet regional instance availability constraints.

Your workload needs will change as the volume of data fluctuates, and the protection methods employed (snapshot-only, snapshot, and streamed). Use AWS Compute Optimizer to identify over-provisioned and under-provisioned resources, and then right-size accordingly.

Related documents: PERF02-BP04 Determine the required configuration by right-sizing.

PERF02-BP05 Use the available elasticity of resources

Commvault recommends decoupling your Access Nodes (which backup and restore workloads), from your centralized command-and-control instance, the CommServe.

Data movement components like MediaAgents and Access Nodes experience more demand for resources than the CommServe in normal operation.

Commvault **auto-scales Access Nodes** during backup operations to automatically scale out the number of resources to meet the backup RPO.

Commvault does not auto-scale resources for restore operations, consider a buffer-based approach that reserves a minimal headroom for MediaAgents to service multiple concurrent restore events.

Related documents: PERF02-BP05 Use the available elasticity of resources.

PERF02-BP06 Re-evaluate compute needs based on metrics

Commvault recommends taking a data-driven approach to continual re-evaluation and optimization of compute instances.

Use AWS Compute Optimizer with **enhanced infrastructure metrics** (additional cost) to gain a historical view of resource utilization over the last 3 months.

Use Amazon CloudWatch to record and potentially archive metrics with a much longer time window than Compute Optimizer.

If a resource sits under 40% utilized for four weeks or more, halve it (**Tips for Right-Sizing**).

If a resource exceeds 100% consumption and workload performance is not meeting business RPO/RTOs, increase the instance size to attain more resources (CPU, ram, network).

Remember that over-achieving on your business's agreed performance measures is an <u>over-provisioned</u> system and may require right-sizing.

Related documents: PERF02-BP06 Continually evaluate compute needs based on metrics.

Storage architecture selection

Commvault Backup and Recovery utilizes multiple storage solutions to perform holistic data management across all your AWS and edge locations. At a high level, Commvault recommends and utilizes the following AWS storage solutions:

- **Object storage** (Amazon S3) to store backup copies and backup copy replicas (auxiliary copies) for disaster recovery purposes.
- **Archival storage** (Amazon S3 Glacier) to store long-term retention, regulatory, or compliance-required archival data.
- **Block storage** (Amazon EBS) to store Commvault application binaries, logs, and databases used to manage and optimize data management.
- **File storage** (Amazon FSx for Windows, Amazon EFS) to store centrally managed and shared procedural documentation and software like *runbooks*, *playbooks*, and Commvault *software depots*.

Commvault simplifies the choice of Amazon EBS block storage by supplying AWS Marketplace products for an all-in-one Commvault Backup & Recovery solution, and scale-out Cloud Access Nodes. These instances are pre-configured with the Commvault recommended EBS volume types and sizes.

Shared storage (i.e., Amazon EFS or Amazon FSx for Windows File Server) may be used with Commvault software to provide a shared location for software binaries (i.e., a software depot) and historical logs. Commvault databases and application binaries expect high-performance low-latency block storage to meet the demanding needs of data protection at scale.

PERF03-BP01 Understand storage characteristics and requirements

Commvault publishes a list of **storage characteristics** per data type in the Cloud Architecture Guide (this document), and recommended Amazon S3 storage classes per use-case.

Commvault recommends the use of **Amazon EC2 gp3 block volumes** exclusively for optimal price performance, **tuneability**, and sustainability compared to magnetic storage.

Commvault **Disaster Auto backups** can be located on shared file-systems **Amazon EFS** or **Amazon FSx for Windows Server** for decoupling from the EC2 instance.

Commvault does not require or utilize **instance store** storage.

Commvault data management has different requirements based on the data type being read or written. See below for the storage characteristics to consider for each Commvault data type, and the Commvault recommended storage product:

Commvault data	Latency	Throughput	Shareable	Recommended
Microsoft SQL Server database, tempdb, binaries, logs	Low, consistent	Single	N	Amazon EBS (**gp3**)

Commvault **software binaries**, **logs**, **SoftwareCache**, **DR backups**, **Private metrics Reports**, **Cloud Metrics uploads**, **Download Center packages**, **temp directory**	Low	Single	N	Amazon EBS (**gp3**)
Deduplication Database	Very-low, consistent	Single	N	Amazon EBS (**gp3**)
IndexCache	Very-low, consistent	Single	N	Amazon EBS (**gp3**)
Job Results **Job Results on UNC Path**	Low	Single	N	Amazon EBS (**gp3**) Amazon EFS (**Gen, Max**) Amazon FSx for Windows
Content Indexing and Search index	Low	Single	N	Amazon EBS (**gp3**)
Continuous Data Replicator Logs	Low	Single	N	Amazon EBS (**gp**)
Cloud library (Backups) – active	Low-latency	Web-scale	Y	Amazon S3 (**Intelligent-Tiering**)
Disk Library (Backups) – active	Low-latency	Single	N	Amazon EBS (**gp**)
Disk Library (Backups) – infrequent	Low	Single	N	Amazon EBS (**st1**, **sc1**)
Disk Library (Backups) - active	Low	Multiple	Y	Amazon EFS (**Gen, Max**)
Disk Library (Backups) – infrequent	Low	Multiple	Y	Amazon EFS (**Std IA**)
Cloud library (Archives)	Minutes to hours	Web-scale	Y	Amazon S3 (**Glacier Instant Retrieval**, **Glacier Flexible Retrieval**, **Glacier Deep Archive**)
Disaster Recovery Backups	Low	Single	Y	Amazon EFS (**Std IA**) Amazon FSx for Windows
O365 **Shared job results** path (OneDrive, SharePoint, Exchange Online, Teams)	Low	Multiple	Y	Amazon FSx for Windows

Individual Commvault data management systems will have unique IOPS and Throughput requirements based on business protection SLAs. The following tables provide the Commvault required block size, file size, and IOPS if Commvault mandates a minimum specification. Values are specified as the number of I/O operations per second (IOPS) for a given block size (size per I/O operation). See **How does Amazon EBS calculate the optimal I/O size I should use to improve**

performance on my gp2 or io1 volume?

Linux MediaAgent Instances

Commvault data type	Initial day one IOPS	Initial day one volume throughput (MiB/sec)	Block size
Microsoft SQL Server database, tempdb, binaries, logs (**best practices**)	150	125	**64K**
Deduplication Database *gp3 baseline performance will manage 600TB of written deduplicated data.*	3,000 Plan for 1,000 IOPS per 200TB written	125	**4K**
IndexCache *gp3 baseline performance will manage 3PB of written deduplicated data.*	3000 Plan for 1000IOPS per 1PB written	125	**32K**
Disk Library (Backups) *Commvault recommends a volume throughput of 160MiB/sec or greater.* *Any Amazon EBS volume type may be used as a disk library location.*	No specific IOPS requirement.	160 **recommended** (not mandatory)	**64K**
Search Engine node (context indexing) *gp3 baseline performance is sufficient to service search engines (see **recommendation**).*	3000	125	32K
Search Only node (search only, no new data) *gp3 baseline performance is sufficient to service search-only search engines (see **recommendation**).*	3000	125	32K
Search Engine Node with ContentStore Email Viewer *gp3 baseline performance is sufficient to service search with email viewer (see **recommendation**).*	3000	125	32K

Microsoft Windows MediaAgent Instances

Commvault data type	Initial day one IOPS	Initial day one volume throughput (MiB/sec)	Block size
Microsoft SQL Server database, tempdb, binaries, logs (**best practices**)	150	125	**64K**
Deduplication Database *gp3 baseline performance will manage 600TB of written deduplicated data.*	3,000 Plan for 1,000 IOPS per 200TB written	125	**32K**
IndexCache *gp3 baseline performance will manage 3PB of written deduplicated data.*	3000 Plan for 1000IOPS per 1PB written	125	**32K**
Disk Library (Backups) *Commvault recommends a volume throughput of 160MiB/sec or greater.* *Any Amazon EBS volume type may be used as a disk library location.*	No specific IOPS requirement.	160 **recommended** (not mandatory)	**64K**
Search Engine node (context indexing) *gp3 baseline performance is sufficient to service search engines (see **recommendation**).*	3000	125	32K
Search Only node (search only, no new data) *gp3 baseline performance is sufficient to service search-only search engines (see **recommendation**).*	3000	125	32K
Search Engine Node with ContentStore Email Viewer *gp3 baseline performance is sufficient to service search with email viewer (see **recommendation**).*	3000	125	32K

Related documents: PERF03-BP01 Understand storage characteristics and requirements.

PERF03-BP02 Evaluate available configuration options

Commvault recommends that the storage characteristics and requirements (above) influence the selection of storage technologies and tuning to meet required I/O performance. Consider the following storage configuration options for AWS storage:

- Leverage **Amazon EBS gp3** SSD-backed volumes <u>exclusively</u> for active data to ensure that capacity, IOPS, and throughput can be tuned independently over the life of the volume.

- Commvault does not recommend the use of **Amazon EBS Throughput Optimized** (st1) HDD volumes or **Amazon EBS Cold (sc1)** HDD volumes as the majority of Commvault I/O is small-block random access (see **Inefficiency of small read/writes on HDD**).
- Commvault recommends for small remote-office branch-office (ROBO) deployments that do not require local backup copies, use Commvault **Storage Accelerator** to read and write backups to a remote Amazon S3 bucket. Commvault supports the use of **Amazon S3 Transfer Acceleration** where required.
- Alternatively, when workloads in a region required local recovery copies, deploy a regional **power-managed MediaAgent,** maintain a local recovery copy, and perform an **Amazon S3 cross-region replication (CCR)** to provide a remote **read-only replica library**.
- Commvault recommends reviewing Amazon EFS backup and recovery performance against business SLAs for *recovery time* to determine if a **general purpose or max I/O** performance mode is required. Commvault performs backup and recovery of EFS directly over the front-end protocol interface and therefore contributes to IOPS and throughput demand.

Commvault recommends reviewing Amazon FSx for Windows Server backup and recovery performance against business SLAs for *recovery time* to determine if **file-system throughput** requires tuning. Commvault performs backup and recovery of FSx directly over the front-end protocol interface and therefore contributes to IOPS and throughput demand

Commvault publishes a list of **storage characteristics** per data type in the Cloud Architecture Guide (this document), and recommended Amazon S3 storage classes per use-case.

Commvault consists of three primary storage types or requirements.

- **Block-mode storage** provided by SSD-backed Amazon EBS for software binaries, SQL Server database, logs, indexes, job results, and deduplication databases optimized for low-latency transactional tasks.

- **Object storage** provided by Amazon S3 for backup and archival data, optimized for durable, elastic multi-PB storage pools.

- **File storage** provided by Amazon EFS and Amazon FSx for shared storage use-cases like runbooks/playbooks, disaster recovery backups, and software install depots, optimized for capacity growth.

Block-mode storage

Commvault recommends the use of Amazon EBS gp3 block-based storage exclusively for the best price-performance of SSD-backed volumes along with the tunability of IOPS and throughput independent of capacity.

Commvault supports gp2, io1, io2, standard, st1, and sc1 volume types but does not recommend their use for optimal price-performance for Commvault workloads.

Commvault does not recommend using block-mode EBS volumes for backup or archive data, Amazon S3 is the most appropriate storage type for backup data which is infrequently accessed and grows indefinitely.

Data Protection of Amazon EBS volumes

Commvault provides multiple transport modes used to perform protection of Amazon EBS volumes.

Commvault recommends using **Amazon EBS direct APIs** exclusively, due to improved performance and avoiding the need to mount and unmount volumes to an Access Node to capture item-level information.

> 💡 **Pro-Tip**
>
> In environments where you need a volume to have production-level performance immediately after restoration, use HotAdd transport mode, which will create an empty volume and mount it to an Access Node to recover the data.

Transport Mode

Commvault supports several different methods of performing backup and recovery. When protecting Amazon services Commvault recommends the use of **EBS Direct APIs** for backup and recovery (Transport: EBS-DIRECT-API), this removes the requirement to mount Amazon EBS volumes onto Commvault Access Nodes to perform data movement.

Considerations when using HotAdd Transport Mode

When performing backup or recovery using HotAdd transport mode a **limit of twenty-six (26) concurrent volume mounts** on an EC2 instance will limit the number of concurrent data management actions (backup, restore).

Commvault recommends utilizing **EBS direct read backup method** to avoid this limit. This limit applies to Microsoft Windows and Linux-based MediaAgents HotAdd backup and restores.

EBS optimization

The **Virtual Server Agent** is responsible for protecting client data held in EBS volumes, and streaming backup data to S3 or equivalent Cloud Libraries. The **EBS bandwidth, throughput, and IOPS** as detailed at **Amazon EBS – Optimized Instances** are crucial to backup and restore performance.

> ⓘ **Note**
>
> **Amazon EBS-optimized instances** can support maximum performance for 30 minutes at least once every 24 hours. Commvault recommends the use of EBS-optimized instances as the backup workload is typically peaky and benefits from the reduced cost offered for EBS-optimized instances.

Performance considerations – Streaming backup

For standalone Access Nodes (without the MediaAgent package), monitor the total **Front-End Terabytes (FETB)** protected by a single Access Node and either scale up the specification of an existing Access Node, or add an Access Node for more backup throughput to your MediaAgents. Remember that each EC2 instance provides additional network and **EBS volume bandwidth and network bandwidth**, so scaling horizontally with smaller Access Nodes is recommended over vertical scaling.

> ⓘ **Note**
>
> When utilizing Amazon EBS direct APIs to perform the backup, the network bandwidth is more important than the underlying EBS volume bandwidth.

Commvault defaults auto-scaled Access Nodes to a **c6g.large** instance (default) for optimal cost and performance

Offline data migration (Cloud Seeding)

Cloud Seeding is the process of moving an initial set of data (backups, archives) from their current location to Amazon S3 in a method or process that is different from regular operations. There are two primary methods for seeding data to AWS:

On The Wire (Online Seeding)

This is typically performed in a small logical grouping of systems to maximize network utilization to complete the data movement more quickly per system. Some organizations will purchase "burst" bandwidth from their network providers for the seeding process to expedite the transfer process.

Amazon VPC provides multiple methods for connecting edge-based locations to the AWS cloud including **AWS Direct Connect** and multiple **AWS VPN** options.

Please see the chart below for estimated payload transfer time for various data sizes and speeds.

LINK SIZE								
	1 GB	10 GB	100 GB	1 TB	10 TB	100 TB	1 PB	10 PB
10 Mbit	14 min	2.2 hrs.	22.2 hrs.	9.2 days	92.6 days	-	-	-
100 Mbit	1 min 20 s	13 m 20 s	2.2 hrs.	22.2 hrs.	9.2 days	92.6 days	-	-
1 Gbit	8 s	1 m 20 s	13 m 20 s	2.2 hrs.	22.2 hrs.	9.2 days	92.6 days	-
10 Gbit	0.8 s	8 s	1 m 20 s	13 m 20 s	2.2 hrs.	22.2 hrs.	9.2 days	92.6 days

Drive Seeding (Offline Seeding)

If the data set is too large to copy over the network, or transport over the network is too costly, then physical drive seeding is a valid alternative option. Drive seeding is copying the initial data set to external physical media and then shipping it directly to the external cloud provider for local data ingestion.

Please refer to **Seeding a Cloud Storage Library** for how to perform drive-seeding with Commvault software.

Amazon provides the following offline seeding technologies, all of which are supported by Commvault (per *Seeding a Cloud Storage Library* process above).

- Snowball Edge **AWS Snowball Edge**
- Snowmobile **AWS Snowmobile**
- Snowcone **AWS Snowcone**

A note on AWS Snow family and using Commvault Combined Storage Tier

AWS Snow family devices <u>cannot</u> be used to seed Commvault **Combined Storage Tier** libraries as the Snow devices only support a single S3 storage class. If looking to migrate a large amount of data with AWS Snow family, migrate into Amazon S3 and then perform an **auxiliary copy** of desired data to a Commvault combined storage library. Be sure to model and perform a minimal dataset test to validate the costs of the full dataset read from Amazon S3 using the **AWS calculator** to understand the costs of the copy.

Related documents: PERF03-BP02 Evaluate available configuration options.

Commvault recommends using **AWS Compute Analyzer** to monitor and receive **recommendations** for right-sizing Amazon EBS volume configuration on Commvault control plane and data plane components. Commvault recommends activating **Enhanced infrastructure metrics** (paid features) to use metrics for as long as three months (93 days) when assessing volume utilization patterns.

Commvault provides instructions for baselining and reevaluating block volume performance that meets Commvault requirements, using the open-source **iometer** or **diskspd** tools. See the following topics for guidance:

- **IOPS for Deduplication Database Volumes**
- **Testing IOPS of Search Engine Notes with IOmeter**
- **Using Iometer to Test Index Cache Directory Performance**
- **Disk Library Volume Performance**

Commvault provides a dedicated I/O test tool called **CVDiskPerf** when deploying disk or network-attached storage (NAS) libraries. Commvault does not recommend using block or file-based storage for backup data in AWS, Amazon S3 provides a scalable, durable, high-performance alternative.

Commvault activates Amazon CloudWatch free disk usage monitoring and alerting on your Commvault CommServe® control plane when deploying the Commvault Backup & Recovery AMI (paid, BYOL) from **AWS Marketplace**. Implement Amazon CloudWatch metrics on all storage volumes, file systems, and buckets so that data-driven tuning can be performed.

Commvault recommends **tracking Amazon CloudWatch EBS storage metrics** to understand the sustained and peak load being experienced by high-performance block-mode volumes in addition to traditional file-system utilization metrics.

Monitor the **Amazon EBS metrics for Nitro-based instances** to understand when EBS burstable IOPS, EBSIOBalance%, and EBSByteBalance% credits are exhausted and performance drops back to baseline.

Consult the **EBS optimized by default table** to understand the baseline and burst EBS bandwidth, throughput, and IOPS available for all EBS volumes on an instance.

Block volumes that demonstrate low-performance utilization may be candidates for migration to shared file systems like Amazon EFS or Amazon FSx.

Ensure that tracking of fixed-size storage solutions like Amazon EBS and Amazon EFS alarm utilization low and high-watermarks, and ideally **automatically grow storage**.

Database architecture selection

Commvault utilizes multiple database sub-systems including Microsoft SQL Server, MongoDB, C-Tree, Redis, and Apache Solr. All databases are built-in and distributed with Commvault software and cannot be externalized into AWS-managed cloud database services.

Related documents: PERF03-BP03 Make decisions based on access patterns and metrics.

Commvault uses several database technologies that run on Commvault CommServe and MediaAgent components. These components cannot be relocated to fully managed cloud databases.

Commvault publishes storage performance requirements in IOPS for day-one seed deployments. Track the I/O Operations per Second (IOPS) from Commvault instances and observe growth in storage performance demand.

Commvault highlights unhealthy DDB storage components in the **Health Report: DDB Performance and Status**.

Commvault recommends that the Query & Insert (Q&I) time on a DDB not exceed 2 milliseconds (see **Average Q&I Time in last 3 days**).

Commvault provides the Subclient Association Limits for Deduplication Alert to notify when Q&I time exceeds best practice. Consider using a **command-line notification** to increase the IOPS and throughput using the **aws ec2 modify-volume** command.

Related documents: PERF04-BP01 Understand data characteristics.

PERF04-BP02 Evaluate the available options

Commvault database technologies can be considered embedded databases and cannot be tuned by selecting alternative cloud database technologies.

Options to improve the performance of Commvault instance-lock block-mode databases are to **increase the IOPS and throughput** for a given volume or to increase the instance size to receive additional **EBS instance baseline and burst bandwidth** (bandwidth applies to I/O from all EBS volumes on the instance).

Related documents: PERF04-BP02 Evaluate the available options.

PERF04-BP03 Collect and record database performance metrics

Commvault recommends using the **Amazon CloudWatch metrics for Amazon EBS** for granular visibility of IOPS, throughput, and latency from block-mode volumes.

Volume metrics for volumes attached to Nitro-based instance types provide `EBSReadOps`, `EBSWriteOps`, and `EBSIOBalance%`, `EBSByteBalance%` metrics that provide insight into the exhaustion of burst capacity.

Track these metrics and include them in technical operational **CloudWatch dashboards** for incident response and proactive performance optimization.

Related documents: PERF04-BP03 Collect and record database performance metrics,

PERF04-BP04 Choose data storage based on access patterns

Commvault does not leverage or utilize cloud databases for its internal embedded databases.

Access patterns to Commvault databases (SQL Server, MongoDB, Deduplication Database, IndexCaches) are random with specific block sizes per structured database type.

Related documents: PERF04-BP04 Choose data storage based on access patterns.

PERF04-BP05 Optimize data storage based on access patterns and metrics

Commvault does not leverage or utilize cloud databases for its internal embedded databases.

Alternative indexing, key distribution, data warehouse design, or caching strategies cannot be applied to Commvault structured databases.

Related documents: PERF04-BP05 Optimize data storage based on access patterns and metrics.

Network architecture selection

Commvault has a significant dependency on the latency, throughput, and bandwidth of network connectivity in a distributed data management deployment. Consider the following considerations when designing and selecting

network components for use with Commvault software:

Commvault data management has three primary areas where network performance impacts the customer experience:

- **MediaAgents and Access Nodes** must be able to access protected workloads via **service endpoints**, **VPC endpoints,** or application-specific port/protocol for backup and recovery. Commvault recommends at least one Cloud Access Node (MediaAgent, Access Nodes) per region to reduce latency when accessing regional data copies.

- **End-user access** to the Commvault Command Center™ web-based portal for self-service backup, recovery, and DR failover activities should originate from the CommServe-located region.

- **Edge-based workload** backups/restores, and backup copies to/from the AWS region must have sufficient bandwidth to support deduplicated and fully rehydrated restores.

Commvault Backup & Recovery and **Commvault Auto Recovery (formerly Disaster Recovery)** are dependent on the ability to transfer backup copies or Point in Time (PiT) backup data to Commvault, to replicate that content between Regions, and to perform restores. Network latency, bandwidth, and availability directly impact how quickly data can be captured (i.e., *recovery point objective*) and how quickly data can be restored (i.e., *recovery time objective).* Commvault recommends and uses Amazon EBS direct APIs for the backup and recovery of Amazon EBS volumes. Network throughput available to the MediaAgent performing the transfer will directly impact protection throughput:

Commvault will use your **MediaAgent grid** Access Nodes first (R6, M6 instances), then *auto-scale* additional resources (C7 instances).

MediaAgent Instance Size	Network Bandwidth	GB/hr. (using EBS direct API)***	
		Commvault avg. backup throughput	Commvault avg. restore throughput
c7g.xlarge*	Up to 12.5 Gigabit**	156	183
c7g.2xlarge*	Up to 12.5 Gigabit**	276	184
c7g.4xlarge*	Up to 12.5 Gigabit**	288	186
c7g.8xlarge*	Up to 12.5 Gigabit**	283	188
c7g.12xlarge*	Up to 12.5 Gigabit**	284	190
c7g.16xlarge*	Up to 12.5 Gigabit**	281	175

* **Amazon EBS-optimized instances** can support maximum performance for 30 minutes at least once every 24 hours.
** Amazon EC2 instance has a **baseline and burst network performance**, test in your environment for achievable throughput.
*** Avg. throughput observed in tuned config using **ReadAhead**=256, **WriteBehind**=256 for optimal transfer from EBS direct.
Backup testing was performed in us-east-1 with 24.5GiB & 88.08GiB dataset size and VPC endpoint and 16 threads.
Restore testing was performed in us-east-1 with 53.57GiB dataset size (gp3, 16K IOPS) and VPC endpoint and 16 threads.

For environments where the Amazon EC2 instance(s) under protection and the Commvault Access Node are in the same Availability Zone, Commvault recommends using **Commvault HotAdd** for optimal backup and restore performance. Commvault creates temporary Amazon EBS io2 volumes during backup activity to accelerate reads and achieves faster backup speeds than EBS direct API backup in most test-cases.

Commvault control plane connectivity is also crucial to orchestrating, monitoring, and completing data management activities using AWS native snapshots and optionally network-streamed copies into Commvault-controlled object storage. If the network becomes unavailable during a data management activity, retries will occur, but repeated outages will mark the protection job as failed.

Clients remote to the Commvault control plane and data plane infrastructure can utilize **client-side deduplication database (DDB)** caches to reduce network demands for backup. Recovery activities are always performed in a fully hydrated format, so recovery of a full system will require more bandwidth than backup activity. See **Enable source-side cache** for more information.

Regional Commvault infrastructure deployed in fault-tolerant groups or grids <u>does not</u> require very low-latency network connections. Nodes may be distributed across Availability Zones and receive sub-millisecond latency between independent locations.

Likewise, **Commvault CommServe® LiveSync** used to replicate Commvault control plane state occurs asynchronously and is not dependent on low-latency networks between instances.

Related documents: PERF05-BP01 Understand how networking impacts performance.

PERF05-BP02 Evaluate available networking features

Commvault recommends selecting and implementing the following networking features for optimal data management transfer times:

- Use Amazon EC2 **nitro-based** systems with enhanced security, reduced latency, and increased packet per second (PPS) performance of nitro-presented **Elastic Network Adapter (ENA)**. Network connectivity for Commvault data plane nodes (MediaAgent, Access Nodes) directly impacts the speed of backup, restore, and replication jobs.
- <u>Do not</u> use **Amazon EC2 enhanced networking** N-series instance types with 25Gbps and 100Gbps bandwidth. Commvault is engineered to horizontally scale with multiple smaller nodes with 5 Gigabit, 10 Gigabit, and 12.5 Gigabit of network connectivity. Commvault data plane nodes will not benefit from additional bandwidth of more than 12.5 Gigabit.
- Commvault Access Nodes deployed from the AWS Marketplace (x86_64, ARM64) receive **Jumbo Frames (MTU: 9001)** by default when using VPC endpoints to access EBS or S3 services (at the time of writing).
- Create an **Amazon S3 Gateway endpoint** to maintain an MTU=8400 for optimized data transfer between Commvault data plane instances and the Amazon S3 service.
- Alternatively, create an **Amazon S3 Interface endpoint** to maintain an MTU=9001 between Commvault data plane instances and the Amazon S3 service.

⚠ Important

    If an Amazon S3 endpoint is not created then transfers to and from the Amazon S3 service endpoint will traverse your Internet Gateway (IGW) or NAT Gateway (NGW) and receive an MTU=1500.

See **Check the path MTU between two hosts** for validating your environment as MTU is affected by inter-region VPC Peering (MTU=1500) and **Transit Gateway** (MTU=8400)

Ensure your Commvault data plane and optionally workload Amazon EC2 instances are running the latest Enhanced Network Adapter (EBA) drivers. See **ENA Linux Driver Best Practices and Performance Optimization Guide** for details on using the ENAv3 driver.

Commvault requires **enhanced networking** for MediaAgents and Access Nodes performing data transfer for backup or recovery (vs. snapshot-only protection).

Commvault recommends the use of **AWS Transit Gateway** but also supports **VPC Peering** for connecting VPCs, AWS accounts, and on-premises networks.

Commvault recommends the use of **AWS Direct Connect** to connect on-premises to AWS. The use of **AWS Site-to-Site VPN** and **AWS Client VPN** is also supported with the acceptance of **maximum throughput limits**.

Commvault recommends the use of **AWS PrivateLink** for connectivity to Amazon S3, and Amazon EBS service endpoints (at a minimum) for security and optimal backup and restore performance.

Commvault does not require but can leverage Amazon Route 53 and DNS failover for **public addresses**, and **private addresses** in coordination with CommServe **automatic failover**.

Commvault can leverage Elastic Load Balancers (ELBs) to balance end-user HTTP requests across pools of Commvault Web consoles for high availability of Commvault web services.

Related documents: PERF05-BP02 Evaluate available networking features.

PERF05-BP03 Choose appropriately sized dedicated connectivity or VPN for hybrid workloads

Consider **Amazon Direct Connect** for Production connectivity to on-premises locations with dedicated connectivity available in 1 Gbps, 10 Gbps, and 100 Gbps options.

Consider AWS Site-to-Site VPN as a backup connection for **Non-Production or Development** workloads. Be aware that using AWS Site-to-Site VPN connectivity has two tunnels and each tunnel supports a maximum throughput of up to 1.25 Gbps. If your backup and recovery activities require greater throughput to meet business *recovery time objectives*, utilize AWS Transit Gateway by **implementing equal-cost multi-path (ECMP) over multiple VPN tunnels** or use AWS Direct Connect.

Commvault backup activity to deduplicated Cloud libraries will benefit from client-side deduplication and compression, reducing the data placed on the network. Commvault auxiliary copies of backup data to remote locations with deduplication enabled (i.e., **DASH copies**), will also benefit from reduced network impact due to deduplicating content before placing on the wire.

Commvault restores are re-hydrated or retrieved in a non-deduplicated format, you should ensure that periodic restore testing is performed to ensure that the network can meet your *recovery time*. Consider maintaining local backup copies on-premises using **AMAZON S3 on Outposts** or **Commvault HyperScale™ X scale-out storage appliances** to avoid over-the-wire restore delays.

Commvault recommends establishing a cross-functional team to establish the hybrid networking architecture bandwidth for each site to be connected to AWS using AWS Direct Connect.

Commvault specifically, should be sized with a measured understanding of the daily backup volume, restore volume, and business-required transfer time.

Be sure to plan for both backup (which is deduplicated, compressed, and incremental forever), and restores (which is fully rehydrated before transfer over the network).

Commvault recommends maintaining a localized data store to avoid rehydrated restores occurring over Direct Connect. Consider **Commvault HyperScale X™** scale-out storage or **Amazon S3 on Outposts** for this purpose.

Direct Connect connections are available from 50Mbps up to 10Gbps with multiple connections required to exceed 10Gbps. Consider designing for a Direct Connect partner outage with additional capacity, or failover to Site-to-Site VPN over the Internet.

Related documents: PERF05-BP03 Choose appropriately sized dedicated connectivity or VPN for hybrid workloads.

PERF05-BP04 Leverage load-balancing and encryption offloading

Commvault can utilize **Elastic Load Balancers (ELBs)** to balance HTTP requests to the **Commvault Command Center™** or **Commvault service endpoint**.

Commvault Web consoles may be deployed across Availability Zones (AZs) for **performance and resilience**.

An **Application Load Balancer** (ALB) will distribute incoming requests amongst healthy instances with sticky sessions (session affinity, persistent sessions) enabled. ALBs can perform SSL offloading to reduce the impact on the Webconsole CPU.

Additionally, Commvault Command Center™ the web-based administrative console (i.e., **WebConsole**) can be placed behind an **Application Load Balancer (ALB)** to load-balanced and route incoming requests based on user location, load, or service endpoint being accessed. Commvault recommends assessing the cost-benefit of load-balancing as the performance benefits will only be realized in high concurrent user environments (i.e., managed service provider environments).

Commvault does not require Network Load Balancer (NLBs) configurations.

Commvault software will automatically load-balance data protection jobs across available **Access Nodes**, MediaAgents, and **Cloud Libraries**. This load-balancing will allow leveraging the network resources attached to each Commvault data mover.

Related documents: PERF05-BP04 Leverage load-balancing and encryption offloading.

PERF05-BP05 Choose network protocols to improve performance

Commvault uses the Transmission Control Protocol (TCP) exclusively for all data transfers, as detailed in **Port Requirements for Commvault**.

Commvault requires retransmission (**rfc793**) and high-reliability features of TCP, with the acceptance that this adds overhead for the processing of packets.

Commvault software communication utilizes Transmission Control Protocol (TCP) exclusively for control plane and data plane connections. There is no option to configure Commvault communication for User Datagram Protocol (UDP) communication. See **Port Requirements for Commvault** for details on the use of TCP and UDP when communicating with protected workloads.

Related documents: PERF05-BP05 Choose network protocols to improve performance.

PERF05-BP06 Choose your workload's location based on network requirements

Commvault recommends co-locating protection infrastructure where most end-users are located. Consider the following best practices:

- Place the **CommServe** where most of the end-user access will occur to the Commvault Command Center™ user interface

- Place **MediaAgents** in each region where data resides to be protected and recovered. This is a self-contained, isolated fault domain with rapid regional recovery.

- Place **Access Nodes** in each region where data resides to be protected and recovered. Access Node software can be co-located with the MediaAgents to reduce the initial compute footprint.

- Place **MediaAgents** and **Access Nodes** in edge-based availability zones (**AWS Local Zones**, **AWS Outposts**, **AWS Wavelength**) to avoid rehydrated data restore latency and throughput.

Avoid performing backup or recovery across regions, instead backup within the region and then use a **DASH copy** to replicate the backup to a remote region for Disaster Recovery.

Commvault does not require the use of **EC2 placement groups** and low-latency 25Gbps networks for high-performance backup or recovery.

Commvault does not require Amazon EC2 instances with network bandwidth exceeding 12.5Gbps.

Commvault software protects Amazon EC2, Amazon EKS, and Amazon RDS workloads located in the Region, on **AWS Outposts**, **AWS Local Zones,** and AWS Wavelength.

Commvault can orchestrate snapshot protection in remote or edge-based services (i.e., **AWS Local Zones – Supported Services**), and then optionally perform a streamed backup to Amazon S3 storage locations in the Region or at the edge on AWS Outposts. Backup data that must reside in the edge location may be written to edge-based **Amazon S3**, **Amazon EBS**, and **Amazon FSx** libraries if required.

Consider the following best practices for achieving optimal network throughput and latency for your AWS workload data management:

- Locate the *primary backup copy* in an Amazon S3 Cloud Library in the same Region as the protected workloads.
- Locate the Commvault Access Node and/or MediaAgent in the same Region as the protected workload.
- Ideally, locate the Commvault Access Node in the same Availability Zone as the protected workload to avoid data transfer charges for in-region cross-AZ communication (see **Amazon EC2 Pricing – Data Transfer within the same AWS Region**), set the **bEnableHostDispatch**=1 additional setting to ensure backups select instances in same AZ.
    - *Commvault will only select Access Nodes in the same AZ when bEnableHostDispatch is enabled.*
    - *If no Access Nodes in the same AZ are available, Commvault will select Access Nodes from the same region, incurring cross-AZ data transfer fees.*

> ⚠ Important
>
> Commvault will not perform zonal matching for data transfers to MediaAgents (backups) or restores. These data management activities may occur across availability zones (AZs) incurring data transfer costs based on the location of the MediaAgent, Access Node, and workload.

Commvault does not require placement of MediaAgent nodes, or Access Nodes in **Placement groups** for low-latency network communications. Round trip time (RTT) between Availability Zones (AZs) is considered acceptable for MediaAgent grids and Access Node groups which load-balance data management activity within a Region.

Related documents: PERF05-BP06 Choose your workload's location based on network requirements.

PERF05-BP07 Optimize network configuration based on metrics

Continually monitor your network configuration to allow timely identification and resolution of network-based anomalies. Typically, any network outage or impact will be directly reflected in the workload *recovery point objective* and *recovery time objective degradation.* Enable VPC Flow logs for networks used to perform Commvault data transfers for additional insight.

Commvault recommends enabling **Amazon VPC Flow Logs** to improve the observability of network communication, reachability, and security inspection.

Leverage Amazon CloudWatch Metrics, particularly **Enhanced Network Adapter (ENA) metrics** to observe, visualize and trend network resource utilization is an indicator of when an instance should be scaled for more bandwidth.

Extend metrics tracking for core network resources including NAT gateways, transit gateways, and VPN tunnels for a complete view of network performance.

Related documents: PERF05-BP07 Optimize network configuration based on metrics.

Review

Evolve your workload to take advantage of new releases

The Review process intends to:

- **Stay up-to-date on new resources and services** that can improve the performance of your Commvault data management platform or the internal performance of your team. Improvements can improve network efficiency, monitoring, and alarming processes, or automated test processes.

  As a team, define how you will stay informed of new AWS resources and services, and how they will be safely tested without impacting business SLAs.

- **Define a process to improve workload performance**
  Identify the technologies and services that directly impact your business SLAs for data protection. These are typically compute and containers, storage, and network related but may include observability metrics and alarming. Use this information to filter and prioritize which technologies to validate.

- **Evolve workload performance over time**
  Continually drive the adoption of recent technologies that can benefit your Commvault data management platform experience. Some examples include the adoption of Amazon EBS General Purpose 3 (gp3) volumes for improved performance and tuneability and the adoption of ENAv3 network drivers for enhanced network performance in sixth-generation EC2 instances.

PERF06-BP01 Stay up-to-date on new resources and services

Stay up to date on new networking announcements by periodically checking the **AWS What's New** announcements, particularly <u>Networking & Content Delivery</u> and <u>Compute</u>.

Commvault also publishes updates in the Cloud Architecture Guide (this document) and **Newsletter for New Features in Commvault Platform Release 2022E**.

Test new technologies, features, and enhancements in a pre-production environment and validate improvement against production baselines before wider deployment.

Continually look to review and optimize your selection of technologies used for your Commvault data management platform. This involves staying current with Commvault platform releases and maintenance releases (**What's New**) as well as reviewing the current **Commvault AWS Cloud Architecture Guide** recommendations.

Consider the following approaches to improving the performance and reliability of implementing change as new technology advancements become available:

Infrastructure as code

Use Infrastructure as Code (i.e., **AWS CloudFormation**) to define and maintain your Commvault data management platform infrastructure. Commvault supplies Amazon Machine Images (AMIs) for an all-in-one **Commvault Backup & Recovery Bring Your Own License (BYOL)** deployment complete with an AWS CloudFormation template to deploy into an existing VPC.

Speak with your Commvault sales representative to learn more about how Commvault has helped customers leverage **Commvault workflows** to upgrade Commvault by redeploying from updated AMIs and then reconnecting to Commvault persistent state EBS volumes (i.e., repaving).

Deployment pipeline

Commvault can be orchestrated by your continuous integration/continuous deployment (CI/CD) pipeline to perform configuration management, testing, and then rollout to Production. Use the Commvault command line, Python SDK or REST API to make changes in a repeatable, consistent fashion.

Well-defined metrics

Deliver consistent performance through continual monitoring of business and infrastructure metrics to ensure SLAs are met, and anomalies are actioned before impacting the business. Some areas to monitor include:

- Elapsed time to perform backup activity as compared to business SLA (per application, per region)
- Elapsed time to perform restore activity as compared to business SLA (per application, per region)
- CPU load, memory consumption, network consumption, and EBS I/Os per second on the Commvault control plane and data plane infrastructure.

Metrics must be measured for an acceptable threshold to meet business metrics, but also an acceptable threshold to ensure that infrastructure investment is appropriately sized.

Performance test automatically

Deployment processes should re-baseline the performance of the Commvault data management system after any change. This allows observing changes in backup and restore times, as well as the impacted infrastructure metrics. Commvault can be used to orchestrate a recovery of a subset of applications (i.e., data types) into an isolated VPC, then perform production-grade protection away from the production application. Testing should occur on production-equivalent instances but may use Spot instances to reduce the cost of testing.

Load generation

Load testing for your Commvault Backup & Recovery platform consists of adequate source data to represent production-grade data management and concurrency. Concurrency or parallel data management activities should ideally match the typical workload experienced during the nightly backup window. Ensure that sufficient data management activity can be scheduled to test prioritization and the impact of resource exhaustion.

Performance visibility

Commvault tracks and visualized key metrics like Service Level Adherence (SLA), the number of jobs that failed with error, Number of jobs that failed with a warning. Historical tracking is provided by the **SLA Trend Report** which can be accessed using Single Sign On (SSO) by authorized users or user groups.

The SLA trend report may be scheduled to be sent in CSV format to an extract transform and load (ETL) script or application to extract relevant metrics and publish them to Amazon CloudWatch. Commvault does not supply an example script or application to perform this extraction and push.

Visualization

Commvault provides rich metrics, reports, and trending through **Private Metrics Reports**. Reports include visualizations and interactive dashboards allowing filtering to view specific workloads or periods. Additionally, **Reports Builder** may be used to create custom reports and interactive dashboards that visualize the rich metadata that Commvault collects about protected workloads.

If centralized reporting and **analytics** are performed using Amazon CloudWatch, reports may be scheduled and exported in CSV format for parsing and ingesting into relevant Amazon CloudWatch metrics and relevant **CloudWatch Dashboards**.

Related documents: PERF06-BP01 Stay up-to-date on new resources and services.

PERF06-BP02 Define a process to improve workload performance

Commvault recommends maintaining a register of known key performance constraints within your distributed hybrid network architecture. Continually review your key constraints against AWS and Commvault-announced resource support and enhancements to look for opportunities to improve.

Commvault software is affected by the baseline and burst bandwidth available to a specific EC2 instance family and size (see **EBS optimization – Compute optimized** for more information).

The only workaround is to scale your EC2 instances (MediaAgents, Access Nodes) vertically or horizontally for more bandwidth.

Related documents: PERF06-BP02 Define a process to improve workload performance.

PERF06-BP03 Evolve workload performance over time

Commvault recommends evolving your Commvault data management platform architecture over time. Your number of protected workloads, sites, and competing data management activities will grow and shrink.

Periodically reevaluate your network architecture against the current business needs and look to drive further efficiency or performance for your most critical workloads.

Related documents: PERF06-BP03 Evolve workload performance over time.

Monitoring

Monitor your resources to ensure that they are performing as expected.

Monitoring your as-built environment is key to ensuring consistent reliable performance for your business. Additional details may be found in the Reliability Pillar, but focus on these key areas when designing your monitoring approach:

PERF07-BP01 Record performance-related metrics

Utilize a **monitoring and observability service** to aggregate data from all your AWS services. Commvault recommends centralizing your AWS service, Operating system, and Commvault application logs and metrics in Amazon CloudWatch.

Consider each of your users and the metrics that are important to them, then ensure they are recorded, analyzed, and anomalies automatically detected, some examples include:

- **Application owners** are interested in the accessibility and performance of the Commvault Command Center™ console, to perform self-service data management for their application. Measure the availability and responsiveness of Commvault web services to ensure adequate resource exists.
- **Application owners** are interested in the *recovery point objective* and *recovery time objective* achievable for their applications. Record backup frequency, achieved protection SLAs, and restore activity performance to measure against business SLA.

Commvault recommends aggregating business-level and infrastructure-level metrics in Amazon CloudWatch.

Business-level metrics relating to the average throughput (GB/hr.) achieved during backup and restore are printed in Commvault Logs and may be extracted from Log Files using **Creating metrics from log events using filters capability**.

Infrastructure-level metrics should track **NetworkIn, NetworkOut, NetworkPacketsIn, NetworkPacketsOut**, and **Metrics for the ENA driver** on Commvault MediaAgents and Access Nodes.

After a stabilization period, establish accepted good baselines for business-level and infrastructure-level metrics and define **CloudWatch static threshold alarms** to notify operations when a deviation of 10% from the baseline occurs.

Related documents: PERF07-BP01 Record performance-related metrics.

PERF07-BP02 Analyze metrics when events or incidents occur

Any event, anomaly, or incident that impacts your Commvault data management platform is an opportunity to learn about how your system performs. Use events to validate you are monitoring the correct metrics and have appropriate dashboards to assist in accelerating incident triage. Standard incident response *playbooks* should be enhanced to include a review of performance baseline metrics against the incident, wherever possible.

Visualize your network, application, and instance metrics in **Amazon CloudWatch Dashboards** for accelerated triage and resolution of performance incidents.

Your network architecture and design requirements should elaborate on what is considered a critical networking event, and what duration is considered acceptable for each event type.

Remember that over-delivering on your business's performance expectations represents an over-provisioned system, so ensure relevant business stakeholders are consulted on networking requirements.

Related documents: PERF07-BP02 Analyze metrics when events or incidents occur.

PERF07-BP03 Establish key performance indicators (KPIs) to measure workload performance

Identify the key AWS services and metrics to monitor to gain an accurate view of your workload performance. This will include Commvault data plane OS resource consumption and dependent services like Amazon VPC, Amazon Route53, and Amazon Identity Center (when used to provide Single Sign On). Consider network connectivity outside your VPC as well and measure link utilization for hybrid workload protection. An accurate understanding of the key KPIs for your workload will act as a canary for your service and notify you before an issue occurs.

Commvault recommends establishing key performance indicators (KPIs) and tradeoffs for each data classification.

Each data classification will have a business-agreed Recovery Point Objective RPO) and Recovery Point Objective (RTO). These objectives can be used to determine the expected throughput (GB/hr.) of backup and restore activities.

Commvault provides the **Throughput per Media Agent** report to monitor the visualize the throughput achieved per MediaAgent.

Establish a maximum workload size for agreed KPIs, because as data volume increases, so too does the required GB/hr. to meet the business-agreed RTO.

Large workloads will need to employ a recovery from Amazon snapshot, then restore to a point-in-time to limit network impact.

Gain agreement and sign-off from all business stakeholders for KPIs, and establish metrics that will track, trace, and report compliance.

Related documents: PERF07-BP03 Establish key performance indicators (KPIs) to measure workload performance.

PERF07-BP04 Use monitoring to generate alarm-based notifications

Start with defining thresholds that define acceptable and unacceptable measures for your KPIs. As you gain a deeper understanding of your workloads, refine the thresholds to reduce false positives and increase the accuracy of incident detection. Use incidents as a source for performance data and threshold breaches that affect the customer, the intent is to only be notified for events that will directly impact business service levels.

Remember to leverage intelligence features like **Amazon CloudWatch anomaly detection** to apply automated statistical and machine-learning algorithms to find issues that would be missed by manual inspection.

Commvault recommends using Amazon CloudWatch EC2 instance metrics and ENA network metrics, combined with Commvault application log metrics to generate **static threshold**-based alarms and then **anomaly-based alarms** requiring attention.

Alarms should extract log entries that include key phrases like `MB/sec, GB/hr., Bytes, Time, Average Speed, stat-`

Forward notifications into your service deck or ticketing system for traceability, ownership, and management visibility.

Related documents: PERF07-BP04 Use monitoring to generate alarm-based notifications.

PERF07-BP05 Review metrics at regular intervals

As part of incident resolution, problem resolution, and regular game days tests, review whether the current network performance metrics are sufficient or require enhancement.

If further network performance metrics collection is required to diagnose the root cause, ensure the procedures are added to relevant runbooks and/or playbooks.

Use regular maintenance events and customer incidents to review the metrics being collected. Do the metrics provide data useful in diagnosing faults? Are there additional metrics that need to be manually collected to gain a full understanding of the environment? Continually refine the metrics being collected to enrich the detail being collected and speed incident resolution.

Related documents: PERF07-BP05 Review metrics at regular intervals.

PERF07-BP06 Monitor and alarm proactively

As part of day-to-day operations, build operational dashboards that show actively running workloads, observed backup and restore throughput (GB/hr.) and the number of network packets dropped due to exceeded network baseline for the instance.

Start with manual processes to query and view relevant metrics, and build out into **CloudWatch Dashboards** to streamline incident triage and resolution.

Alarms should be forwarded to an automated ticketing system to ensure a responsible owner is assigned and drives the event to completion. The use of an automated incident management system also provides the ability to escalate the event if not resolved within an acceptable time. Alarms can be used to fire alarms before service-impacting thresholds are breached, so timely investigation and resolution are key to reliable performance for your Commvault data management platform.

Make monitoring of alarms a standard operating procedure (SOP) for your operational teams by equipping them with **Commvault Dashboards** and **Amazon CloudWatch Dashboards** that provide a snapshot of healthy and unhealthy events across protected Regions and edge locations. Consider a *single-view dashboard* that overlays backup throughput, CPU utilization, Memory utilization, and Network consumption metrics to understand how infrastructure performance affects backup time. Consider securely **sharing dashboards** with other platform teams (network, security, risk) to ensure transparent visibility and cross-functional resolution of data management risks.

Related documents: PERF07-BP06 Monitor and alarm proactively.

Trade-offs

Using trade-offs to improve performance

When architecting your Commvault data management platform, consider the following tradeoffs to meet business performance demands, while also considering the cost and complexity of supporting your solution:

- **Use various performance-related strategies**

PERF08-BP01 Understand the areas where performance is most critical

Focus your performance optimization efforts on the areas that will have the most impact on the efficiency of your data management services and your customers. Your *data classification* of business data may help focus efforts on resources dedicated to your most critical business workloads.

Where your Commvault data management platform resources are shared across all workloads and leverage Commvault **job priorities and precedence** to schedule activity based on priority, use your metrics to identify *hot spots*. Commvault data management streamed backup is heavily dependent on Amazon EC2 per-instance **network baseline** performance, ensuring you are monitoring network consumption and alarm exhaustion conditions.

Commvault recommends reviewing your MediaAgent performance, specifically **Deduplication Database (DDB) performance** and overall health by **SLA**, **Strike Count**, and **Fallen behind** replication. Additionally, your Access Nodes (**auto-scaling** or manual) are responsible for communicating with your protected workloads and will directly impact data management performance.

As part of day-to-day operations and incident or problem resolution, identify areas or instances where performance is most critical.

Consider the mixture of mission-critical, production, and non-production workloads being protected by a shared MediaAgent.

Consider segregating mission-critical workloads onto dedicated MediaAgent instances when repeated performance issues are encountered impacting critical applications.

Related documents: PERF08-BP01 Understand the areas where performance is most critical.

Consult the Commvault **Reference Architectures**, **Design Principles, and Best Practices** published in the Cloud Architecture Guide for network performance optimizations.

Commvault is largely dependent on the network bandwidth provided by the Amazon EC2 instance utilized, and the amount of concurrency permitted by the service being protected (i.e., the maximum number of **GetSnapshotBlock** requests per account, the maximum number of GetSnapshotBlock requests per snapshot).

Use **Commvault documentation** and the **Amazon Builders' Library** to understand the recommended patterns and services used in designing and operating an optimal data management platform.

Some patterns that are particularly relevant to designing highly available performant data management services include:

- **Minimizing correlated failures in distributed systems**
- **Reliability, constant work, and a good cup of coffee**
- **Building dashboards for operational visibility**
- **Implementing health checks**

Options available to impact the performance of your Commvault data management platform include:

- Upgrading your Commvault data plane Amazon EC2 instance size for improved network throughput of streamed and replicated backups.
- Utilizing Amazon EBS gp3 SSD-backed volumes to tune capacity, IOPS, and throughput for **high-IOPS deduplication database (DDB)** demand.
- Consider instance size or service performance characteristics of protected workloads (i.e., Amazon EFS performance mode impact on backup times).
- Consider which Amazon S3 storage class matches your business demand for data in a data-loss event (i.e., Does a one-zone copy of data provide sufficient durability for your backups?). Storage choices are often made on storage cost in isolation which can dramatically impact recovery time.

Remember performance also includes the performance and effectiveness of your team to detect and respond to performance incidents in your environment.

Related documents: PERF08-BP02 Learn about design patterns and services.

Identify tradeoffs and architectural and design decisions made during the original deployment or ongoing optimization of your Commvault data management platform.

For example, choosing to segregate mission-critical workloads to dedicated infrastructure with increased network bandwidth might require a **Compute Saving Plans** or **Reserved Instance** commitment to offset the cost impact of a large instance.

Use metrics to identify performance hotspots in your environment and then evaluate the available design patterns and services that can assist in mitigating or removing the hotspot. Remember that solutions to isolated events could drive increased complexity and cost of delivering your data management service. Use your *data classification* framework to understand the business importance of a proposed performance-related improvement.

Related documents: PERF08-BP03 Identify how tradeoffs impact customers and efficiency.

PERF08-BP04 Measure the impact of performance improvements

Always measure the impact of a performance improvement on the business service or SLA, the workload, and the customer experience. Existing performance metrics should be used to measure the impact and evaluate the cost-benefit of the performance improvement in Production and deployment to other affected workloads.

Ensure that any performance change is measured and tested in pre-production and then again in the production environment.

Improvements should deliver the expected performance improvement, with business-accepted cost impacts. If the improvement fails to deliver the required performance of the agreed cost profile – the change should be reverted.

Related documents: PERF08-BP04 Measure the impact of performance improvements.

PERF08-BP05 Use various performance-related strategies

Multiple proposed performance improvements should be theorized and tested in non-production environments before deciding on the final improvement. A 'slow' recovery time could be mitigated by multiple approaches including:

- Increasing the shared Commvault data plane Amazon EC2 size to gain additional network baseline bandwidth.
- Migrating the workload protection approach to use an Amazon EC2 IntelliSnap® snapshot as the first rapid recovery point.
- Enabling Fast Snapshot Restore (FSR) on critical snapshots for rapid snapshot recovery at a per-hour snapshot cost increase.
- Reducing the volume of data protected and restored by excluding cache or data that can be easily recreated.

Promote a 'no bad ideas' approach to performance improvement within your operational team. Identify options, test and assess with performance metrics, and then implement the approach that meets performance and business needs.

Commvault recommends employing multiple strategies to optimize network performance. Examples being:

- Locating a data copy onsite to avoid large restores over Direct Connect or VPN connections.
- Segregating backup and restore MediaAgent infrastructure to better isolate workloads that are impacting network performance for multiple users or workloads.
- Segregating and upgrading instance size for high-value workloads to deliver improved network baseline and subsequently achieved throughput.
- Removing easily re-creatable data from the backup scope.
- See **Commvault Risk Analysis**,

Related documents: PERF08-BP05 Use various performance-related strategies.

Additional Resources

- **Performance Efficiency Pillar: Well-Architected**
- **The Amazon Builders' Library**
- **AWS re:Invent 2021 - {New Launch} Deep dive into AWS Graviton3 and Amazon EC2 C7g instances**
- **AWS re:Invent 2021 - Optimizing resource efficiency with AWS Compute Optimizer**
- **AWS re:Invent 2021 - Advanced Amazon VPC design and new capabilities**
- **AWS re:Invent 2021- Get insights from operational metrics at scale with CloudWatch Metrics Insights**
- **Commvault Store – Performance, Trending & Prediction Reports**
- **Commvault Store – Alerts & Notifications**

Cost Optimization Pillar

Delivering holistic data management services across your AWS Regions and edge locations includes initial and continual cost optimization of used services. A cost-optimized Commvault platform will achieve business *recovery point objectives* and *recovery time objectives* with the least possible price point.

Practice Cloud Financial Management

Practicing mature Cloud Financial Management (CFM) means gaining an understanding of your business data protection needs and your financial objectives for mitigating data-loss risk. Understanding business needs allows continual monitoring of the business value of consumed AWS services, and the ability to cost-optimize when misalignment is identified.

COST01-BP01 Establish a cost optimization function

Commvault recommends establishing a cost optimization function within your business with finance, technology, and business stakeholder inclusion.

Develop KPIs for the team that focuses on cost management and cost optimization across all cloud spending. Ensure that an executive sponsor exists to assist in the resolution of conflict or misalignment between internal policy and standard.

For example, legal counsel may dictate that all backup data is retained indefinitely, and the cost awareness function assists in educating on data classification, and industry compliance to optimize data retention while meeting regulations.

Commvault publishes a list of **Cost Optimization Features** that should be monitored and assessed for implementation.

Ensure that an individual or team is established with the accountability to drive organizational cost awareness, cost management, and cost optimization for both business applications and your Commvault data management platform. This team should be measured on meeting business cost objectives for fixed monthly costs or specific workload efficiency metrics (i.e., cost per protected workload).

The team or individual can be accountable for directly driving optimizations where a centralized platform approach is used to provision and manage workloads. Alternatively, in large, distributed organizations, establishing cost-optimized technology standards, cost transparency reporting, and per-business unit alerts can be used.

Regardless of whether this function is provided by the team responsible for designing and operating your Commvault data management platform, the business's financial objectives and measures must be made available to influence architectural and design decisions.

Related documents: COST01-BP01 Establish a cost optimization function.

COST01-BP02 Establish a partnership between finance and technology

Establish a regular cadence for finance and technology teams to review the state of cloud and usage, including any items exceeding expected spend.

Provide an avenue for costs to be tracked and understood, including baseline monthly costs and active projects or initiatives.

The technology team has the responsibility to explain pricing models, discounting options, and cost optimization practices are undertaken (i.e., using **power management**, **auto-scaling**, and **deduplication savings**)

Technology and Finance teams work closely together to build operational cost and usage dashboards and trending reports to drive discussion.

Engage **Commvault Technology Consulting** or **Commvault Enterprise Support** your representative for guidance on optimizing your data protection and data management practices in AWS and beyond.

Create cross-functional teams that include financial and technology leaders to establish guardrails for accelerated cloud-fueled innovation, while respecting business financial objectives. Ensure that your *risk management function* is included in these virtual teams to establish any business, industry vertical, or compliance demands on data protection and data retention.

Commvault data management is most deployed in a *shared services* model where the infrastructure costs and ongoing management are borne by the centralized platform team. Individual business unit owners and even third-party operators should be provided guardrails on AWS service adoption to drive an acceptable-use approach to new services and technology.

Establishing well-communicated guardrails and policies for consumption allows a standard operating procedure (SOP) when a business unit has unique requirements (i.e., an exception-handling process). Should a single business unit or application impose a significant cost increase on cloud spending, it can be assessed by the virtual team for a cost-optimized solution or additional funding.

Commvault can segregate specialized workloads onto dedicated Amazon S3 storage and utilize dedicated Amazon EC2 infrastructure for unique workload needs. This approach allows greater visibility of the cost impact of the single application or business unit, which facilitates a business value analysis.

Related documents: COST01-BP02 Establish a partnership between finance and technology.

COST01-BP03 Establish cloud budgets and forecasts

Budgeting and forecasting processes should be adjusted to account for a business-accepted variable consumption threshold.

Pay particular attention to storage, compute, and network transfer costs which grow as increased data protection and disaster recovery solutions are activated.

Use **AWS Cost and Usage** reporting (current, historical, forecast) and internal business reporting to build a hybrid forecast of future consumption.

Commvault recommends building a workload data protection calculator that reflects the typical costs generated to protect a workload at each tier of your data classification plan. New initiatives and business plan approval should include workload protection costs.

Commvault provides the **Storage Cost Optimization** report to help model changes to **storage class** on storage spend.

Business budgets and planning processes need to account for the variability in new workload deployment and the impact on Commvault data management services. Consider a *data management tax* on workloads that budgets a baseline reservation for protecting a business workload to the *data classification* needs of the business.

Too often data protection needs are not assessed until the *Operational Readiness Checks (ORC)* are performed to transition into production. Ensure budgets and forecasts are integrated with a nearly real-time view of provisioned workloads and use **Cost Explorer** to forecast future costs from active services. As the number of workloads and

Regions and edge locations grows, so too does the cost to protect those workloads. Depending on your industry and business policy, data retention costs may be significant and require a cost-optimization exercise to control.

Related documents: COST01-BP03 Establish cloud budgets and forecasts.

COST01-BP04 Implement cost awareness in your organizational processes

Embed cost awareness into existing processes and develop new processes where cost awareness insight is required.

Any process that will create or delete a resource should be enhanced with **cost awareness** and optimization checks (e.g., **automated power down of resources nightly**).

For example, the creation of new compute resources by line-of-business developers should include recommended instance types, cost benefits, and sustainability benefits to aid in cost-optimized selection.

Additionally, incident and problem management processes should be augmented to ensure that any resources created are powered down when not required, and then terminated.

Commvault recommends that the termination of resources include a final long-term retention, and low-cost archival backup if the associated resource data cannot be easily recreated.

Cost awareness needs to become a key skill set for your technical leaders and engineers, not just your financial leaders. Reuse and extend your existing processes that measure and report on cost consumption. Consider enhancing existing practices with the following practices to better align with the velocity of cloud consumption:

- Ensure that change management includes a cost consideration and consider including a financial representative in *change approval* processes.
- Measure and alarm cost increases or anomalies in the same way performance is managed.
- Build teams with cost-optimization skills and experience via training, certification, and continual improvement initiatives.

Related documents: COST01-BP04 Implement cost awareness in your organizational processes.

COST01-BP05 Report and notify on cost optimization

Commvault recommends using **AWS Budgets** notifications on all AWS accounts, to notify responsible finance, technology, and line-of-business owners of low and high watermark consumption.

Consider sending notifications directly to the service desk or incident management systems if a cost mitigation action is required.

Use **AWS Cost Explorer** dashboards and reports to identify both optimized and uncontrolled costs. Promote learnings across technical and line-of-business teams.

Compare optimized and uncontrolled workloads to understand key architectural decisions that led to optimal cost consumption (i.e., excluding **recreate-able content**, **selective replication** vs. replicate-everything, **auto-scaled** vs. static resource allocation).

Use **AWS Cost Explorer** and **AWS Budgets Reports** to proactively report cost and usage. Commvault data management will drive compute and/or container, networking, and storage costs as protected workloads grow. Create **Budget alerts** to allow proactive review of costs before an Amazon invoice arrives.

Related documents: COST01-BP05 Report and notify on cost optimization.

COST01-BP06 Monitor cost proactively

Commvault recommends frequent monitoring of cost vs. waiting for AWS Budget **notifications**. Use AWS Cost Explorer **Dashboards, Trends, and Reports**, and publish relevant extracts throughout your office and intranet to inform users on cloud consumption.

As shared services or data management teams learn more about optimizing resource selection and configuration, share learnings with the relevant cost-saving metrics to application owners.

Cost and usage reporting should be made widely available to technical resources that are building workloads for your business. Transparency of consumption drives awareness and a continuous improvement mentality owned by all layers of the business.

Showcase application owners that have implemented cost-optimized solutions and create *technical architecture forums* where all responsible parties can learn how to cost-optimize their solutions (including Commvault data management).

Related documents: COST01-BP06 Monitor cost proactively.

COST01-BP07 Keep up-to-date with new service releases

Ensure you are subscribed to **AWS Blog** (AWS Cloud Financial Manager Category) to be updated on new advancements for managing your cloud costs.

Review new versions of the Cloud Architecture Guide (this document) and the **Newsletter for New Features in Commvault Platform Release 2022E** for details on new products and features relevant to cost optimization.

Actively participate in **Commvault Community** and **AWS re:Post** to share ideas and best practices across Commvault and AWS users.

Related documents: COST01-BP07 Keep up-to-date with new service releases.

COST01-BP08 Create a cost aware culture.

Commvault recommends that part of any Commvault Backup & Recovery Operational Readiness Review (ORR) include the establishment of cost transparency reporting and ongoing awareness and optimization KPIs. Commvault platform owners should have access to view **AWS Cost & Usage Reports**, and optionally **AWS Cost Explorer**. Commvault provides granular storage usage and trending, and deduplication and compression storage savings to compliment AWS reporting.

Additionally, change control processes should include a cost element that establishes if a change will incur additional costs and seek approval from relevant business stakeholders if exceeding a pre-agreed threshold. Commvault platform owners should continually update themselves via **AWS Skill Builder**, **AWS Solutions-focus Immersion Days**, and **AWS What's New** blog to identify opportunities to reduce costs over time. Solutions that are identified to reduce cost should be celebrated and documented for re-use across the organization. Right-sizing resources using tools like **AWS Compute Optimizer** should be part of regular maintenance events (i.e., monthly, quarterly).

AWS Budgets should be established before moving into production, with automated alarms into platform owner ticketing systems for immediate attention. Uncontrolled or runaway costs should be curbed by automating actions (i.e., shutting down high-cost instances automatically to allow for investigation and potential budget re-baseline).

Related documents: COST01-BP08 Create a cost-aware culture.

Commvault recommends that financial reporting include the benefits of cost optimization.

Use the **Storage Cost Optimization** report, and **AWS Calculator** to model the impact of storage tiering optimization. Model the before and after effects of Amazon EC2 runtime costs when enabling **power management**.

Commvault recommends using **Compute Savings Plans** and **Reserved Instances** for shared Commvault components like the CommServe and MediaAgents, model and promote the savings impacts using AWS Calculator.

Consider savings offered by adopting more cloud-native approaches, for example implementing hands-free operations through **automation**, and investing in employee **education** in cloud services and cost-optimal solutions.

Understand the business value or impact of cost optimization. *Cost avoidance* or *cost reduction* of **powering off a resource when not in use**, or **terminating a resource** when no longer used are well understood and measurable. Measure the impact of cost optimization on business outcomes and consider improved *recovery time objectives* or *reduced recovery time* which directly impact the time to recover a business service from an unplanned outage.

Cost optimization also includes operational efficiencies, for example:

- Re-assessing application *data classification* to relax required data protection resources.
- Consolidating backup and replication windows to allow powering off shared resources earlier.
- Automating previous manual processes and freeing up resources for more valuable tasks.

A common operational efficiency achieved by implementing Commvault Backup & Recovery is automated snapshot lifecycle management for workloads. This approach replaces manual 'scripted' solutions that can often result in orphaned or invisible data copies that drive monthly cost that is misaligned with workload business value.

Related documents: COST01-BP09 Quantify business value from cost optimization.

Expenditure and usage awareness

Seek to understand your organization's costs both at a macro-level and down to an individual business unit and workload level or micro-level. Ensure that workloads are onboarded to Commvault data management with a *data classification* associated with the workload. *Data classification* is an indication of the criticality and sensitivity of the workload and can be used to review cost drivers and opportunities for cost optimization.

Governance

Consider implementing cost governance and control functions in the following areas:

Commvault recommends developing cloud usage policies and publishing and reporting usage compliance per business unit or application (workload). Initially, this will be a discovery process on required locations, services, and features.

Policies should guide which regions, compute instances, and storage types should be used based on the business data classification for a workload (e.g., mission-critical, production, dev-test).

Commvault helps keep compliance simple by directing recoveries to **approved regions** and using **Plans** that perform data protection to approved storage classes based on AWS Resource tags applied to resources. You can be

confident that your compliance controls stay in place, as Commvault protects and recovers AWS tags as part of AWS workload backup & recovery.

Individual users and user groups can be permitted access to specific regions based on **role-based access controls**.

Develop and publish organizational policies for cloud usage, mandatory data protection, and retention. Policies can start high-level with supported Regions, preferred instance types and sizes, and acceptable hours of operation.

Policies can mature over time to include resource creation, modification, and decommissioning processes. Processes should include *data management considerations* to ensure that any change captures the workload and retains the backup or archive for the business-required period. Consult your business Governance Risk and Compliance (GRC) function, and Legal department for guidance on how long data should be retained.

Related documents: COST02-BP01 Develop policies based on your organization requirements.

COST02-BP02 Implement goals and targets

Develop cost and usage goals that are delivered to line-of-business leaders and individual workload owners.

Consider broad measurable objectives with agreed time windows, allowing time for experimentation and optimization (i.e., reduce cloud spending by 5% for every 20% increase in cloud consumption, measured quarterly).

Commvault recommends establishing data protection resourcing and cost estimates that indicate baseline costs and expected increases for adding DR replicas, improving restore performance, and adding additional retention.

Goals should include the ability to establish cost targets per application, based on business value. A new capability aimed to grow new revenue may have a lower cost optimization target vs. a dev-test experimentation initiative.

Data protection cost and usage goal setting must include legal counsel and governance, risk & compliance representative to ensure goals are realistic based on compliance requirements.

Set cost and usage goals for all cloud consumers within the organization. Additionally, publish cost and usage targets that allow business units and leaders to measure their consumption to business objectives. Targets are not fixed and often require continual optimization over time. For example, a business target may be *a 20% increase in protected workloads results in a less than 5% increase in data management costs.* By leveraging centralized, shared data management infrastructure, Commvault can deliver cost optimization across all business units and Regions.

Related documents: COST02-BP02 Implement goals and targets.

COST02-BP03 Implement an account structure

Establish an account structure with a single managed account (parent) and multiple member (child) accounts. Ensure all AWS consumption is included in this account structure and use **AWS Organizations consolidated billing** to manage and visualize spending.

Ensure workloads are only deployed to the member account. Commvault recommends that Commvault resources are isolated in a dedicated AWS account per environment type (production, DR, development, test). Establish an account separation policy for your business considering risk containment, the sensitivity of data, cost transparency, and required per-workload granularity.

Centralizing billing provides visibility and allows maximizing of discounting and even cost-saving analysis (i.e., applying Compute Savings Plans for stable resource consumption).

Ensure that a single-parent account (known as the *management account*) is maintained as the *payer account* with all business units and workloads being contained within *linked accounts*. Utilize **AWS Organizations** and **consolidated billing** to ensure that all cloud consumption is visible. Commvault

performs **cross-account protection**, allowing data protection costs to be consolidated within a single shared-services account.

Related documents: COST02-BP03 Implement an account structure.

COST02-BP04 Implement groups and roles

Create logical groups of users with associated **AWS IAM roles** that permit or enforce the creation, modification, and decommission of approved services and resources per organizational policy.

Commvault software uses these roles or machine identities to perform backup, recovery, and replication of AWS snapshots between accounts and regions.

Consider using Commvault as a control or enforcement point that can restrict the **users and groups** that can create new instances, referred to as an **out-of-place restore**.

Define AWS Organizations' **organizational units (OUs)** to allocate permissions (i.e., AWS IAM Roles) that perform common functions like service provisioning, workload management, or financial reporting. Commvault provides prescriptive **AWS IAM Policies** for protecting specific AWS services, allowing granular allocation of permission to appropriate groups.

Related documents: COST02-BP04 Implement groups and roles.

COST02-BP05 Implement cost controls

Ensure that **AWS Budgets** notifications to resource owners on spend exceeding agreed thresholds. Implement a low watermark and high watermark notification, which allows prompt action by the resource owner.

Utilize AWS Identity & Access Management (**AWS IAM**). AWS Organizations Service Control Policies (**SCPs**) to enforce a policy restricting creation in approved **regions**, with approved instance and storage types.

Consider **Increase in data size** predefined alerts to resource owners when significant data growth occurs, which may lead to unanticipated costs.

Define monthly budgets in **AWS Budgets** and be automatically alerted when budget thresholds are exceeded. Alerts can be defined at the organizational level, or down to individual accounts, services, or tags. Budget alerts can be sent to an **Amazon SNS topic** for a more sophisticated, centralized notification approach.

Commvault applies **AWS Resource** tags to all resources that are created as part of a backup, restore, or replication activity. The following tags may be used to identify Commvault-created resources:

Tag key	Example Tag value	Commvault created resources
_GX_BACKUP_	Amazon EC2 instance-id for AMIs *Blank for other resources*	Amazon Machine Images (AMIs), Amazon EBS snapshots, Amazon EBS volumes, Amazon RDS snapshots, Amazon DocumentDB snapshots
CV_Retain_Snap		Amazon EBS snapshots (Commvault Auto Recovery – Periodic Replication)
CV_Integrity_Snap	AMI ID	Amazon EBS snapshots (Commvault Auto Recovery – Periodic Replication)
_GX_AMI_	AMI ID	Amazon Machine Images (AMIs)

Tag key	Example Tag value	Commvault created resources
Description	Snapshot_created_by_Commvault_for_job_n_at_nnnnnnnnn._Source_Volume_vol-*VolumeId*_from_*CommServeName*	Amazon EBS snapshots
Description	Snapshot_created_by_Commvault_for_job_nnnn_Source_RDS_instance_(instance-name)	Amazon RDS snapshots
Description	Snapshot_created_by_Commvault_for_job_nnnn_Source_DocumentDB_(instance-name)	Amazon RDS snapshots
Description	Volume_created_by_Commvault_for_job_6_at_1661277681_from_cs32	Amazon EBS volumes, during Amazon EC2 Full instance restore
Name	CV_CBT_Snap	Amazon EBS snapshots
Name	SP_x_yyy_zzz	Amazon DocumentDB snapshots
Name	SP_x_y_z_nnnnn	Amazon EBS snapshots
Name	_GX_BACKUP_SnapJob_5_vol-07a265a4227709c00	Amazon EBS snapshot in use for a Full Instance restore
Name	(new-instance-name)	Amazon EC2 Full Instance restores
CV_Subclient	*Subclient ID*	Amazon Machine Images (AMIs)
AMI name	*AMI_FOR_SP_x_y_z_nnnnn*	Amazon Machine Images (AMIs) *x = commcell id*
AMI Source	*nnnnnnn/AMI_FOR_SP_x_y_z_nnnnn*	Amazon Machine Images (AMIs)
Volume name	*_GX_BACKUP_SnapJob_5_vol-07a265a4227709c00*	During restore of a full instance

ⓘ **Note**: Commvault does not tag snapshots created for Amazon Redshift or Amazon DynamoDB.

It should be noted that Commvault discovers AWS Resource tags during AWS workload protection, protects them, and recovers them as part of workload recovery. This means your existing tag-based management processes and governance controls automatically apply to any resources created by Commvault.

Controls – Enforcement

A robust Identify and Access Management process should be defined to ensure that users and user groups have only the privileges they require to perform their duties (i.e., **least privilege**). These privileges should extend to a user or user group's **role** within Commvault Backup & Recovery. Commvault has a rich role-based access control (RBAC) system that allows granular assignment of permissions to individual users and workloads.

Use AWS Organizations **Service Control Policies (SCPs)** to further restrict what operations can be performed

by users, but take care to ensure that your Commvault data management system is given the Commvault required **user permissions** per data management **use-case**.

> ⚠ Important
>
> Service Control Policies will <u>override</u> permissions granted at the individual IAM user or role level. Ensure that SCP controls include Commvault permissions and test all SCP changes do not adversely affect your Commvault data protection services. For example, ensure that Commvault-created AWS Resource tags are <u>permitted by AWS Organizations Service Control Policies</u>

Controls – Service Quotas

Service quotas can be used to effectively limit the over-provisioning of resources resulting in a bill shock event. Be aware that some service quotas will directly impact how quickly Commvault can perform backup, recovery, and replication activity (i.e., **EBS Direct API Backups for Amazon Web Services**). Controls should prevent increases in service quotas by unauthorized users or organizational units (OUs), Commvault administrators must be permitted to increase quotas where they limit data protection *recovery objectives*.

Related documents: COST02-BP05 Implement cost controls.

COST02-BP06 Track project lifecycle

Ensure that every workload has a defined owner and data classification attached via **AWS Resource Tags**. Commvault will protect and recover these tags for supported workloads.

Audit workloads by business unit, workload, and data classification through their lifecycle from development, to production, and then decommissioning.

Commvault **Plans** adjust the storage location and storage costs associated with a workload as its business value changes.

Commvault recommends that as projects and resources are decommissioned a final <u>long-term retention/archive</u> backup is taken to a **Combined Storage Tier** location to allow simplified recovery before terminating resources.

Implement a process for tracking and reflecting the status of a workload (e.g., active, inactive, decommissioned, archived). Use Amazon resource tags to reflect a workload business value or *data classification*, as Commvault uses **resource tags** to discover and protect workloads per business policy. You can be confident workload status will be maintained on all workload copies, as Commvault protects and recovers tags for supported AWS resources.

Commvault can take a 'last backup' for AWS services as a streaming backup (e.g., **Amazon EC2 streaming backup**) that does not rely on service snapshots, and can be used to delete (terminate) the workload and all further service costs (excluding the Amazon S3 storage to retain the final backup).

See below for a high-level summary of the AWS services Commvault can protect with snapshots and streamed service-independent backup copies. See Commvault protection of AWS Cloud Products for full coverage.

AWS Service	Snapshot-backup	Streaming-backup
Amazon Aurora (including v1, v2 serverless)*	●	●
Amazon DocumentDB	●	

Amazon DynamoDB		●
Amazon EC2 and Amazon EBS* **	●	●
Amazon EFS (NFS)		●
Amazon EKS and Amazon EBS*	●	●
Amazon FSx for Windows (SMB)		●
Amazon FSx for NetApp ONTAP (SMB, NFS)		●
Amazon RDS*	●	●
Amazon Redshift	●	
Amazon S3*		●
Amazon Workspaces		●

* Includes workload on AWS Outposts

** Snapshots must be stored in AWS Region due to API limitation on AWS Outposts

Related documents: COST02-BP06 Track project lifecycle

Monitor Cost and Usage

To promote a culture of cost awareness and continual *cost optimization*, teams must be equipped with cost transparency. Consider the following areas for providing detailed reporting that allows teams to drill down and understand where cost is being generated:

COST03-BP02 Add organization information to cost and usage

Commvault recommends hourly granularity be enabled in AWS Cost Explorer to allow a more granular understanding of cost peaks (i.e., auto-scaled resources during backup windows). Ensure AWS Resource Tags are attached with granular organization information to permit per-organization team/unit reporting. Alternatively in AWS Organizations-based environments, Organizational Units (OU) may be used to report and review costs via org. units.

Ensure your **AWS Cost and Usage Report** includes resource ids, and automatic refresh with hourly granularity (this will additional cost, review AWS Cost Explorer pricing).

Ensure that Amazon CloudWatch Log ingestion occurs with at least hourly granularity to allow traceability between usage reports and Commvault software activities.

Commvault recommends dedicated Cost and Usage Reports (CURs) dedicated to the shared Commvault data management platform compute, network, and storage resources.

Use **hourly granularity** in AWS Cost Explorer to understand hourly costs down to a resource level. Hourly granularity is particularly important with Commvault data management as backup and restore activity can occur ad hoc and drive unplanned or unforeseen costs. Use resources, accounts, and *locations* to isolate the data management activities that are driving cost.

Related documents: COST03-BP01 Configure detailed information sources.

COST03-BP02 Identify cost attribution categories.

Finance and technology teams should work together to determine how the cost will be attributed within the business for **chargeback** or **showback**.

Commvault **Chargeback Details** report can be used to apportion cost across business units (client groups) and workloads (clients).

Shared organizational costs like online and in-person education and hackathon initiatives should also be considered.

Commvault recommends <u>multiple categorizations</u> are implemented based on financial, technical, and data management needs.

Consider recording business-unit, environment types, and data classification for the workload as categories used to report cost. Commvault can reflect these categories using **entity tags**, and use them in **reports** and **custom dashboards**.

Ensure that all cloud costs are appropriately assigned to the consuming department, user, or shared function (i.e., data protection service). Ensure all costs are considered, including workload resources, education & training, and ad hoc or on-demand support services provided by Amazon or another third party.

Determine how *data management costs* will be attributed to each business unit or application. Commvault recommends consolidating backups into global deduplication pools that reduce the cost of stored and replicated backup data. One approach to attributing data management infrastructure and storage cost is by the size of protected data as measured by the client (before compression or deduplication). Commvault reports client data size as Application Size in the **Backup Job Summary Report** and **Restore Job Summary Reports**. Another cost attribution approach is considering the total number of data management activities that were performed for the workload, each activity (snapshot backup, streaming backup, replication) has an associated infrastructure and service cost.

Related documents: COST03-BP02 Add organization information to cost and usage.

COST03-BP03 Identify cost attribution categories

Commvault recommends defining data protection metrics to reflect the business value of costs associated with backup and recovery.

Backup and recovery SLA compliance should be measured using the <u>recovery readiness report</u>, which details the configured and achieved RPO, and the configured and achievable RTO against business policy.

Some examples include the **SLA Health** for an application or environment (met/missed workloads), the actual RPO (RPA), and the actual RTO (RTA) achieved during restore activities (See **Recovery Readiness Report**).

Data management typically has three (3) primary costs to consider which roll up the infrastructure, networking, and underlying storage to provide the service:

- **$ per protected TB** which is an indicator of the cost of providing data protection for the workload at its given *data classification*. A business-critical application will have snapshots + Amazon S3 backup copies typically in multiple regions, whereas a development workload, may only have an Amazon S3 copy.
- **$ per recovery minute** which is an indicator of the cost of *recovering an application*. There may be multiple measures here dependent on whether the recovery is using a snapshot or occurring in an alternate Region with differing costs to Production location.
- **$ per replicated TB** which is an indicator of the cost of providing a replicated disaster recovery copy of the the workload in an alternate Region for Disaster Recovery.

Related documents: COST03-BP03 Identify cost attribution categories.

COST03-BP04 Establish organization metrics

Ensure that every team capable of creating AWS resources has business-relevant reporting, notifications, and trending for their workloads. Establish best practices for acceptable consumption and growth, allowing org. units to self-manage.

Promote and educate teams on **AWS Cost Explorer**, **AWS Budgets**, and optionally cost analysis using **Amazon Athena** and **Amazon QuickSight**.

Using AWS Organizations and AWS Control Tower to place users in a cost optimization group that grants access to the required Cost and Usage reports.

Ensure that **AWS Resource Tags** are applied to all resources to allow traceability of owning team, purpose, and status (production, non-production, test). Consider configuring **Service Control Policies** that ensure that all newly created resources must have a minimum set of AWS tags to facilitate billing and cost management.

Ensure that each user or user group that has permission to create billable resources has the access to consume the following **AWS Billing and Cost Management** functions to inform and action cost optimization initiatives:

- Real-time, historical, and forecast cost and usage *reports.*
- Ability to define and receive *notifications* when costs exceed agreed business thresholds.
- Actual use *tracking* and *analysis* against business cost and usage targets.

> ☀ Pro-Tip
>
> Be aware then when shared storage is used by Commvault to provide organization-wide backup and archive services, per workload storage reporting will not be possible using AWS Billing and Cost Management tools. If granular per workload costs are required, consider creating dedicated **Cloud Storage Pools** per workload or line-of-business, this will reduce overall savings achieved by Commvault Global Deduplication.

Related documents: COST03-BP04 Establish organization metrics.

COST03-BP05 Configure billing and cost management tools

Ensure all resources created during development, testing, or during recovery events, are **tagged**. Develop a tagging strategy that includes at a minimum data classification, environment type, financial owner, technical owner, and owning business unit. Consult **AWS tagging best practices** for additional guidance. Enable **AWS Budgets** to alarm usage approaching organizational thresholds. Enable **cost allocation tags** to allow per-team granular cost management.

Commvault automates tagging by **restoring tags** as part of AWS resource restores to new regions or AWS accounts. You can also customize tags during the recovery of some resources, for example adding, deleting, or modifying tags during an Amazon EC2 instance recovery.

Commvault enforces consistency between tags and resource utilization by directing backups to business-approved storage technology based on tags applied to resources.

Commvault implements **tagging** of created resources to allow visibility to resources created by Commvault. Additionally, apply business context tags that show per-workload *data classification* to indicate data value, criticality, and therefore 'application/ service value' to the business. All resources should be tagged as *Production* or *Non-Production* to ensure appropriate protection policy is applied, but appropriate data handling occurs.

Use **Tag policies** to ensure that mandatory tags are applied to all new workloads created by the business.

While **AWS Cost Categories** can assign costs based on AWS accounts, services, charge types, and cost categories, Commvault recommends implementing tag policies to reflect workload importance and data classification as a minimum

Related documents: COST03-BP05 Configure billing and cost management tools.

COST03-BP06 Allocate costs based on workload metrics

Consider creating hybrid aggregated business metrics that reflect business value for a cost. Consider measuring the effectiveness of your Commvault data protection platform with metrics like:

- **$ per protected TB** as an indicator of the cost of providing data protection for the workload at its given data classification.
- **$ per recovery minute** as an indicator of the cost of recovering an application.
- **$ per replicated TB** as an indicator of the cost of providing on-offsite disaster recovery copy for a workload.

Consider engaging **Commvault Technical Consulting** to assist in building relevant business dashboards and/or reports.

Combine static and variable costs along with data management service metrics (cost per protected TB, cost per recovery minute) to gain a business outcome focused on costs. Focus on the least-efficient cost drivers. For example a restore to a Region without an associated VPC Peering relationship might drive an increased egress to Internet cost that can be optimized through network re-design.

Commvault recommends allocating streaming backup costs based on the front-end terabytes (FETB) or *storage as measured at the client before deduplication and compression*. An application with more data will incur more impact on shared MediaAgent infrastructure in deduplication lookups, and network impact to backup or recover data.

Snapshot-only backup costs can be identified by the Amazon EBS snapshots created and held during the measurement period.

Related documents: COST03-BP06 Allocate costs based on workload metrics.

Decommission Resources

When resources are no longer active, remove or archive them.

COST04-BP01 Track resources over their lifetime

Ensure that all resources are tagged with the owning workload or application name.

Additionally, ensure that the workload status (e.g., active, inactive, testing, decommissioned) is reflected in AWS resource tags. Some resources may allow daily power-down or even termination.

Use workload throughput (**network**) or load (**CPU**) monitoring to identify workloads that are no longer active and can be decommissioned.

Using tags to reflect workload data classification allows data protection for a workload to stop or redirect to lower-cost resources when a workload reaches an inactive state. Commvault will restore tags that were collected during backup to any restored resources, ensuring resource right-sizing or automation will apply consistently across your AWS resources.

Ensure that operational processes both manual and automated (via continuous integration/continuous deployment pipelines) tag and re-tag resources as their operational status changes. Commvault will use these tags to apply appropriate backup and data retention policies. This means that re-tagging a resource will apply the most appropriate protection policy from the next backup, effectively cost-optimizing protection. Resources with inactive or 'archival' tags can be used to automate 'last backup' and resource termination.

Related documents: COST04-BP01 Track resources over their lifetime.

COST04-BP02 Implement a decommissioning process

Ensure a process exists for identifying unused resources, tagging, or flagging them for termination, and what actions should be taken to decommission the resource. The process must include the business unit responsible for each step of the process to ensure decommissioning can be actioned promptly to realize cost optimizations.

Commvault recommends that when a resource is tagged as inactive or end-of-life that a decommissioning process is executed to avoid future spending.

Decommissioned processes may differ based on the value and sensitivity of the data stored with the workload. Use **Commvault® Risk Analysis** to identify sensitive data in the workload and tag resources for appropriate handling.

Decommissioning processes may differ on data type as well. All archive backups should be streaming backups independent of the originating cloud service.

Ensure those database backups are taken as a database-native dump or **export-based backups** (Aurora, MariaDB, SQL Server, MySQL, Oracle, PostgreSQL).

Related documents: COST04-BP02 Implement a decommissioning process.

COST04-BP03 Decommission resources

Proactively decommission resources using a runbook that drives a one-time archive of inactive resources to the Commvault **Combined Storage Tier**, and then terminates all related resources.

Use data classification tags to control the cost and number of copies taken before decommissioning, Commvault can enforce these controls through dedicated decommissions **Plans**. Additionally, if any previously-archived resources are restored, their data classification status will be maintained and reflected via the applied data classification tags.

Commvault assists in decommissioning by providing long-term retention or archival backup before resource termination. Commvault can write these backups to Amazon S3 Glacier Flexible Retrieval and Amazon S3 Glacier Deep Archive, but allows **simplified automated recall** if required. After the final backup is completed, the business unit or application owner can terminate the resources from their AWS account. Additionally, Commvault now supports **S3 Batch Operations** for increased throughput restores from S3 Glacier, meaning you can extract more value out of your long-term retention archives (see faster restores for S3 Glacier).

Commvault® Risk Analysis provides detailed analytics to identify duplicate, unused, or insecure unstructured data that can be archived. Uncontrolled unstructured or 'dark' data can often account for 80%+ of data in an organization. Use Storage Optimization to identify inactive data, and **archive it** by performing a backup and then deleting it from the source file server or storage service.

Commvault Risk Analysis also identifies critical or sensitive data that requires additional governance to ensure your most sensitive data is kept safe. **Commvault® Risk Analysis** (Data Governance) can **archive sensitive files** and delete them from the source file server or storage service.

Commvault Risk Analysis assists in decommissioning data assets from Amazon EFS, and Amazon FSx for Windows, and traditional file servers running on Amazon EC2.

Related documents: COST04-BP03 Decommission resources.

COST04-BP04 Decommission resources automatically.

Enhance decommissioning processes with automation to identify resources by tag, backup, then terminate using **AWS Systems Manager Run Command** automation. This process can be combined with **Business Logic Workflows** to require active approval before decommissioning begins.

Consider automated decommissioning for resources that demonstrate a consistently low network, and CPU utilization (see **Create Alarms to Stop, Terminate, Reboot, or Recover and Instance**).

Commvault software automates cost optimization of its resources and **automatically decommissions auto-scaled Amazon EC2 resources** that are no longer required to perform backup activities. When the next backup runs, Commvault auto-scales the required resources to protect the workloads based on data volume and *recovery point objective*.

Commvault Data Intelligence implements automated archival (data backup then deletion) along with the protection of active business approval (see **Request Manager** for more information).

Related documents: COST04-BP04 Decommission resources automatically.

COST04-BP05 Enforce data retention policies.

Commvault Server Plans implement your business data retention policies for all backup and archival copies (including AWS native snapshots, where supported). Commvault will automatically manage the creation, copying, replication, and deletion of data copies per configured plans.

Additionally, Commvault tiers data between higher-cost snapshot copies designed for rapid recovery, to Amazon S3 frequently accessed and infrequently accessed **storage classes** (including S3 Glacier). Commvault recommends the use of Commvault **storage copies** instead of Amazon Data Lifecycle Manager to ensure optimal recall times for Commvault format backup data. Commvault provides a **Combined storage tier** that distributes Commvault data across frequent-access and infrequent-access storage classes to optimize storage costs and simplify recalls.

Commvault does not recommend or require the creation of lifecycle configurations on Amazon S3 buckets, Commvault will manage the entire lifecycle of any created objects from creation through to eventual expiry (i.e., deletion).

Related documents: COST04-BP05 Enforce data retention policies.

Cost-effective resources

Cost optimization of your Commvault data management platform starts with selecting the most appropriate resources to meet business data protection SLAs with the least cost. Consider the following when selecting AWS resources:

Evaluate Cost When Selecting Services

Make cost one of the key criteria when evaluating and adopting AWS services. This includes

COST05-BP01 Identify organization requirements for cost

Commvault publishes **sizing guidelines** for the least cost and best price-performance instances for performing data protection of AWS workloads.

Work with business owners to understand the strategic priority for a workload to determine if it should be optimized for cost or performance, depending on its function.

Commvault recommends starting simple and using your data classification framework to identify whether workloads should be cost-optimal or performance-optimal.

Use AWS Resource Tags to identify the current preference for the workload.

Selecting resources for your application workloads and your Commvault data management platform is a a balance between performance, reliability, and cost. Use *data classification* of protected workloads to inform and prioritize the selection of the most appropriate resources for your data management system. In rare cases, resources may be selected and isolated for use by business-critical or mission-critical applications.

Related documents: COST05-BP01 Identify organization requirements for cost.

COST05-BP02 Analyze all components of this workload

Commvault recommends using hotspot analysis for cost optimization. Use AWS Cost and Usage Reports (CURs) to understand the workloads that are driving the most cost. Then, for each workload break down the underlying components (compute, network, storage) and analyze the sub-components driving the most cost.

Review each component against the published **Design Principles and Best Practices** for opportunities to reduce cost.

Ensure that all AWS services involved in delivering your Commvault data management services are included in cost analysis and optimization. Consider the active runtime of resources that may only be powered on during data protection activities (i.e., **power-managed MediaAgents**). Consider temporary or ephemeral resources that are terminated after use (i.e., **auto-scaled access nodes**). Also, consider the cost of variable costs like recovery exercises that might result in inter-availability zone network transfer costs. Most importantly, ensure that analysis is continually re-assessing the current data volumes and right-sizing infrastructure to meet business protection needs.

The following table provides a summary of the services that are utilized in a typical Commvault deployment, consult the pricing page for each service to understand the cost implications of usage.

AWS service	Usage
Amazon EC2	Used to provide Commvault application servers which perform AWS and hybrid data protection. Perform optimized network transfer to/from Amazon S3 and between regions.
	Commvault powers down **MediaAgent** components when not in use.
	Commvault **auto-scales Access Nodes** for backups, and terminates when no longer needed. Commvault restore activities use Access Node software on power-managed MediaAgents.
Amazon EBS	Used by Commvault compute instances to store persistent data required to perform data management.
Amazon VPC	Provides internet gateway (GW) or equivalent for accessing AWS global service endpoints.
	Provides VPC endpoints for optimized and secure communication to AWS service endpoints.
	Provides VPC Peering and AWS Transit Gateway to access and protect AWS resources in other accounts and Regions.

AWS service	Usage
	Provides connectivity to hybrid locations through VPN tunnels, Direct Connect, and Transit Gateway.
Amazon S3	Used to store Commvault-optimized backup and archival data copies for operational recovery, discovery recovery, and regulatory compliance needs. Typically deployed per region, to allow regions to operate independently.
Amazon CloudWatch	Used to monitor, analyze, and alarm Commvault compute instances for breaches in key Operating System and Application metrics.
AWS IAM	*Zero-cost service used to secure access to AWS resources to users, groups, and machine identities.*
AWS KMS	Used to create and manage customer-managed keys (CMKs) used to encrypt and decrypt data stored in AWS and optionally Commvault. Typically deployed per region, to allow regions to operate independently.
AWS CloudFormation	Used to deploy and manage Commvault Backup & Recovery infrastructure often deployed from Commvault-supplied AWS Marketplace AMI-based products.
AWS Systems Manager	Used to perform agentless file recovery into running Amazon EC2 instances.
AWS CloudTrail	Used to record, monitor, and audit user activity and events within and across your AWS accounts and services. *Commvault does not configure any Paid Tier AWS CloudTrail services.*

Related documents: COST05-BP02 Analyze all components of the workload.

COST05-BP03 Perform a thorough analysis of each component

Consider the Total Cost of Ownership (TCO) for a component, not just the monthly costs incurred on your AWS bill.

Determine if resources are driving elevated levels of operational burden requiring automation investment or require continual right-sizing and could be consolidated with other workloads.

Commvault is a modular data protection platform that can segregate functional workloads like MediaAgent, Access Nodes, and Search Engine, or combine them. As data needs change your Commvault platform can grow and shrink to match demand.

Consider managed services to reduce or remove the cost of operating your Commvault data management platform. **Commvault Remote Managed Services (RMS)** provides remote monitoring, management, and tuning of your Commvault data management platform 24x7x365. Commvault RMS may be purchased via AWS Marketplace and provides financially backed guarantees for backup and restoration success.

Consider **Amazon Managed Services (AMS)** for your Amazon workloads, including your Commvault data management platform.

Related documents: COST05-BP03 Perform a thorough analysis of each component.

COST05-BP04 Select software with cost-effective licensing

Commvault recommends selecting components with the least-cost software licensing costs.

- Commvault recommends the use of Amazon Linux 2 or Red Hat Enterprise Linux for MediaAgents and Access nodes deployed from **AWS Marketplace** for this reason.
- Cost provides cost comparisons of Linux vs. Windows-based deployments for Commvault components in the **Well-Architected Cost Optimization Pillar**.
- Linux-based MediaAgents provide live browse support for Linux and Windows file systems (see **Live Browse for Restores Using a Linux MediaAgent**), consolidating the infrastructure needs for recovery in Cloud.
- Refer to the **Live Browse Process** used to select a MediaAgent to perform a live browse activity, Commvault supports Live Browse without the need to mount temporary volumes to an Access Node, by leveraging EBS Direct APIs for Live Browse (see **EBS Direct API Restores for Amazon**).

Commvault-embedded Microsoft SQL Server database software is included with your Commvault software purchase price and cannot be replaced with alternative open-source technologies.

Commvault software includes licensing for any bundled components (i.e., **Microsoft SQL Server**) as part of your Commvault purchase price. Selecting Operating Systems that leverage open-source Linux vs. Microsoft Windows is recommended (where supported) to reduce licensing costs. The following are the Commvault data management platform components, and whether they require Microsoft Windows or Linux.

Commvault component	Linux supported	Microsoft Windows supported
CommServe® instance (including **Web Console and Command Center**)	●	●
MediaAgent [1] (including **IntelliSnap®**)	●	●
Web Server	●	●
CommCell Console (standalone)	●	●
Cloud Controller [1]	●	●
Access Node (standalone, auto-scaling) [1, 3, 4, 5]	●	●
Cloud Apps (standalone, auto-scaling) [1, 2]	●	●

1 Typical MediaAgent deployments include Cloud Controller/Access Node, IntelliSnap, and Cloud Apps components.
2 **Cloud Apps (Linux)** and **Cloud Apps (Windows)** provide differing coverage see coverage (below).
3 File search can use a **Linux Access Node** to index Windows instances basic disks, and NTFS file systems.
4 File search must use a **Windows Access Node** to index dynamic volumes, FAT/FAT32/ReFS, and Storage Spaces file-systems
5 Virtual Server Agent for Kubernetes (Amazon EKS protection) is not supported on AWS Graviton instances at the time of writing..=

When protecting Cloud Applications (Cloud storage, Cloud databases, and SaaS applications) protection coverage using open-source Linux or Microsoft Windows is as follows:

Cloud Apps protected workload	Linux supported	Microsoft Windows supported
Cloud databases - **Amazon Aurora**, **Amazon RDS**, **Amazon Redshift**, **Amazon DynamoDB**, **Amazon DocumentDB** - **Microsoft Azure cloud databases**, **MongoDB Atlas**,	●	●
Object Storage (including Amazon S3, Amazon S3 on Outposts) + Includes protection for **Alibaba**, Azure (**Blob**, **File**, **Data Lake**), **Google**, **IBM Cloud**, **Oracle**, and **OpenStack Swift** object storage.	●	●
Azure DevOps and **GitHub**	●	●
Salesforce	●	
Office 365 (including Exchange Online, SharePoint Online, OneDrive for Business, and Teams)		●

Linux support includes ARM64 (recommended) and x86_64 (AMD, Intel).
Linux support on ARM64 excludes Amazon EKS (Kubernetes) protection, Amazon RDS export-based backup, and Salesforce backup.

Amazon EC2 Protection – Linux vs. Windows Access Nodes

A summary of the supported features when using a Linux vs. Windows-based Access Node to perform Amazon EC2 backup and recovery is shown below, if in doubt consult the **VSA Feature Comparison Matrix**.

Amazon Linux-based Access Node Features		
Feature	Linux	Windows
Streaming backups	✓	✓
IntelliSnap backups (snapshot)	✓	✓
Backup copy	✓	✓
Agentless File Recovery	✓	✓
Alerts	✗	✓
Application-aware backups	✗	✓
Attach disks to existing VM	✓	✓
Auto Commit on Kill	✓	✓
Automated Retry for Failed VM Backups	✓	✓
Automatic Discovery of Virtual Machines	✓	✓
Backup set and Subclient filtering (Virtual Machine, Disks)	✓	✓
Cloud MediaAgent Power Management	✓	✓
Commvault Command Center™	✓	✓

Cross-account Snapshot Replication	✓	✓
Failover and Failback Orchestration - VMware to Amazon	✗	✓
Failover Groups (Application orchestrated failover)	✓	✓
File Indexing for Virtual Machines	✓	✓
IntelliSnap®	✓	✓
Live Browse (without Collect File Details during backup)	✓	✓
Live Sync (from AWS to AWS)	✓	✓
Use Resources from the admin account	✓	✓
Multi-VM Restores, Restore full VMs, guest files & folders, agentless file recovery, attach the disk to VM (existing, new)	✓	✓
Access Node Teaming for load distribution	✗	✓
Restore full VMs	✓	✓
Restore guest files and folders	✓	✓
ServiceNow	✓	✓
Snap Replication (cross-region, cross-account)	✓	✓
Subclient filtering (disks)	✓	✓
Subclient filtering (VMs)	✓	✓
View All Versions (If the backup is done with collect File details enabled)	✗	✓
VM Conversion from Azure to Amazon	✓	✓
VM Conversion to Azure	✗	✓
Unicode Support	✓	✓

Source: **Linux Access Node Support for Amazon Web Services**, **Feature Comparison Matrix**

Specific use-cases requiring the use of Microsoft Windows Access Nodes

- If protected Windows guests utilize **dynamic disks** or **ReFS file systems,** a Windows-based Cloud Access Node or MediaAgent will be required to perform a granular browse and restore of files (see **Restoring Guest Files and Folders**).
- If protected Windows guests utilize **Microsoft Windows disk encryption**, a Windows-based Cloud Access Node or MediaAgent will be required for performing granular browsing and restoration of those file systems.

You can model licensing impacts of using Linux-based Commvault instances using the **AWS calculator**.

You should consider the **Best practices for deployment of SQL Server on Amazon EC2**, particularly **Avoid CPU cure mismatches** for your Commvault CommServe® instance.

Commvault has a built-in license manager and does not utilize AWS License Manager to track Commvault software licenses.

Related documents: COST05-BP04 Select software with cost-effective licensing.

Commvault provides cost-based recommendations for all components of your Commvault data protection platform. These recommendations prevent waste by avoiding resources that would result in waste, specifically:

- Use AWS Compute Optimizer to right-size over-provisioned compute instances regularly.
- Use AWS Graviton-based instances for best price-performance compute for streaming MediaAgents.
- Use AMD EPYC x86_64 instances for a 20% price reduction from comparable x86_64 Intel instances.
- Use Linux-based instances for CommServe, MediaAgents, and Access Nodes.
- Use auto-scaling Access Nodes to incur instance runtime costs only during backup activities.
- Use power-managed MediaAgents to reduce runtime costs of MediaAgent instances when unused.
- Use deduplication and compression (default) for all backup transfers to reduce network transfer costs.
- Use Amazon EBS gp3 storage exclusively, for the best price-performance and tuning capability over data lifetime.
- Use Amazon S3 (all storage classes supported) exclusively for Commvault-optimized backup and archive data.
- Use **Commvault Combined Storage Tier** exclusively for Commvault-optimized workload archives.
- Use VPC Endpoints to avoid data transfer to/from AWS service endpoints traversing internet/NAT gateways.

Consider adjacent processes and technologies that your Commvault data management platform will leverage to deliver consistent service to your organization. Leverage application-level services like Amazon Simple Queue (SQS) and Amazon Simple Notification Service (SNS) to centralize and simplify event notifications for your enterprise.

Consider centralizing push notifications via Amazon SNS and your organizational messaging platform (see **How do I use webhooks to publish Amazon SNS messages to Amazon Chime, Slack, or Microsoft Teams?**).

Related documents: COST05-BP05 Select components of this workload to optimize cost in line with organization priorities,

Commvault recommends automatically tuning resources used to protect a workload as it ages. For example, as backup data ages, it becomes less likely to be required in a daily operational recovery.

Commvault recommends taking a tiered approach to data protection of all your AWS services. For example,- maintain a small number of native snapshots for rapid recovery, followed by weekly backups written to Amazon S3 Standard (S3 Standard), followed by yearly backups written to Amazon S3 Infrequent-Access (S3 Standard-IA).

> ⓘ **Note**
>
> When performing **automated storage tiering** with Commvault, data is not moved between storage classes, but **selective copies** are taken to an alternate deduplicated, compressed, and encrypted **Combined Storage Tier**. Commvault recommends using **Amazon S3 Intelligent-Tiering** for automated storage optimization of backup data between 30 days and 90 days of retention.

As innovations in AWS Compute, Network, and Storage services are released, Commvault updates its recommendations in the Cloud Architecture Guide (this document).

Ensure that constant monitoring of your Commvault data management platform costs is occurring. This allows identification of AWS resources that are driving the most cost, and where to spend time on resource optimization.

Ensure you are monitoring Commvault and AWS What's New documentation to identify opportunities for improvement.

Some recent examples of improved AWS service capabilities that could benefit Commvault data management:

- **AWS Free Tier Data Transfer Expansion – 100 GB From Regions and 1 TB From Amazon CloudFront Per Month**
- **Announcing new Amazon EC2 C7g instances powered by AWS Graviton3 processors**
- **Introducing auto-adjusting budgets**

Related documents: COST05-BP06 Perform cost analysis for different usage over time.

Select the Correct Resource Type, Size, and Number

Commvault consists of three (3) primary components in most deployments:

- **CommServe instance** which provides command & control, monitoring, alerting, and provides the self-service Command Center user interface.
- **MediaAgents are** responsible for reading, writing, and replicating backup data to Amazon S3 in the appropriate region.
- **Access Nodes** that communicate with protected workload to perform backup and restore activities (may be co-located with the MediaAgent). Backup access nodes are **auto-scaled** on-demand.

Commvault provides prescriptive guidance on selecting the most appropriate resource type, size, and number of resources to meet availability, performance, and **cost optimization** needs.

A continual **right-sizing approach** should be adopted to ensure that cost and performance needs are met as workload needs change.

Consider the following approaches to understanding and selecting the right resources over time:

COST06-BP01 Perform cost modeling

Perform cost modeling on new workloads to determine the cost of delivering business-required data protection policies.

Ensure that as new workloads are added to your shared data protection platform, service levels for existing workloads are not affected, as this will adversely affect the cost for multiple workloads.

Commvault has performed cost modeling on recommended day-one **seed instances** and then a broad array of compatible compute-optimized, memory-optimized, and general-purpose compute offerings available for day-two **scale-out instances**.

Cost modeling shows some clear trends to consider in your selection of cost-effective resources:

- Linux-based instances are more cost-effective than Microsoft Windows-based instances.
- Memory-optimized instances (R6a, R5a, R6i, R5) represent the least cost option for an all-in-one CommServe + MediaAgent + Access Node deployment.
- AWS Graviton-based instances (C7g, C6g) offer the best price-performance for compute-optimized workloads typically used in streaming MediaAgent use-cases.
- Instance sizes up to 128GB memory have been used with Commvault software but are typically considered ~~cost prohibitive~~ in comparison to scaling horizontally with multiple smaller instances.

CommServe® Least Cost vs. Most-Performant Summary

For CommServe® instances, deploy the smallest recommended instance size (r5a.xlarge) and then scale vertically only when AWS Compute Optimizer indicates the instance is under-provisioned. Commvault recommends the following current generation instance types for the CommServe® instance:

Commvault Scale-out CommServe® instance			
Selection criteria	Network-streamed	Snapshot-based	General Purpose
Least cost (sorted by performance)	c6a.4xlarge c6a.8xlarge ~~c6a.12xlarge~~ ~~c6a.16xlarge~~ c5a.4xlarge c5a.8xlarge ~~c5a.12xlarge~~ ~~c5a.16xlarge~~	**r6a.xlarge** (seed instance) r6a.2xlarge r6a.4xlarge **r5a.xlarge** (seed instance) r5a.2xlarge r5a.4xlarge	m6a.2xlarge m6a.4xlarge m6a.8xlarge m5a.2xlarge m5a.4xlarge m5a.8xlarge
Most performant	c6i.4xlarge c6i.8xlarge ~~c6i.12xlarge~~ ~~c6i.16xlarge~~ c5.4xlarge c5.9xlarge ~~c5.12xlarge~~ ~~c5.18xlarge~~	**r6i.xlarge** (seed instance) r6i.2xlarge r6i.4xlarge r5.xlarge r5.2xlarge r5.4xlarge	m7i.2xlarge m7i.4xlarge m7i.8xlarge m6i.2xlarge m6i.4xlarge m6i.8xlarge m7i-flex.2xlarge m7i-flex.4xlarge m7i-flex.8xlarge m5.2xlarge m5.4xlarge m5.8xlarge
	Compute Optimized	**Memory Optimized**	**General Purpose**

Commvault recommends using the Commvault-maintained Microsoft Windows-based AMI (Commvault Backup & Recovery BYOL) to deploy your new CommServe instance. To reduce cost, you may self-build a Linux-based CommServe® instance (see **Linux CommServe – Requirements**). The following are the indicative costs for the recommended CommServe instance sizes at the time of writing.

CPU and RAM load can be monitored in the AWS Management Console, or the CommCell Command Console using the Infrastructure Load Report (see Commvault **Infrastructure Load Report** for details).

Commvault CommServe® instance Cost Modelling

Best used for network streamed backups	Annual cost USD$ 👤	🟦	Best used for snapshot backup	Annual cost USD$ 👤	🟦	Best used for a mixture of snapshot and streaming backup	Annual cost USD$ 👤	🟦
Best Price-Performance (latest generation preferred for best price-performance)								
c6a.4xlarge	3,309.97	9,757.33	**r6a.xlarge**	**1,226.66**	**2,838.50**	m6a.2xlarge	1,869.12	5,092.80
c6a.8xlarge	6,619.93	19,514.65	r6a.2xlarge	2,453.24	5,676.92	m6a.4xlarge₁	3,738.33	10,185.69
~~c6a.12xlarge~~	~~9,929.90~~	~~29,271.98~~	r6a.4xlarge	4,906.56	11,353.92	m6a.8xlarge	7,476.57	20,371.29
~~c6a.16xlarge~~	~~13,239.86~~	~~39,029.30~~	**r5a.xlarge**	**1,165.08**	**2,776.92**	m5a.2xlarge	1,769.52	4,993.20
c5a.4xlarge	3,171.12	9,618.48	r5a.2xlarge	2,330.16	5,553.84	m5a.4xlarge	3,547.80	9,995.16
c5a.8xlarge	6,342.24	19,236.96	r5a.4xlarge	4,660.32	11,107.68	m5a.8xlarge	7,086.84	19,981.56
~~c5a.12xlarge~~	~~9,522.12~~	~~28,864.20~~						
~~c5a.16xlarge~~	~~12,693.24~~	~~38,482.68~~						
Best Performance (latest generation preferred for best price-performance)								
c6i.4xlarge	3,677.71	10,125.07	**r6i.xlarge**	**1,362.88**	**2,974.72**	m7i.2xlarge	2,180.71	5,404.39
c6i.8xlarge	7,355.42	20,250.14	r6i.2xlarge	2,725.85	5,949.53	m7i.4xlarge	4,361.34	10,808.70
~~c6i.12xlarge~~	~~11,033.22~~	~~30,375.30~~	r6i.4xlarge	5,451.70	11,899.06	m7i.8xlarge	8,722.68	21,617.40
~~c6i.16xlarge~~	~~14,710.93~~	~~40,500.37~~	r5.xlarge	1,296.48	2,908.32	m6i.2xlarge	2,076.82	5,300.50
c5.4xlarge	3,504.00	9,951.36	r5.2xlarge	2,592.96	5,816.64	m6i.4xlarge	4,153.64	10,601.00
c5.9xlarge	7,884.00	22,390.56	r5.4xlarge	5,194.68	11,642.04	m6i.8xlarge	8,307.37	21,202.09
~~c5.12xlarge~~	~~10,512.00~~	~~29,854.08~~				m7i-flex.2xlarge	2,071.48	: 5,133.97
~~c5.18xlarge~~	~~15,759.24~~	~~44,772.36~~				m7i-flex.4xlarge	4,143.04	10,268.03
						m7i-flex.8xlarge	8,286.08	20,536.07
						m5.2xlarge	1,979.76	5,203.44
						m5.4xlarge	3,959.52	10,406.88
						m5.8xlarge	7,910.28	20,805.00
Compute Optimized			**Memory Optimized**			**General Purpose**		

Pricing is computing instance only, deployed in us-east-1, EC2 Instance Savings Plan – 1 year, All upfront. Seed instances (day one deployment) are in **bold**. Use **AWS Pricing Calculator**.

MediaAgent Least Cost vs. Most Performant Modelling

At the time of writing (September 2022), Commvault recommends using Amazon C7g family instances exclusively for MediaAgent / Access Nodes. See below for the pricing for both Linux and Windows-based Access Nodes. Commvault recommends using Linux exclusively unless your workload requires a Windows-based Access Node (e.g., Microsoft Office 365 protection)

Instance size	Annual USD$ 🔥	Instance size	Annual USD$ 🔥 🟦	Instance size	Annual USD$ 🔥 🟦
c7g.large	745.44	c6a.large	780.48 / 1,586.40	c6i.large	855.00 / 1,660.92
c7g.xlarge	1,380.60	c6a.xlarge	1,450.68 / 3,062.52	c6i.xlarge	1,599.60 / 3,211.44
c7g.2xlarge	2,650.80	c6a.2xlarge	2,790.96 / 6,014.64	c6i.2xlarge	3,088.80 / 6,312.48
c7g.4xlarge	5,191.20	c6a.4xlarge	5,471.52 / 11,918.88	c6i.4xlarge	6,067.20 / 12,514.56
c7g.8xlarge	10,272.00	c6a.8xlarge	10,832.64 / 23,727.36	c6i.8xlarge	12,024.00 / 24,918.72
c7g.12xlarge	15,352.80	c6a.12xlarge	16,193.76 / 35,535.84	c6i.12xlarge	17,980.80 / 37,322.88
c7g.16xlarge	20,433.60	c6a.16xlarge	21,554.88 / 47,344.32	c6i.16xlarge	23,937.60 / 49,727.04

In some AWS regions and Availability Zones, the latest generation instance types may not be. The following shows cost modeling for previous-generation instance types ideal for MediaAgent and/or MediaAgent + Access Node workloads.

Instance size	Annual USD$ 🔥	Instance size	Annual USD$ 🔥 🟦	Instance size	Annual USD$ 🔥 🟦
c6g.large	706.08	c5a.large	784.92 / 1,590.84	c5.large	855.00 / 1,660.92
c6g.xlarge	1,301.76	c5a.xlarge	1,459.44 / 3,071.28	c5.xlarge	1,599.60 / 3,211.44
c6g.2xlarge	2,493.12	c5a.2xlarge	2,808.48 / 6,032.16	c5.2xlarge	3,088.80 / 6,312.48
c6g.4xlarge	4,875.84	c5a.4xlarge	5,506.56 / 11,953.92	c5.4xlarge	6,067.20 / 12,514.56
c6g.8xlarge	9,641.28	c5a.8xlarge	10,902.72 / 23,797.44	c5.9xlarge	13,513.20 / 28,019.76
c6g.12xlarge	14,406.72	c5a.12xlarge	16,298.88 / 35,640.96	c5.18xlarge	26,916.00 / 55,929.12
c6g.16xlarge	19,172.16	c5a.16xlarge	21,695.04 / 47,484.48	c5.24xlarge	35,851.20 / 74,535.36

Commvault MediaAgent Cost Comparison

Commvault recommends using AWS Graviton instances exclusively for streaming MediaAgent workloads, as they offer the best price-performance for CPU, memory, and network bandwidth utilized in network-heavy streaming data workloads.

Commvault has performed **performance benchmarking** for MediaAgents and Access Nodes that leverage the Amazon EC2 C7g family to help align achievable RPO/RTO with differing sizes.

Commvault Snapshot + Streaming MediaAgent Grid – Indicative Compute Savings Plan Annual Price					
Instance size	Max stored data (TB)	1 node	2 nodes	3 nodes	4 nodes
♨ c7g.xlarge	400	$960.12	$1,920.24	$2,880.36	$3,840.48
♨ c7g.2xlarge	800	$1,809.84	$3,619.68	$5,429.52	$7,239.36
♨ c7g.4xlarge	1200	$3,399.76	$6,799.52	$10,199.28	$13,599.04
♨ c7g.8xlarge	2400	$6,909.91	$13,819.82	$20,729.73	$27,639.64
♨ c7g.12xlarge	3600	$12.308.79	$24,617.58	$36,926.37	$49,235.16
♨ c7g.16xlarge	4800	$13,708.55	$27,417.10	$41,125.65	$54,834.20

Estimates USD$ in the us-east-1 region as of Sep 2022 with 295GiB gp3 storage. Use the AWS calculator to validate pricing in your region.
Prices include compute and base 115GB gp3 Amazon EBS storage. IOPS and throughput values are excluded from volume estimates.

In environments where the **cost per protected TB** reduction is the primary selection criteria for resources, an AWS Graviton-based memory-optimized or general-purpose-based instance may be more suitable to meet cost optimization demands.

> ⓘ Note
>
> Any backup and recovery workload can be serviced by any instance type, indications (below) of network vs. snapshot usage are recommendations only.

Commvault Scale-out MediaAgent Instance – Indicative Compute Savings Plan Annual Price					
Network-streamed		Snapshot-based		General Purpose	
c7g.4xlarge	$3,682.96 ♨	r6g.xlarge	**$1,478.06 ♨**	m6g.2xlarge	$2,137.69 ♨
c7g.8xlarge	$7,082.71 ♨	r6g.2xlarge	$2,672.93 ♨	m6g.4xlarge	$3,991.31 ♨
c7g.12xlarge	$10,481.59 ♨	r6g.4xlarge	$5,063.53 ♨	m6g.8xlarge	$7,699.42 ♨
c7g.16xlarge	$13,881.35 ♨				
c6a.4xlarge	$10,346.51 ▥	r6a.xlarge	**$3,238.47 ▥**	m6a.2xlarge	$5,582.82 ▥
c6a.8xlarge	$20,409.74 ▥	r6a.2xlarge	$6,193.75 ▥	m6a.4xlarge	$10,882.45 ▥
c6a.12xlarge	$30,473.05 ▥	r6a.4xlarge	$12,104.38 ▥	m6a.8xlarge	$21,481.70 ▥
c6a.16xlarge	$40,536.28 ▥				
Compute Optimized		**Memory Optimized**		**General Purpose**	

Estimates USD$ in the us-east-1 region as of September 2022 with 295GiB gp3 storage. Use the AWS calculator to validate pricing in your region.
Seed instances are shown in bold.
Excludes Amazon EBS snapshots, Amazon EBS direct API backup estimates, and Amazon S3 and Amazon EBS direct endpoint costs.
Excludes Amazon S3 backup data storage and associated API (GET, PUT), and network transfer fees.
Excludes **Amazon Support Costs**.
Excludes Commvault IOPS requirements for high IOPS volumes (deduplication database, IndexCache).
* Instance is EBS optimized with burstable EBS and network performance for 30-60 mins per day.

Consider modeling your environment by experimenting with multiple configurations that use single node configuration without resilience, and then day-two configurations distributing the load across multiple instances for performance and high availability.

Also, consider all costs when performing cost modeling, for example, see the following modeling on the Total Cost of Ownership (TCO) for a simple Commvault deployment.

Using publicly available pricing for AWS resources (as of March 2020) the cost of performing protection in AWS by utilizing any combination of iDataAgents and coupled with Commvault IntelliSnap for AWS snapshot management, the following becomes a rough estimate of the cost of the AWS infrastructure required for 1 year with a Partial upfront payment.

The single largest contributor is the **S3-IA backup storage** followed by the **CommServe** Amazon EC2 instance. Commvault can help control the cost of your **MediaAgent** by powering down MediaAgents during periods of inactivity.

Related documents: COST06-BP01 Perform cost modeling.

COST06-BP02 Select resource type, size, and number based on data

Commvault recommends that all resource selections start with the smallest (i.e., cheapest) resource type and size supported. Use **AWS Compute Optimizer** and the **Recovery Readiness Report** to determine if business-agreed RPOs and RTOs can be met.

Commvault recommends leveraging **memory-optimized** instances for the least cost, and scaling to **compute-optimized** instances for workloads performing streaming network transfers, or **general purpose** for mixed-mode environments.

Commvault provides prescriptive guidance on selecting instance type and size for initial or **seed deployments** and day 2 or expansion **scale-out deployments**. Recommendations have been developed based on matching instances to Commvault software requirements, lab-based validation, and observed customer deployments.

There are four recommendations split across **day one seed** deployments and **day two scale-out** deployments.

- **Seed CommServe® instance**
- **Seed MediaAgent Grid**
- **Scale-out CommServe® instance**
- **Scale-out MediaAgent Grid**

ⓘ Note

The cost modeling provided below excludes Amazon EBS capacity, IOPS, and throughput beyond the day one configuration, as growth needs are unique to data protection type and data volume. Additionally, data transfer and data storage costs are excluded as they are heavily dependent on the data protection type (snapshot, streaming) and overall data capacity.

Seed CommServe® instance

As an example, the day-one **seed architecture** costs for a **Commvault Backup & Recovery BYOL** server from the AWS Marketplace vary by the Amazon EC2 instance size selected.

Refer to the **Design Principles and Best Practices – CommServe Sizing** for detailed instance information.

Commvault Seed CommServe® instance – Indicative Price Comparison			
Amazon EC2 instance size	Configuration	Annual on-demand	Annual Compute savings
r5a.xlarge* least cost Up to 10 Gigabit network bandwidth*** 504GB/hr. avg. backup throughput*** 1424GB/hr. avg. restore throughput***	35 GiB Amazon EBS root volume (gp3) 260 GiB Amazon EBS app volume (gp3) *Protect up to 3.6K EBS volumes per account, per supported region, per day, see* **service quotas*** *Protect up to 12TB via Commvault streamed backup per day.* *Restore up to 34TB via Commvault streamed recovery per day (burst).*	$2,262.96 🪣 $3,874.80 🪟	$1,623.48 🪣 $3,235.32 🪟
r6a.xlarge* Up to 12.5 Gigabit network bandwidth*** 449GB/hr. avg. backup throughput*** 443GB/hr. avg. restore throughput***	35 GiB Amazon EBS root volume (gp3) 260 GiB Amazon EBS app volume (gp3) *Protect up to 3.6K EBS volumes per account, per supported region, per day, see* **service quotas*** *Protect up to 10.7TB via Commvault streamed backup per day.* *Restore up to 10.6TB via Commvault streamed recovery per day.*	$2,269.92 🪣 $3,881.76 🪟	$1,626.63 🪣 $3,238.47 🪟
r6i.xlarge* most performant Up to 12.5 Gigabit network bandwidth*** 449GB/hr. avg. backup throughput*** 561GB/hr. avg. restore throughput***	35 GiB Amazon EBS root volume (gp3) 260 GiB Amazon EBS app volume (gp3) *Protect up to 3.6K EBS volumes per account, per supported region, per day, see* **service quotas*** *Protect up to 10.7TB via Commvault streamed backup per day.* *Restore up to 13.4TB via Commvault streamed recovery per day.*	$2,490.72 🪣 $4,102.56 🪟	$1,775.90 🪣 $3,387.74 🪟

* **Amazon EBS-optimized instances** can support maximum performance for 30 minutes at least once every 24 hours.

** Commvault software creates 5 snapshots at a time by default, to avoid exceeding **Amazon EBS resource quotas**.

*** Amazon EC2 instance has a **baseline and burst network performance**, test in your VPC for achievable throughput.

*** Avg. throughput used tuned configuration with **ReadAhead**=256, **WriteBehind**=256 for optimal transfer speed.

Estimates USD$ in the us-east-1 region as of September 2022. Use the AWS calculator to validate pricing in your region.

Excludes Amazon EBS snapshots, Amazon EBS direct API backup estimates, Amazon S3, and Amazon EBS direct endpoint costs.

Excludes Amazon S3 backup data storage and associated API (GET, PUT), and network transfer fees.

Excludes **Amazon Support Costs**.

Seed MediaAgent Grids

As an example, the day-one **seed architecture** costs for a **Commvault Cloud Access Node ARM** or **Commvault Cloud Access Node x86_64** server from the AWS Marketplace vary by the Amazon EC2 instance size selected and the **backup type** being used.

Commvault uses MediaAgent grids to orchestrate Amazon EC2, Amazon RDS, and many other AWS cloud database snapshots within and across regions and accounts. Additionally, grids can take network-streamed backups of your AWS services that are independent of the AWS service and/or region. Your MediaAgent grid must be sized based on the predominant type of protection being performed:

- **Snapshot-only** grids are used in environments performing snapshot-based protection exclusively.
- **Snapshot + Streaming** grids are used in environments performing a mixture of snapshot-based and network-streamed protection.

You can use the Amazon EC2 **Change instance type** functionality to convert between the EC2 instance types as your data protection needs change.

Commvault recommends deploying **deduplication-enabled MediaAgent grids** <u>only</u> in AWS, as the use of deduplication and compression reduces data transferred, and therefore conserves valuable network burst credits.

Commvault recommends using Amazon Graviton exclusively for MediaAgent grids due to the best price performance of any general-purpose compute in Amazon EC2.

A **seed MediaAgent grid** is used to expand an existing Commvault data management platform by:

- Expanding the volume of data managed within a region by growing the number of MediaAgent nodes performing data management in highly-available MediaAgent grids.
- Expanding the volume of data managed across regions by providing regionalized MediaAgent infrastructure to manage data into and out of the new region. Providing MediaAgent grids within a region allow each grid to operate independently, providing fault isolation of regional events.

Your CommServe® instance contains the MediaAgent grid software and functionality, you only need to expand as your protection data volume or data management availability needs to change.

Snapshot-only Seed MediaAgent Grid

Snapshot-only MediaAgent grids are used to protect Amazon EC2, Amazon RDS, Amazon Redshift, and Amazon DocumentDB resources using Commvault IntelliSnap® snapshot protection <u>only</u>.

Look to the Snapshot and streaming seed MediaAgent grid if performing **backup copy** or cross-region **periodic replication**.

Refer to the **Design Principles and Best Practices – MediaAgent Sizing** for detailed instance information.

Commvault Seed MediaAgent (Snapshot only) – *Indicative Price Comparison*			
Amazon EC2 instance size	Configuration	Annual on-demand	Annual Compute savings
t4g.small* (**least cost**) Up to 5 Gigabit network bandwidth	**t4g.small** 10 GiB Amazon EBS root volume (gp3) 105 GiB Amazon EBS app volume (gp3) *Protect up to 3.6K EBS volumes per account, per supported region, per day, see* **service quotas**.	$257.52 🐧	$209.39 🐧
t3a.small* (**AMD EPYC alternative.**) Up to 5 Gigabit network bandwidth	**t3a.small** 10 GiB Amazon EBS root volume (gp3) 105 GiB Amazon EBS app volume (gp3) *Protect up to 3.6K EBS volumes per account, per supported region, per day, see* **service quotas**.	$275.04 🐧 $436.32 🪟	$221.65 🐧 $382.84 🪟
t3.small* (**Intel alternative.**) Up to 5 Gigabit network bandwidth	**t3.small** 10 GiB Amazon EBS root volume (gp3) 105 GiB Amazon EBS app volume (gp3) *Protect up to 3.6K EBS volumes per account, per supported region, per day, see* **service quotas**.	$292.56 🐧 $453.84 🪟	$233.04 🐧 $394.22 🪟

* **Amazon EBS-optimized instances** can support maximum performance for 30 minutes at least once every 24 hours.

Backup testing was performed in us-east-1 with 50 instances, 360GiB dataset size.

Pricing for Amazon EC2 instance with gp3 volumes, **Compute Savings Plan**, 1-year commit, All upfront.

It should be noted that a **snapshot-only** MediaAgent cannot be used for any streamed backup as the memory specification does not meet Commvault minimum requirements.

It should be noted that Amazon EC2 instance sizes below a T4 or T3 small (i.e., nano, micro) cannot be used for Commvault snapshot protection as they do not meet Commvault minimum CPU and RAM requirements.

Snapshot and streaming seed MediaAgent Grid

Snapshot and streaming MediaAgent grids are to be used for Amazon EC2, Amazon RDS, Amazon Redshift, and Amazon DocumentDB IntelliSnap® snapshot-based backup and recovery. They may also be used for network-streamed protection such as Amazon RDS export-based protection, Amazon DynamoDB, Amazon EFS, Amazon FSx, and Amazon S3 backup.

Use streaming MediaAgent grids to perform network-optimized replicated **backup copy** and cross-region **instance/backup replication** for Disaster Recovery. Commvault recommends **compute-optimized** EC2 instances for best price-performance, and **memory-optimized** instances for least cost MediaAgents.

Refer to the **Design Principles and Best Practices – MediaAgent Sizing** for detailed instance information.

Commvault Seed MediaAgent (Snapshot + Streaming) – *Indicative Price Comparison*			
Amazon EC2 instance size	Configuration	Annual on-demand	Annual Compute savings
c7g.xlarge* **least cost** Up to 12.5 Gigabit network bandwidth** *156GB/hr. avg. backup throughput**** *183GB/hr. avg. restore throughput****	**c7g.xlarge** 10 GiB Amazon EBS root volume (gp3) 105 GiB Amazon EBS app volume (gp3) *Protect up to 3.6K EBS volumes per account, per supported region, per day, see* **service quotas****** *Protect up to 3.7TB via Commvault streamed backup per day.* *Restore up to 4.3TB via Commvault streamed recovery per day.*	$428.40 👤	$323.27 👤
c6a.xlarge* **AMD EPYC alt.** Up to 12 Gigabit network bandwidth** *Baseline throughput not benchmarked.*	**c6a.xlarge** 10 GiB Amazon EBS root volume (gp3) 105 GiB Amazon EBS app volume (gp3) *Protect up to 3.6K EBS volumes per region and account, per day, see* **service quotas******	$780.48 👤 $1,586.40 🪟	$562.42 👤 $1,368.34 🪟
c6i.xlarge* **Intel alt.** Up to 12.5 Gigabit network bandwidth** *Baseline throughput not benchmarked*	**c6i.xlarge** 10 GiB Amazon EBS root volume (gp3) 105 GiB Amazon EBS app volume (gp3) *Protect up to 3.6K EBS volumes per region and account, per day, see* **service quotas******	$855.00 👤 $1,660.92 🪟	$612.61 👤 $1,418.53 🪟

* **Amazon EBS-optimized instances** can support maximum performance for 30 minutes at least once every 24 hours.

** Commvault software creates 5 snapshots at a time by default, to avoid exceeding **Amazon EBS resource quotas**.

*** Amazon EC2 instance has a **baseline and burst network performance**, test in your VPC for achievable throughput.

*** Avg. throughput used tuned configuration with **ReadAhead**=256, **WriteBehind**=256 for optimal transfer speed.

Pricing for Amazon EC2 instance with gp3 volumes, **Compute Savings Plan**, 1-year commit, All upfront.

Scaling Guidance

Commvault recommends starting with the smallest instance size for CommServe, MediaAgent, and Auto-scaling Access Nodes, and then increasing instance size only when AWS Compute Optimizer indicates the instance is under-provisioned. Additional nodes may be added to a MediaAgent grid up to a maximum of four (4) per grid, nodes may also be vertically scaled.

Commvault recommends using multiple minimally sized Access Nodes to gain improved performance and improve the resiliency of backup operations vs. vertically scaling deployed resources

Consult the **AWS Pricing Calculator** for the best pricing in your AWS region, Commvault has provided two recommended EC2 instance sizes below for each component. Commvault has selected instances for best baseline performance and cost.
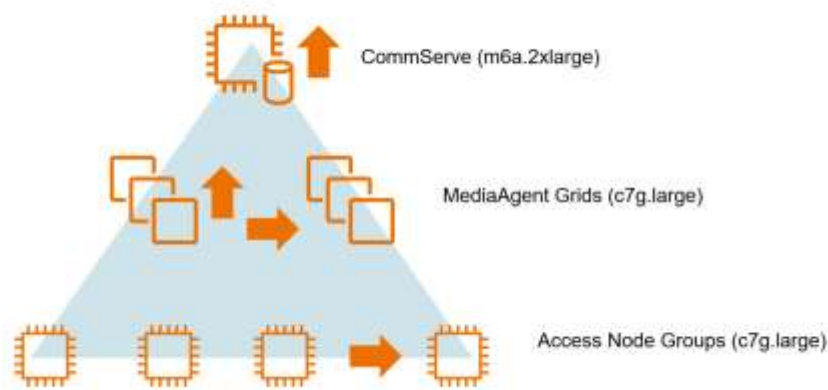
Commvault expects you will scale in three dimensions:

1. A single CommServe will be scaled vertically to a limit of **20,000 virtual instances** protected.
2. One or more MediaAgents will be scaled vertically to a maximum of **1000TB of managed cloud data**, then scaled horizontally to a **maximum of 4000TB** per MediaAgent grid.
3. Access Nodes (also referred to as Virtual Server Agents) will be scaled horizontally with smaller instance types for cost control. Vertical scaling will only be required when the size of a given EC2 instance exceeds desired backup window.

It should be noted that **Automatic scaling for Amazon Access Nodes** (Backup only) removes the requirement to plan and maintain Access Nodes, they are provisioned, used, then terminated after a period of non-use.

See the **Well-architected – Cost Optimization – Cost-effective resources** section for a price comparison of recommended MediaAgent configurations.

MediaAgent Scaling Guidance

MediaAgents grids provide a pooled resource of CPU to process data and a locally hosted Deduplication DataBase (DDB) which is responsible for deduplicating a subset of the data for the grid. MediaAgent scalability is highly dependent on the amount of parallel backup, recovery, and replication activity and data types.

Commvault recommends starting with the smallest instance size (c7g.xlarge) and using the following guidelines to determine the most appropriate scaling option:

- If the instance network is under-utilized, but the deduplication query and insert (Q&I) time is exceeding 2 milliseconds best practice, **add a second DDB volume**.

- If the host network is under-utilized, but CPU and/or RAM is exhausted, consider **increasing the instance size**.

- if the host network is over-utilized, or high availability for data management is desired, **consider adding another node of the same size** to distribute the workload.

If current generation instances are not available in your target region, previous generation instances may be utilized.

The following table models the total written backup data (in TB) that each C7 family instance can theoretically manage. The table is based on observation of real-world environments and varies based on the data types, retention period, and change rate of data between backups.

The relative activity of a deduplication database (DDB) is critical in planning, a c7g.xlarge instance is capable of handling large volumes of managed data where a subset of managed data is *active*, but the remainder consists of infrequently updated long-term retention or compliance copies.

Indicative total written deduplicated backup data in Amazon S3 (1 DDB volume / 2 DDB volumes)			
Grid Configuration	c7g.xlarge	c7g.2xlarge	c7g.4xlarge
1 Node	25 / 50	50 / 100	150 / 300
2 Node	50 / 100	100 / 200	300 / 600
3 Node	100 / 200	200 / 400	450 / 900
4 Node	200 / 400	400 / 800	600 / 1200

Derived from **Hardware Specifications for Deduplication Mode**

Note: Commvault does not support MediaAgent grids larger than four (4) nodes.

If you do not know your **stored backup data** size, multiply your client data size by 1.6x for an estimate of back-end data for a 90-day retention period. Commvault provides the Commvault Solution Design Tool (CSDT) to partners to help model the storage and compute costs in multi-region deployments, speak with your Commvault sales representative for assistance in modeling your environment.

Commvault recommends **compute-optimized instances** for MediaAgent/Access Nodes as these instances are designed with an ideal mix of compute, memory, and networking resources to service the compute-intensive data management processes that Commvault software performs.

Occasionally a time-based data management event may occur, like indexing a long-term archive for an eDiscovery event or performing a one-time data migration with Commvault. During these events, the day-to-day C Class instances may not meet the combined compute or memory requirements, and **memory-optimized** or **general-purpose** instances may be utilized.

> ⓘ **Note**
>
> All configurations are sized with either one or two 25GiB deduplication database (DDB) volumes. DDB volumes will require growth per your specific data patterns. Indicative DDB sizing is found at **Hardware Specifications for Deduplication Mode** but should be used for paper-based planning only.

Scale-out CommServe® instance

As your protected data grows you will need to scale your CommServe® instance to handle the increased data management activity. Commvault observes growth in one of three (3) primary growth modes:

- **Streamed data management** growth is observed when the volume of network-streamed backup copies, DASH copies, and disaster recovery replication increases across protected regions and edge locations.
- **Snapshot management** growth is observed when the volume of AWS-native snapshots being created, shared, and copied across regions and accounts increases.
- **Data intelligence** growth is observed as the volume of data intelligence or data insight analytics, alerting, and reporting grows across protected regions and edge locations.

Commvault provides holistic data management across regions and edge locations and has used its extensive footprint of customer environments to identify the following scaling patterns. Using Amazon EC2 **Change Instance Type** functionality a CommServe server may be scaled to meet your business data management needs

16X 12X 8X — Network-Streamed CommServe® Server (C6, C6i) 4X

8X 4X 2X — Seed CommServe® Server (R6a, R6i) XL

8X 4X — Data Intelligence CommServe® Server (M6s, M6i, M7i M7i-flex) 2XL

Streamed protection and replication of Amazon compute, container, database and storage

Snapshot-based protection of Amazon compute and databases. Some streamed backup and replication.

Data intelligence focus on indexing protected data, identifying waste, sensitive, and PII data.

Selecting a particular instance size should be driven by exhaustion of compute, memory, or both as observed in **Amazon CloudWatch** and/or **AWS Compute Optimizer**.

Consult the **Commvault CommServe® Instance Cost Modelling** for modeling on CommServe scale-out options.

Scale-out MediaAgent Grids

Commvault MediaAgent grids are responsible for collecting backup data, optimizing for reduced storage and replication cost, and then writing to Amazon S3. MediaAgents grids refer to multiple MediaAgents working in a pool, where the loss of one or many nodes redistributes backup and recovery workload to remaining instances.

Refer to the **MediaAgent Cost Modelling** for an indication of the costs associated with scaling from a single node to multi-node grids for additional managed storage. Nodes are added for high availability and increasing network throughput of backup and recovery operations.

Commvault has performed benchmark testing using a combined MediaAgent + Access Node configuration and tested the throughput and cost when using single larger nodes vs. multiple smaller nodes. The results show that starting smaller will minimize day one cost, and deploying multiple smaller nodes results in an identical cost to a single larger instance. An additional benefit gained by scaling horizontally to multiple nodes is the added resilience of backup and recovery operations, which are restarted on healthy nodes in the event of a failure.

Amazon EBS gp3 volumes should be used exclusively to hold a localized deduplication database (DDB) used to reduce data transfers during backup operations only. When the DDB Query & Insert (Q&I) time exceeds Commvault's best practices (less than 2 milliseconds), additional IOPS and/or throughput may be provisioned to the volume.

When instance maximum EBS IOPS and bandwidth limits are reached, the instance size may be increased.

> ⚠ **Important**
>
> Commvault instance sizes and IOPS recommendations are guidelines only. You may choose to continue using a fully consumed resource or EBS volume, with the acceptance of reduced backup and recovery performance. Commvault can be sized to meet your performance or cost needs.

Related documents: COST06-BP02 Select resource type, size, and number based on data.

COST06-BP03 Select resource type, size, and number automatically based on metrics

Commvault provides **automatic-scaling for Access Nodes** which automatically selects instance type, architecture, operating system, and size based on the data volume to protect and configured *recovery point objective (RPO)*. Auto-scaling will automatically and dynamically select the most appropriate number of resources at backup runtime, and then power-down, and then terminate resources when no longer required.

Commvault will **auto-scale** the required number of resources required for backup and backup copy operations based on the Recovery Point Objective (RPO), throughput rate (GB/hr.), and the concurrent number of backup operations configured.

Commvault will default to the use of Amazon Graviton (arm64) instances but will revert to using x86_64 (AMD, Intel) when Graviton instances are unavailable.

Commvault makes the scaling decision once, during the initiation of the backup, and will not reevaluate backup metrics until the next scheduled backup.

Commvault recommends using Amazon CloudWatch **instance metrics** to measure, monitor, and alarm static thresholds for CPU, memory, and network utilization to influence upgrades of your CommServe and MediaAgent grids. MediaAgents also benefit from monitoring **EBS performance** and **ENA network interfaces** to determine when available resources are exhausted.

Related documents: COST06-BP03 Select resource type, size, and number automatically based on metrics.

Select the Best Pricing Model

COST07-BP01 Perform pricing model analysis

Commvault recommends using **AWS EC2 Savings Plans** and **AWS Reserved Instances (RIs)** for Commvault CommServe and MediaAgent instances which are typically running for extended periods.

Compute Savings Plans offer the most flexibility over regions and usage types, whereas Reserved Instances are linked to specific locations/regions, instance types, and operating systems.

Auto-scaled access nodes do not require savings plans as they are provisioned and terminated based on backup activity and grow and shrink based on the number of workloads to protect.

> ⓘ **Note**
>
> Commvault resources cannot be run as Amazon EC2 Spot instances.

Related documents: COST07-BP01 Perform pricing model analysis.

COST07-BP02 Implement Regions based on cost

AWS pricing varies by region, consider selecting regions to reduce the cost of compute, network, and storage for non-primary workloads like DR data bunkers.

Be sure to consider the latency to your users and network transfer costs for getting the data back to your preferred or primary region(s) after a recovery event.

Related documents: COST07-BP02 Implement Regions based on cost.

COST07-BP03 Select third-party agreements with cost-efficient terms

Commvault pricing is not tied to the consumption of AWS resources or AWS spend. Commvault pricing & licensing terms are aligned with the total amount of protected workloads.

Related documents: COST07-BP03 Select third-party agreements with cost-efficient terms.

COST07-BP04 Implement pricing models for all components of this workload

Commvault recommends that a mixture of **pricing models** be employed based on the observed usage of a component.

CommServes are running 24x7 and will benefit from Savings Plans and/or **Reserved Instances** with a one or three-year commitment.

MediaAgents for baseline protection workloads that operate for at least 50% of the time can also benefit from Savings Plans and Reserved Instances.

Access Nodes typically only run during the backup window (nightly) and are typically an on-demand purchase. As your data volume grows and a baseline usage can be observed, Reserved Instance (RI) commitments can be made.,

Consider **Savings Plans** when the instance family, size, OS, or Region choices are fluid and may change during the commitment period.

Consider migrating temporary or project-based work out onto on-demand MediaAgents grids where costs can be better observed.

Related documents: COST07-BP04 Implement pricing models for all components of this workload.

COST07-BP05 Perform pricing model analysis at the master account level

Regularly review the about of sustained consistent compute usage across all AWS accounts using **AWS Cost Explorer**.

Finance teams may opt to convert on-demand costs to 1 or 3-year commitments after a pattern of usage is observed in AWS Cost & Usage Reporting.

Consider whether a workload is stable or still in the growing phase and will require a change in instance type, size, or OS over the commitment period (**Savings Plans**).

CommServe instances typically experience little change, whereas MediaAgents and Access Nodes will fluctuate in instance type and size based on data growth.

AWS Cost Explorer will make recommendations for commitments based on observed **on-demand instance usage**.

Related documents: COST07-BP05 Perform pricing model analysis at the management account level.

Plan for Data Transfer

COST08-BP01 Perform data transfer modeling

Ensure you are aware of the data transfer flows and corresponding costs of data transfer in your Commvault data management platform.

- Minimal command and control and indexing data are transmitted by Commvault data management activities between the CommServe, MediaAgents, and Access Nodes.
- When streamed backup copies are taken, unique workload blocks are transferred to and deduplicated on the Access Node (same AZ), and then to the MediaAgent (same region, and optionally the same AZ).
- When streamed backup copies or service-independent data copies and cross-region copies are created, data is replicated between MediaAgents in a deduplicated format (see **DASH Copies**).
- Restores are fully hydrated at the MediaAgent and then transferred to the target region and availability zone, for this reason, a regional data copy is always advised to avoid **cross-region data transfer fees**.

Review the **AWS pricing pages** and ensure that backup and recovery network transfers by Commvault align with the agreed business value. Data transfer frequency directly impacts achievable application RPOs and RTOs.

In rare cases where a rapid exit of an owned or leased facility is required by the business, backups for an edge location may be stored in the AWS Region. All restores in this scenario will incur a **Data Transfer OUT** egress fee.

Perform modeling based on the best-case and worst-case restore scenarios to review with your finance controllers and risk management representatives.

Related documents: OST08-BP01 Perform data transfer modeling.

COST08-BP02 Select components to optimize data transfer cost

Commvault recommends using **compute-optimized** Amazon C7g/C6g instances when optimizing for data transfer. These instances offer the best price-performance and network baselines across instance types.

Commvault will **attempt to use Access Nodes** that are in the same Availability Zone (AZ) as the workload being backed up or restored, effectively optimizing for minimal cross-AZ data transfer costs.

Commvault Access Nodes will read and write data to in-region MediaAgent grids, and Commvault will distribute backup data across nodes using block hashes. Commvault recommends cross-AZ MediaAgent Grids Cloud Storage Pool scalability, resilience, and performance.

Commvault recommends using **source-side deduplication** (on the client or Access Node) and compression for all data transfers to reduce the amount of data placed on the network and conserve **network credits**.

Related documents: COST08-BP02 Select components to optimize data transfer cost.

COST08-BP03 Implement services to reduce data transfer costs

Commvault recommends the use of regional **Amazon S3** and **Amazon EBS** endpoints provided by **AWS PrivateLink** to prevent backup and recovery transfers from transiting Internet Gateway (IGW) or NAT Gateway (NGW) devices and reduce transfer costs.

Commvault recommends implementing **AWS Direct Connect** where bandwidth demands can no longer be met with **AWS VPN** connectivity.

Commvault recommends locating at least one MediaAgent + Access Node per region to ensure that primary backups are written to Amazon S3 in-region, then replicated using deduplicated **DASH Copies**. This approach minimizes the **network egress costs** incurred by providing an *operational recovery point* and a remote out-of-region *disaster recovery point*.

Be aware that every decision to reduce transfer costs may impact the speed of transfer, impacting your Recovery Time Objective (RTO). Be sure to measure changes against business-agreed SLAs.

Related documents: COST08-BP03 Implement services to reduce data transfer costs.

Manage demand and supply resources

COST09-BP01 Perform an analysis on the workload demand

Commvault recommends maintaining a workload calendar within operational documentation. Cloud operations and platform administrators should factor known peaks in workloads like end-of-quarter backups and end-of-year archives into workload cost profiles.

Platform operations should be aware of high-impact seasonal activities (i.e., increased load sales events) due to the potential impact on shared MediaAgent infrastructure.

Consider the ability for an unanticipated workload or demand on the commvault platform. What is the impact of the business creating one hundred new workloads to protect?

Platform operations and finance teams will be able to plan for the most cost-effective resources to meet the business needs when patterns of consumption are identified.

Related documents: COST09-BP01 Perform an analysis on the workload demand.

COST09-BP02 Implement a buffer or throttle to manage demand

Commvault operates in a buffer-based mode for all activities. Commvault accepts new work to back up, restore, and replicate data and attempts to locate compute resources to service the requirement. If no resource is available, the request is queued and processed in a prioritized order as the resource is made available to the running activity.

Commvault provides **Job throttling** to perform throttling at a specific client or group of clients by the total number of parallel backup activities. Clients and client groups may also be throttled by **network bandwidth** to smooth out shared resource utilization.

Related documents: COST09-BP02 Implement a buffer or throttle to manage demand.

COST09-BP03 Supply resources dynamically

Commvault delivers backup and backup copy resources using demand-based dynamic delivery through **automatic scaling for access nodes**.

Commvault also delivers access to stored data via **power-managed MediaAgents**, which are powered down during periods of inactivity.

Consider whether heavily used peaky workloads would benefit from investment in automated instance size uplift and then downgrade during periods of high activity (see **Change the instance type**).

Related documents: COST09-BP03 Supply resources dynamically.

Optimize over time

COST10-BP01 Develop a workload review process

Commvault recommends developing a workload review process that identifies both protected workloads and Commvault data management components that drive more than 10% of the total platform bill.

Define the frequency of review to allow for early detection of issues, experimentation, and resolution before the next planning or review cycle.

Be aware that cost optimizations that reduce the retention of data may not be immediately realized for finance team cost & usage reporting. Previous backup data must age out and then be expired (deleted) from storage.

Related documents: COST10-BP01 Develop a workload review process.

COST10-BP02 Review and analyze this workload regularly

Regularly review the most expensive protected workloads and the Commvault workload components to determine if there are opportunities to optimize the volume of data transferred, the RPO/RTO for the workloads, and the underlying data management architecture (data flows, storage classes).

Related documents: COST10-BP02 Review and analyze this workload regularly.

Automating Operations

COST11-BP01 Perform automations for operations

Commvault automates your end-to-end data protection, audit & compliance, and cyber-resilience protections to reduce the effort for operations in AWS. Many of these enhancements also reduce Cloud resource consumption (i.e., runtime), helping save cost in both resource utilization and cloud operations efficiencies. Some examples include:

- **Virtual Server Agent for AWS** automates the snapshot protection, copying, sharing, and replication between AWS accounts and Regions for AWS compute, database, and storage services.
- **Threat Scan** proactively identifies malware threats to avoid reinfection during recovery. Threat Scan analyzes backup data to find encrypted or corrupted files, ensuring users recover trusted versions of their data quickly.
- **Commvault Data Platform** uses AI and automation to use historical insights to optimize backup, replication, and recovery scheduling to ensure Service Level Objectives are met. Built-in AI and machine learning identify anomalies in operational patterns that could indicate insecure or malicious activity in both production data environments and backup data.
- **Auto-scaling Access Nodes** & **Power-managed MediaAgents** automate the deployment, use, and termination of ephemeral resources or power-manage (power-down, power-up) Cloud compute in resource to data management activities.

Related documents: COST11-BP01 Perform automations for operations.

Additional Resources

- **Cost Optimization Pillar: Well-Architected**
- **AWS re:Invent 2021 - Cost control and governance at scale**
- **AWS re:Invent 2021 - Improving cost visibility and allocation**
- **AWS re:Invent 2021 - Best practices for cost optimization with Amazon S3**
- **AWS re:Invent 2021 - Optimize compute for cost and capacity**

Sustainability Pillar

The **Sustainability Pillar** focuses on integrating sustainability objectives into your cloud architecture, designs, and potential impacts and trade-offs of a more sustainable solution. Sustainability is concerned with the long-term environmental, economic, and societal impacts of your business activities and the AWS services you use to deliver those products.

Most organizations have sustainability objectives that are developed from regional commitments or broader global commitments to reduce overall carbon emissions (see **Greenhouse Gas Protocol** for the definition of carbon emissions).

Cloud sustainability

Cloud sustainability is the act of including emissions accounting in the architecture, design, and operation of cloud-based solutions. Three emission scopes are considered when accounting for the overall emission impact of a workload:

- **Scope 1:** All direct emissions of an organization
  (e.g., fuel combustion by data center backup generators).
- **Scope 2:** Indirect emissions from electricity purchased and used to power data centers and facilities
  (e.g., emissions from commercial power generation)
- **Scope 3:** All other indirect emissions from activities of an organization from sources it does not control
  (e.g., data center construction, manufacture, and transportation of IT hardware deployed in data centers).

Source: **AWS Well-Architected Framework – Sustainability Pillar – Cloud Sustainability**.

For sustainability architecture and design, your workloads and the Commvault data management platform contribute to *Scope 3: All indirect emissions* vary based on the type, size, number, and duration of AWS used by your organization.

The sustainability pillar helps you make architectural decisions and apply best practices to your Commvault data platform that will minimize the impact of your Commvault-powered data management services.

- **The shared responsibility model**
  Sustainability is a shared responsibility between customers and AWS.

  AWS is responsible for optimizing the sustainability *of the cloud* which includes AWS-operated facilities, sourcing renewal power sources, and water stewardship.

  Customers are responsible for sustainability *in the cloud* which includes optimizing workloads to reduce the number and type of resources utilized to meet your business needs. For cloud data management this includes selecting efficient compute types, powering off instances when not in use, and reducing the size of data for transmission and storage.
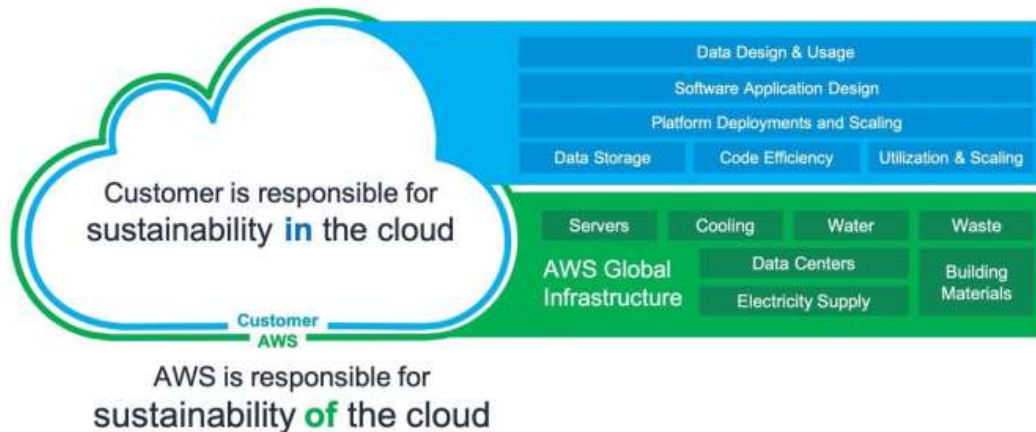
- **Sustainability through the cloud**

  AWS Cloud can be used to address broader sustainability challenges that are challenging for most businesses to address with owned or leased infrastructure

  AWS cloud compute, storage, and network resources are designed to lower indirect emission footprint through economy of scale deployment, and on-demand just-in-time use of resources vs. long-running over-provisioned resources.

  Commvault recommends migrating large-scale multi-TB and multi-PB data archives to more sustainable Amazon S3 storage classes designed for very infrequent access. Commvault **Risk Analysis** can be used to identify data to migrate to archival or delete.

  Continual data landscape analysis and optimization is possible and recommended, due to the elastic on-demand nature of Amazon EC2 compute services.

- **Design principles for sustainability in the cloud**

  Consider the following design principles when architecting your Commvault data management platform to maximize sustainability and minimize environmental impact.

  o **Understand your impact**

  Measure the impact of your overall Commvault data management platform and model impact as your AWS workload adoption increases. Be sure to consider the entire lifecycle of workload protection, from initial active workload usage to decommissioning, to eventual data destruction. Consider baselining data management activities with varying emissions profiles, then establish *key performance indicators (KPIs)* that align emissions impact to the value of data to the business.

  AWS provides the **Customer Carbon Footprint Tool** to measure and model the carbon footprint of solutions across services and geographies.

  Commvault recommends measuring the consumption of your Commvault data management resources to provide accurate modeling for your current and future platform growth.

- o **Establish sustainability goals**
  Identify and publish sustainability goals for your line-of-business teams, application owners, and IT shared services owners. Commvault consolidates the data management resources across your organization and can be used to deliver reduced compute and storage requirements at an organizational level.

- o **Maximum utilization**
  Right-size both application and Commvault data management resources to drive high utilization (greater than 60%) of compute and storage investments. Review resources monthly and reduce, remove, or power-down resources with observed utilization of less than 40%. AWS Compute Optimizer will provide insight into resource utilization of the previous four weeks and before.

- o **Anticipate and adopt new, more efficient hardware and software offerings**
  Stay aware of Commvault updates and enhancements by subscribing to **Commvault RSS feeds** and implementing new releases promptly to achieve to latest updates to improve the performance and efficiency of cloud-based data management.

  See **New Features in Commvault Platform Release 2022E**.

- o **Use managed services**
  Commvault recommends migrating your data management system to the AWS cloud for the shared sustainability impact of the common data center, compute, network, and storage resources. An immediate impact on the data center footprint, management overhead, and ongoing lifecycle costs can be achieved by moving backup and archival data to Amazon S3. Coupling storage elasticity with on-demand computing provided by Amazon EC2 allows for a self-adjusting and sustainable data management platform for your business.

- o **Reduce the downstream impact of your cloud workloads**
  Commvault continually reviews and recommends more energy-efficient resources to be performing your cloud-based data management. This includes, but is not limited to using AWS Graviton resources, adopting tunable Amazon EBS gp3 volumes, and leveraging Amazon EBS direct APIs to reduce the time to identify new and changed data to protect.

Improvement process

Improving the sustainability of your Commvault data management platform in the cloud is a continuous improvement process that will begin with the following two key focus areas:

- **Eliminate waste** by removing or right-sizing low-utilization resources.
- **Maximize the value** or efficiency of the resources that remain after waste is removed.

The following steps are recommended to be applied in a continual improvement feedback loop for all architectural, design, and operational changes throughout the life of your Commvault data management platform:

- **Identify targets for improvement**
  Review the Amazon-**published best practices** for sustainability and Commvault solution to each recommended best practice. Commvault recommends starting with AWS Compute Optimizer to review all Commvault EC2 instances for evidence of over-provisioning and right-size as recommended. After compute

resources are optimized, reviewing the frequency, number, and retention of backup copies can yield additional areas for focus.

- **Evaluate specific improvements**

  Before modifying your environment a clear understanding of the resources being consumed to deliver a specific service. Use *AWS Cost and Usage Reports* to understand where your largest cost drivers exist across compute, network, and storage, and identify areas for focus.

  Evaluating an improvement should identify the resource (compute, network, storage) and the metric to measure the impact before and after the change. You will have both technical metrics (minutes to perform backup) and business metrics (maximum acceptable minutes to perform backup). Identify the areas that have the greatest potential for improvement without impacting business SLAs.

  For example, you may want to reduce the cost of Commvault compute resources by using burstable instances. You will measure that time to perform the backup without and without burstable instances and compare it to an acceptable total elapsed backup time from your business SLA. You can use the Commvault **backup job summary report** and **restore job summary report** to measure the elapsed time of data protection.

- **Prioritize and plan improvements**

  Prioritize the areas for focus by the greatest expected impact and least risk on your organization-wide data management services. For example, choosing to utilize Amazon S3 Intelligent-Tiering instead of Amazon S3 Standard-Infrequent Access for primary backups can be tested for only a subset of workloads through a dedicated Plan and then easily measure the performance and/or cost savings for the fixed set of test workloads.

- **Test and validate improvements**

  Perform small and self-contained tests, initially in non-production or pre-productions environments, then in production with low-risk workloads (i.e., dev-test). Calculate and qualify the resource reduction and relative business impact against expected results.

- **Deploy changes to production**

  If the test fails to meet business accepted 'success' criteria, move to roll the change out to production in phased deployed (non-production, production, then business-critical workloads). Use **Blue/Green deployments** to phase workloads from one environment configuration to another. In the context of your data management system, a blue/green deployment may be the workload **plan**, **VM group**, or target **cloud storage**.

- **Measure results and replicate successes**

  Continue to measure the impact of the change or improvement as it rolls out to each grouping of production workloads. Some workloads may not experience the benefits identified in an isolated lab test, requiring roll-back and re-testing to identify the root cause for improvement failure. *This step is a feedback loop that begins the improvement cycle again with a newly identified improvement to test.*

Sustainability as a non-functional requirement

Sustainability should be included in the listed business requirements or architectural non-functional requirements when reviewing and assessing new products for introduction into your business application landscape.

Be aware the more sustainable solutions may affect the business outcome of the service being delivered, but testing should validate the actual impact by classifying the impact as one of the following:

- **Adjust the quality of the result**: Consider trading the Quality of Results (QoR) for a reduction in workload intensity. For example, a dev-test workload may only require a single Amazon EBS snapshot for recovery vs. a streaming backup copy in an EBS-independent format to meet recovery SLAs.
- **Adjust response time:** A slower response time can reduce the carbon footprint by requiring fewer processing resources to provide a timely response to demand. For example, aligning backup copy SLAs allows the reuse of ephemeral Access Node resources, while potentially delaying backup time frequency for some workloads. Utilize the **backup job summary** and **restore job summary** reports observing the impact of changes over time.
- **Adjust availability:** Evaluate the true availability needs of your Commvault data management platform. For example, CommServe Disaster Recovery can involve provisioning a new and restore from the latest DR backup or an always-running static stability approach that requires idle resources awaiting a failover event.

Best practices for sustainability

Amazon publishes a set of **best practices for sustainability in the cloud**, the following section details how these recommendations can be applied to a Commvault data management platform.

Region Selection

SUS01-BP01 Choose Regions near Amazon renewable energy projects and Regions where the grid has a published carbon intensity that is lower than other locations (or Regions)

Commvault resources may be located remote to protected workloads if longer recovery points and recovery time objectives are acceptable to the business.

Related documents: SUS01-BP01 Choose Region based on both business requirements and sustainability goals.

User behavior patterns

SUS02-BP01 Scale infrastructure with user load

Commvault recommends using **AWS Compute Optimizer** to continually right-size resources based on observed usage.

Commvault resources **automatically power down** when not in use and power up when required for backup, recovery, and replication.

Commvault **auto-scales** (deploys) backup compute resource at runtime and terminates after the job is complete.

Related documents: SUS02-BP01 Scale workload infrastructure dynamically.

SUS02-BP02 Align SLAs with sustainability goals

Commvault recommends ensuring your data protection SLAs meet business recovery and sustainability objectives, not exceed them.

Consider reduced recovery time and improved sustainability outcomes for less critical workloads.

Related documents: SUS02-BP02 Align SLAs with sustainability goals.

SUS02-BP03 Stop the creation and maintenance of unused assets

Commvault recommends using a **single data management solution** for all your data management needs, reports, and analysis.

Commvault reduces redundant and unused data, point solution systems, and reporting practices across your hybrid data landscape.

Related documents: SUS02-BP03 Stop the creation and maintenance of unused assets.

SUS02-BP04 Optimize geographic placement of workloads for user locations

Commvault recommends centralized command and control with regional data copies for timely access and recovery.

Commvault manages your backup and archive data as a distributed data store, **located closest to the business** or users that require it.

ⓘ **Note**: May conflict with SUS01-BP01.

Related documents: SUS02-BP04 Optimize geographic placement of workloads based on their networking requirements.

SUS02-BP05 Optimize team member resources for activities performed

Commvault recommends a remote management approach using SSH and/or RDP, moving processor and memory-intensive workloads to the cloud resources that can be more easily scaled.

Related documents: SUS02-BP05 Optimize team member resources for activities performed.

SUS02-BP06 Implement buffering or throttling to flatten the demand curve

You may implement **job throttles**, network **bandwidth throttles,** or concurrent threads (**readers**, **writers**) throttles to artificially constrain Commvault impact on resources to flatten the demand curve and reduce provisioned capacity to meet sustainability goals.

Related documents: SUS02-BP06 Implement buffering or throttling to flatten the demand curve.

Software and architecture patterns

SUS03-BP01 Optimize software and architecture for asynchronous and scheduled jobs

Commvault queues protection requests to best price-performance sustainable AWS Graviton compute instances.

Commvault creates **on-demand instance workers** to maximize backup throughput, then terminates when the batch workload is complete.

Commvault <u>can</u> schedule data management activities during times of day when carbon intensity is lowest using **blackout windows**.

Related documents: SUS03-BP01 Optimize software and architecture for asynchronous and scheduled jobs.

SUS03-BP02 Remove or refactor workload components with low or no use

Commvault recommends using **AWS Compute Optimizer** to continually monitor compute instances for over-provisioning.

As data access patterns change, deduplication databases (DDBs) and search indexes may be consolidated to reduce wasted resources.

Related documents: SUS03-BP02 Remove or refactor workload components with low or no use.

SUS03-BP03 Optimize areas of code that consume the most time or resources

Commvault continually monitors and tunes data management components for optimal speed and available cloud resources.

Commvault recommends Amazon Linux 2 running on AWS Graviton3 instances for the best price-performance for cloud data management.

Commvault reduces resource consumption for **storage copies**, by scheduling data copies asynchronously to initial snapshot operations.

Related documents: SUS03-BP03 Optimize areas of code that consume the most time or resources.

SUS03-BP04 Optimize impact on customer devices and equipment

Commvault utilizes **source-side deduplication** to reduce client compute and network bandwidth demand over the life of the customer device.

Computationally intense activities like re-hydrating deduplicated and compressed data are performed server-side to reduce client impact.

Commvault Command Center™ paginates and filters large data environments to limit the impact on client browsers and edge-based devices.

Related documents: SUS03-BP04 Optimize impact on devices and equipment.

SUS03-BP05 Use software patterns and architectures that best support data access and storage patterns

Commvault utilizes technologies to minimize data processing (i.e., **EBS direct API change block tracking**) and transfers and stores data in efficient deduplicated and compressed format.

Deduplication indexes are stored in an optimized on-disk database requiring only 0.002% of total data managed in Amazon S3.

Related documents: SUS03-BP05 Use software patterns and architectures that best support data access and storage patterns.

Data

SUS04-BP01 Implement a data classification policy.

Commvault implements your business data classification policy as **Plans**, which ensures data is retained, moved, and deleted per business outcomes. Data is **copied** to more energy-efficient storage before removal from frequent access locations.

Plans utilize **AWS resource tags** to identify workload data classification and automatically adjust protection as a workload changes classification. Additionally, AWS tags are both protected and recovered so workload data classification is maintained for all data copies.

Commvault also protects **untagged or unclassified** data allowing periodic auditing of unclassified resources.

Related documents: SUS04-BP01 Implement a data classification policy.

SUS04-BP02 Use technologies that support data access and storage patterns.

Commvault publishes storage selection best practices in the Cloud Architecture Guide (this document) to align the data access patterns of Commvault software with the most appropriate technology.

Commvault advocates and provides automated data migration from frequent access to infrequent access archival storage (i.e., Amazon S3 Glacier Archive / Deep Archive) using **storage copies**.

Related documents: SUS04-BP02 Use technologies that support data access and storage patterns.

SUS04-BP03 Use policies to manage the lifecycle of your datasets

Commvault manages the automated lifecycle of aggregated data through lifecycle policies called **Plans**. Plans automatically enforce the creation, deletion, and migration of data to appropriate **storage classes** or **tiers** based on user-defined policy.

Commvault does not recommend or support the use of *S3 Lifecycle Policies* to expire or delete Commvault data.

Related documents: SUS04-BP03 Use policies to manage the lifecycle of your datasets.

SUS04-BP04 Use elasticity and automation to expand block storage or file system.

Commvault can use object, file, or block-storage to store backup data. Commvault recommends using Amazon S3 object storage exclusively for all backup and archival copies due to its durability, scalability, and security. Commvault infrastructure (CommServe, MediaAgents, Access Nodes) uses a mixture of block-mode storage, and optionally file-mode storage for shared components like software caches.

Commvault recommends minimal volume and file-system sizes that sit at 50-60% utilized in normal operation, and then automatically scale capacity, throughput, and latency when service levels can no longer be met. Opting for minimally sized capacity avoids the effort of shrinking file-systems, which can be complex with block-mode storage.

Commvault recommends configuring Amazon CloudWatch alarms with automated **Amazon EventBridge** actions that resize file-system or IOPS and throughput settings when sustained performance exceeds 80% of the current thresholds. Commvault will auto-manage consumption to available disk space for working file-systems like index caches, jobResults, and temporary cache directories.

See **Automating Amazon EBS volume resizing with AWS Step Functions and AWS Systems Manager** for guidance on automating block-mode volume resizing.

Related documents: SUS04-BP04 Use elasticity and automation to expand block storage or file system.

SUS04-BP05 Remove unneeded or redundant data.

Commvault deduplicates backup and archive data at the block-level to remove unneeded redundant data being transferred or stored.

Block-level deduplication delivers greater storage reduction than file and object-level deduplication.

Commvault performs incremental and incremental forever backups that deduplicate data before transmission and storing in Amazon S3.

Related documents: SUS04-BP05 Remove unneeded or redundant data.

SUS04-BP06 Use shared file systems or object storage to access common data

Commvault places all data in shared storage to provide secure multi-tenant access to retained organizational backup and archive data.

Commvault implements **caches** to avoid access to shared storage, content in caches has automated time-to-live (TTL) to automatically expire old or unused data.

Related documents: SUS04-BP06 Use shared file systems or storage to access common data.

SUS04-BP07 Minimize data movement across networks

Commvault allows **storing data in proximity to the user** or protected workload, automatically discovers workload location, and writes to the nearest configured storage.

Commvault uses block-level duplication and compression by default for all data transfers across the network.

ⓘ **Note**: May conflict with SUS01-BP01.

Related documents: SUS04-BP07 Minimize data movement across networks.

SUS04-BP08 Back up data only when difficult to recreate

Commvault Plans allow the creation of a backup policy that enforces which data to **include and exclude** from protection operations.

Commvault Plans allow the exclusion of ephemeral, temporary, or cache files and volumes from backup protection.

Related documents: SUS04-BP08 Back up data only when difficult to recreate.

Hardware patterns

SUS05-BP01 Use the minimum amount of hardware to meet your needs

Commvault uses **automated horizontal scaling** to scale hardware only when required for backup activities.

Commvault advocates the use of AWS Compute Optimizer to right-size hardware based on observed sustained and cyclical utilization patterns.

Commvault advocates the consumption of the smallest instance size that will meet business protection. SLAs.

Related documents: SUS05-BP01 Use the minimum amount of hardware to meet your needs.

SUS05-BP02 Use instance types with the least impact

Commvault advocates the use of AWS Graviton instances exclusively for the most energy-efficient cloud data management available at the time of writing.

Commvault **instance recommendations** have been optimized to align with compute-optimized Graviton instances for high-performance reduced impact data management.

Commvault advises the use of **burstable instances** (T Class family) for workloads not requiring sustained compute performance.

Related documents: SUS05-BP02 Use instance types with the least impact.

SUS05-BP03 Use managed services

Commvault recommends migrating your application workloads from self-managed to AWS-managed instances where feasible.

Commvault protects both self-managed applications and databases running on Amazon EC2 compute, and AWS-managed alternatives like **Amazon Aurora**, **Amazon RDS**, and **many others**.

Commvault-embedded databases cannot be migrated to AWS-managed cloud databases.

Related documents: SUS05-BP03 Use managed services.

SUS05-BP04 Optimize your use of GPUs

Commvault does not utilize high-power consumption Graphics Processing Units (GPUs) in cloud data management.

Related documents: SUS05-BP04 Optimize your use of hardware-based compute accelerators.

Development and deployment process

SUS06-BP01 Adopt methods that can rapidly introduce sustainability improvements

Commvault recommends frequent baseline testing and optimization of your AWS-based data management platform.

Utilize Spot instances and minimal workload clones to validate production-like use-cases before making changes to production.

> ⓘ Note
>
> Commvault protects Amazon EC2 Spot instances but cannot restore to Spot-provisioned instances at this time.

Related documents: SUS06-BP01 Adopt methods that can rapidly introduce sustainability improvements.

SUS06-BP02 Keep your workload up-to-date.

Commvault allows **automated software download** and phased **automated upgrade** policies and schedules to keep your Commvault software updated.

The latest General Availability (GA) releases of Commvault will include the latest performance optimizations that impact overall platform sustainability.

Related documents: SUS06-BP02 Keep your workload up-to-date.

SUS06-BP03 Increase utilization of build environments.

Commvault pre-production and test environments may be deployed with AWS CloudFormation on-demand.

Commvault recommends the use of EBS-optimized, burstable, and Spot instances to perform optimized data management testing.

Related documents: SUS06-BP03 Increase utilization of build environments.

SUS06-BP04 Use managed device farms for testing.

Not applicable.

Related documents: SUS06-BP04 Use managed device farms for testing.

Additional Resources

- **Commvault Sustainability – Greenhouse Gases - Commvault**
- **Commvault 2022 Corporate Social Responsibility Report**
- **Sustainability Pillar: Well-Architected**
- **AWS Customer Carbon Footprint Tool Overview**
- **AWS re:Invent 2021 – Architecting for sustainability**
- **AWS re:Invent 2021 - Sustainability in AWS global infrastructure**
- **AWS re:Invent 2021 Breakout Sessions - Sustainability**

Reference Architectures

Cloud-native backup with Commvault Backup & Recovery

Cloud-native data protection with Commvault Backup and Recovery

This reference architecture describes how Commvault Backup and Recovery is implemented across multiple accounts to protect multiple services in an automated way.
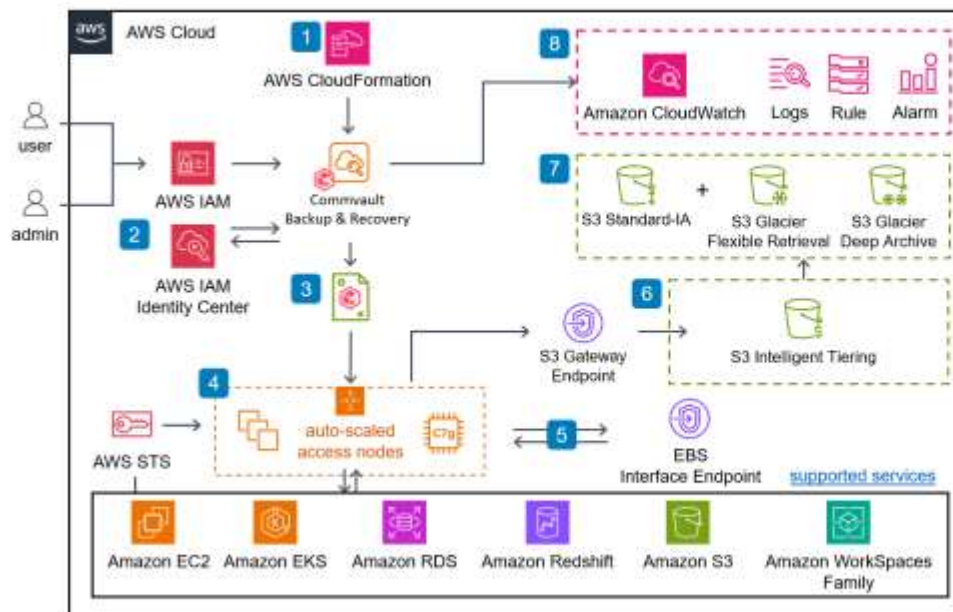
1. Use AWS CloudFormation to provision the components that Commvault uses in this architecture from AWS Marketplace.

2. Users authenticate via AWS IAM Identity Center using SAML 2.0 assertions, then Commvault authorizes using role-based access.

3. Server plans define the frequency, retention period, lifecycle, backup copy destination and resources to protect.

4. Auto-scaled access nodes perform multi account protection using STS AssumeRole temporary credentials.

5. Amazon EC2 protection uses Amazon EBS Direct APIs to read/write data via an EBS direct interface endpoint.

6. Deduplicated backups are written to Amazon S3 Intelligent Tiering bucket with Archive Instant Access Tier enabled. (Use Commvault Combined storage to use S3 Glacier Archive/Deep Archive)

7. Archival data is copied by Commvault to Commvault Combined Storage Tier bucket for optimized retrieval.

8. Amazon CloudWatch agent is pre-installed and configured to monitor and alerts on key operating system metrics.

Reviewed for technical accuracy Jul 11, 2023
© 2023, Commvault

AWS Reference Architecture

Cross-Account and Region data protection with Commvault Backup & Recovery and AWS Organizations

Cross-account and Region data protection with Commvault Backup & Recovery and AWS Organizations

This reference architecture enables customers to implement a consistent secure backup strategy using multiple AWS accounts and Regions, and copies backups between them using Commvault Backup & Recovery automated and policy-driven data management.
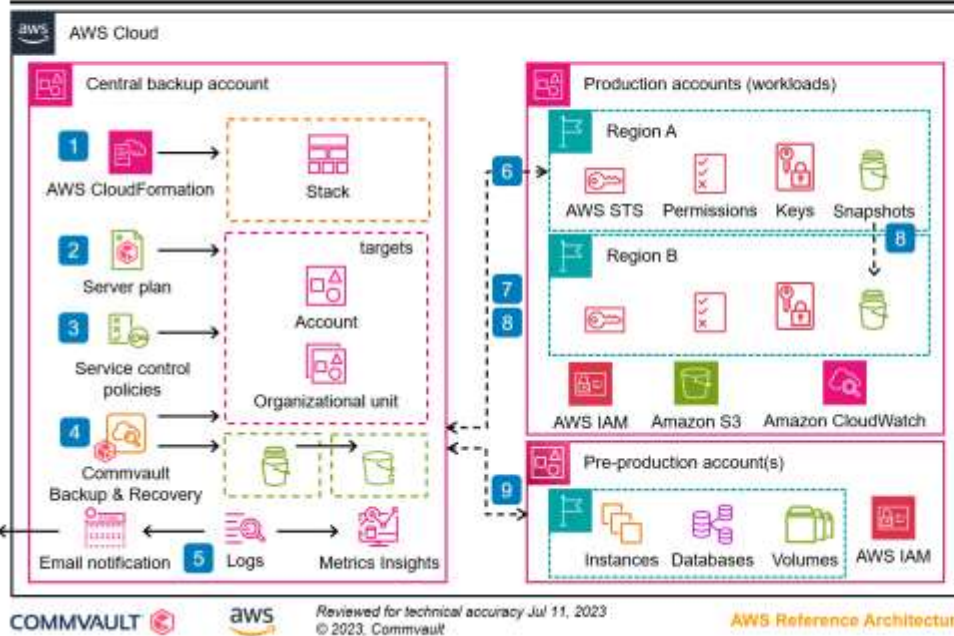
1. Use AWS CloudFormation to provision Commvault resources, including shared **Amazon EC2**, **AWS KMS** and **Amazon S3** buckets (per region).

2. Create server plans that define the frequency, retention, lifecycle, backup copy settings to protect AWS workloads.

3. AWS Organizations SCPs and Tag policies enforce mandatory tags per account to grant identity trust and permissions, and auto-protect workloads.

4. Cross-account and regional Recovery Readiness can be monitored through SAML 2.0 SSO authenticated **Commvault Command Center™** console.

5. Commvault and workload logs are centralized in the management account for **Amazon SNS** email notification and anomaly detection

6. Create each workload account in Commvault with **STS AssumeRole** trust to management (admin) account and a VM group to enable auto-discovery and protection of workload resources.

7. Auto-scaled Access Nodes use STS AssumeRole temporary credentials to perform backup jobs, recovery points are copied to management account.

8. Cross-account and cross-region backups will re-encrypt backup copies with KMS key from the target account and region.

9. Software update deployment testing can validate changes in canary and blue/green deployments to dedicated access nodes.

AWS Control Tower and Commvault Backup and Recovery

AWS Control Tower and Commvault Backup & Recovery

This reference architecture describes how Commvault Backup & Recovery is implemented in multi-account landing zones provisioned by AWS Control Tower to provide a well-architected multi-account AWS environment with automated data protection.
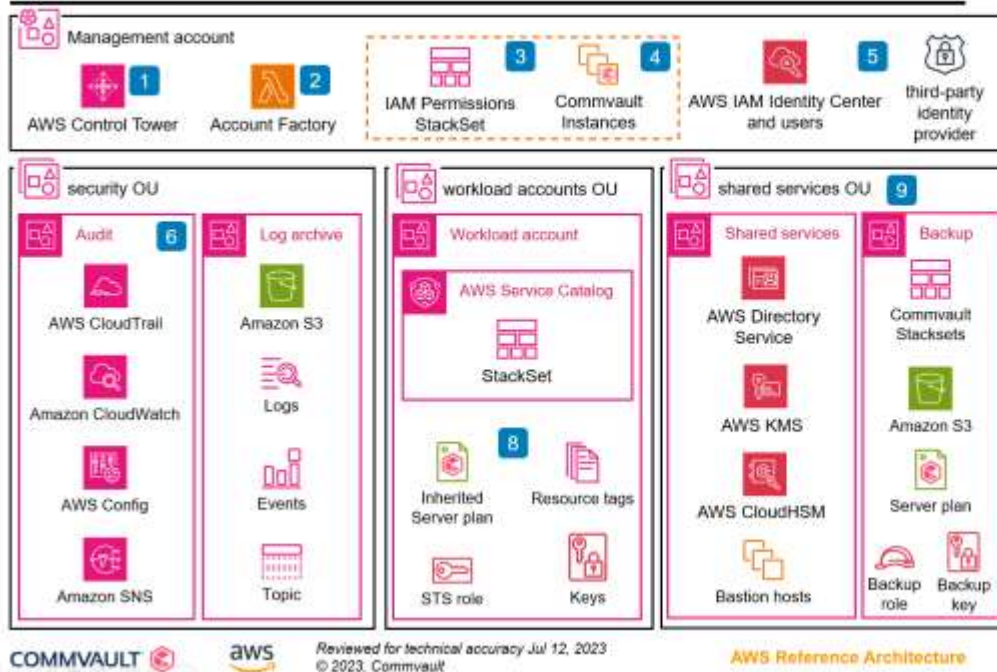
1. Use AWS Control Tower to deploy a multi-account landing zone with a central shared services OU + backup account.

2. Enable Account Factory to automate provisioning of workload accounts and applying controls or guardrails (service control policy).

3. Enable **AWS CloudFormation StackSets** for your Organization to centrally deploy Commvault resources.

4. Enable **Commvault Backup & Recovery** for protection of all resources created across your Organization.

5. Centrally manage SSO access to your environment using **AWS IAM Identity Center**, SAML 2.0 and optional third-party identity providers (IdPs).

6. AWS Control Tower provisions an **Audit** account to provide cross-account auditing capability.

7. AWS Control Tower provisions an **Log archive account** to provide logs of all API activities and resource configs (back up your logs).

8. Workload accounts will provision **STS roles**, **AWS KMS keys** and **resource tags** to automate per-account back up.

9. **Central backup account** owns shared Commvault infrastructure, backups, admin role and server plans. Access granted via centralized identity and optional bastion hosts.

257

Creating immutable backups with Amazon S3 Object Lock and Commvault Backup & Recovery

Creating immutable backups with Amazon S3 Object Lock and Commvault Backup & Recovery

This reference architecture describes how Commvault Backup & Recovery is implemented with Amazon S3 Object Lock to provide an immutable data bunker that follows the principles of least privilege in a multi-account AWS Organization.
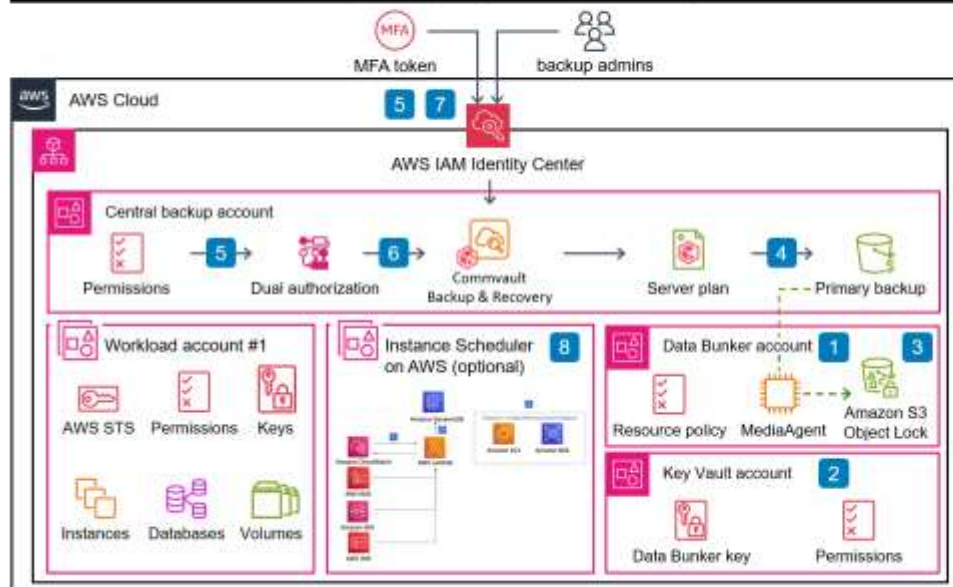
1. Create a Commvault Data Bunker account, and an Amazon S3 bucket. Create a S3 resource policy that limits Put / Delete actions to the bunker AWS account, MediaAgent IP, and optional date/time.

2. Create a Customer Managed KMS key in a separate Key Vault account, and share the KMS key to the Data Bunker for encrypt, decrypt of data at rest. Enable controls requiring MFA for critical KMS APIs.

3. Create Commvault Data Bunker cloud storage with Amazon S3 Object Lock enabled, accessible only by the Data Bunker MediaAgent with attached machine identity using STS temporary credentials.

4. Add a Storage Copy to existing Server Plans to perform backup copy operations to the Data Bunker account in accordance with business policy.

5. Restrict the access to the Commvault Data Bunker credential and configuration, and restores from the bunker to a limited users via Commvault SSO, MFA, and Roles.

6. Enable dual-authorization approval for high-risk configuration changes to the bunker plans, copies, or cloud storage.

7. Restrict the access to the Data Bunker account to specific admin users via AWS SSO and MFA, that follows a Break Glass Workflow for temporary access.

8. (Optional) Use Instance Scheduler on AWS from a dedicated account to power-down and power-up the Data Bunker MediaAgent to provide a logical airgap

COMMVAULT | aws | Reviewed for technical accuracy Jul 13, 2023 © 2023, Commvault | AWS Reference Architecture

Ransomware protection with Commvault Backup & Recovery

Ransomware mitigation with Commvault Backup & Recovery

This reference architecture describes how to protect and recover your AWS resources from ransomware and malware attacks using Commvault Backup & Recovery automated data protection with preemptive monitoring, alarming, and automation action.
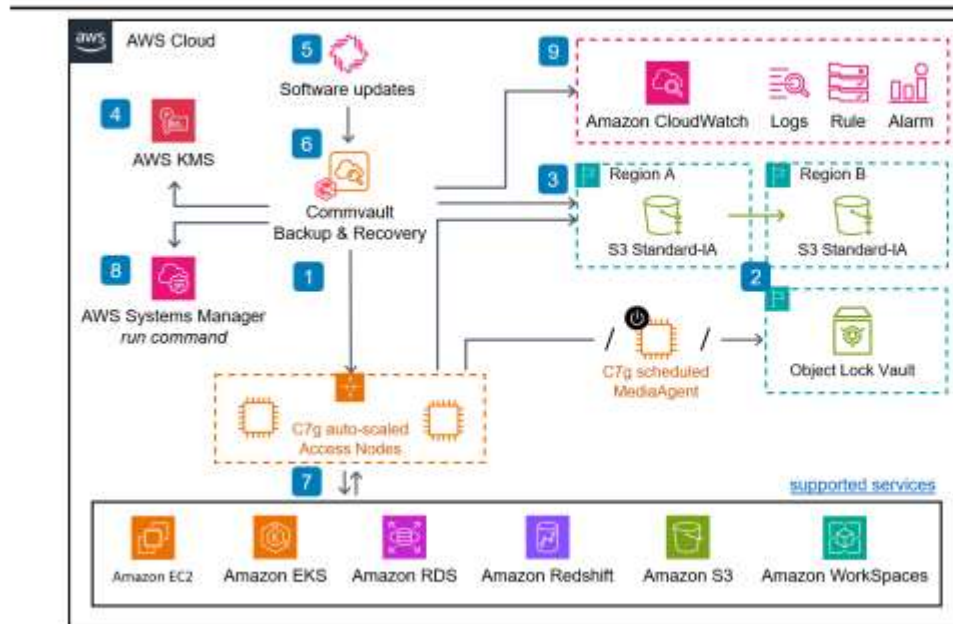
1. Create cross-account backups to allow recovery and prevent bad actor deletion.

2. Create an air gapped S3 Object Lock'ed offline backup copy in a Data Bunker accessible only by a Instance scheduled MediaAgent.

2. Replicate backup copies to another availability zone or Region for Disaster Recovery. Recover on-failure or perform periodic replication and failover.

3. Ensure Commvault DR backups are exported to a dedicated DR account / Amazon S3 bucket for Disaster Recovery

4. Encrypt everything. Workloads, primary backups, and offline backup copies to protect against data exfiltration threats.

5. Enable automatic software update download and deployment, to receive fixes from common vulnerability exposures (CVEs) and bug fixes.

6. Harden your Commvault CommServe with CIS Level 1 benchmarks for additional controls for your central backup instances.

7. Enable ransomware detection on workload instances using honeypot, file anomalies, and file encryption activity monitoring.

8. Automate anomaly response with Amazon Systems Manager and Alert notifications. (e.g. disable network access on infection)

9. Enable the File Activity Anomaly Alert and forward into your SecOps teams via email, SNS, or Amazon CloudWatch logs.

COMMVAULT | aws | Reviewed for technical accuracy Jul 13, 2023 © 2023, Commvault | AWS Reference Architecture

Hybrid backup with Commvault Backup & Recovery and AWS Outposts

Hybrid backup with Commvault Backup & Recovery and AWS Outposts

This reference architecture describes how to protect and recover your AWS Region, AWS Outposts, and corporate data center workloads using Commvault Backup & Recovery automated data protection.
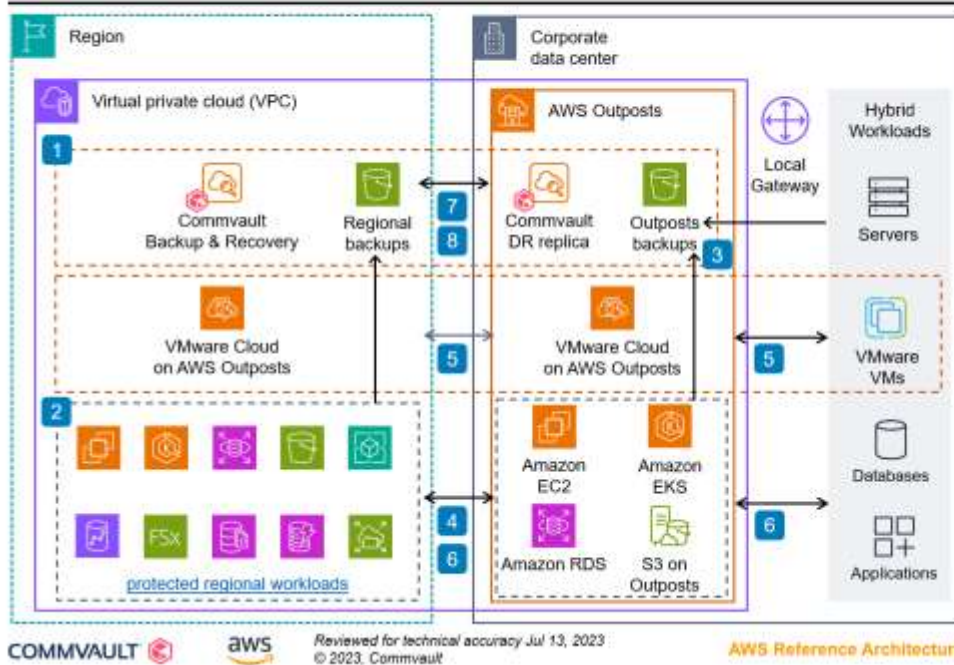
1. Extend your existing VPC with one or more AWS Outposts racks, which appear as a special availability zone in your Region.

2. Protect your AWS Region workloads using native snapshots and service independent backup copies to regional Amazon S3.

3. Protect your AWS Outposts and data center workloads to Amazon S3 on Outposts or snapshots stored in the Region (Amazon EC2, Amazon RDS).

4. Protect, replicate and migrate Amazon EC2, RDS, EKS, S3, and VMware resources between the Region and Outposts.

5. Protect, replicate, migrate and perform Disaster Recovery of VMware VMs from Outposts / Data Center to/from the Region.

6. Perform Disaster Recovery replication for traditional databases, file-systems, big data, and S3 Object storage

7. Replicate backups to the Region for off-site DR copies or on-demand disaster recovery.

8. Enhance Commvault with HA/DR replication to failover the CommServe Server between the Region and Outposts.

Reviewed for technical accuracy Jul 13, 2023
© 2023, Commvault

AWS Reference Architecture

Near real-time Auto Recovery to AWS

Near real-time Auto Recovery to AWS (formerly Disaster Recovery)

This reference architecture describes how to perform on-demand disaster recovery for your AWS workloads between Regions and edge-locations using Commvault Backup & Recovery and Auto Recovery intelligent automation.
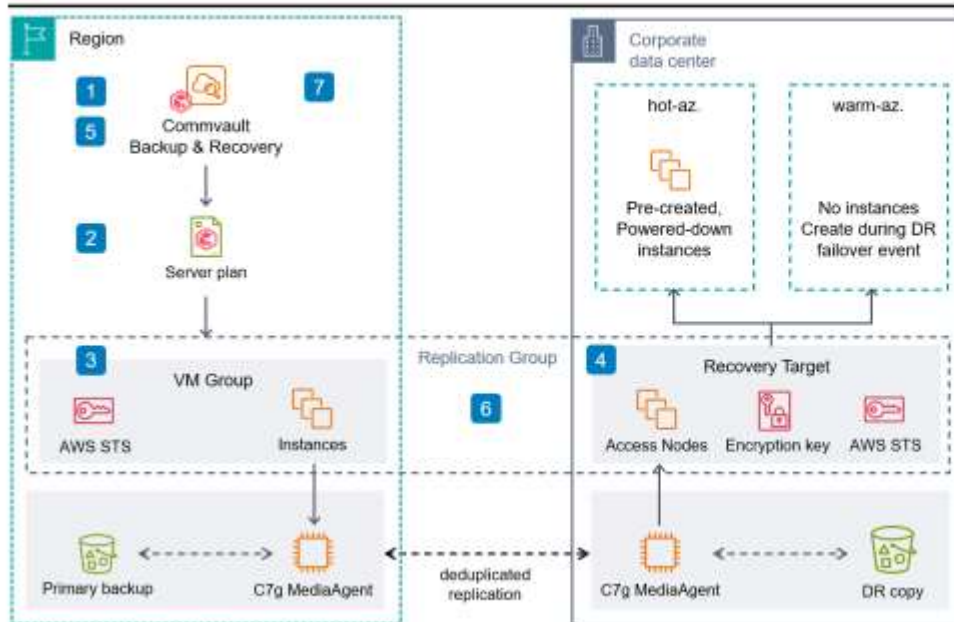
1. Ensure a reliable highly available CommServe® Server exists to coordinate backup, replication, and manual or automated failover (optimal passive replica)

2. Configure a server plan to backup your mission-critical instances to a regional Amazon S3 bucket and periodically copy to an Amazon S3 bucket in your DR region

3. Configure your Production AWS Region credentials and a VM group to select instances to protect

4. Create a recovery target to supply your DR AWS Region credentials, access nodes, and encryption key to encrypt new instances

4. Create a replication group to configure replication method, frequency, and whether to create DR instances (hot site) or delay until failover (warm site)

5. Monitor replication status using the Periodic Replication Monitor. Alerts will be generated if replication falls behind or fails.

6. Perform Disaster Recovery Testing including planned failover/failback, unplanned failover, failback, and test failover (bubble network)

7. Optionally configure cross-platform disaster recovery from VMware VMs directly into Amazon EC2 instance, or databases, file storage, object storage, and Hadoop DR replication.

Reviewed for technical accuracy Jul 13, 2023
© 2023, Commvault

AWS Reference Architecture

259

Scaling from day-one to multi-region Commvault Backup & Recovery

Scaling from day-one to multi-region Commvault Backup & Recovery

This reference architecture describes how to deploy an initial day-one Commvault Backup & Recovery instance and scale to supporting multi-region multi-account protection.
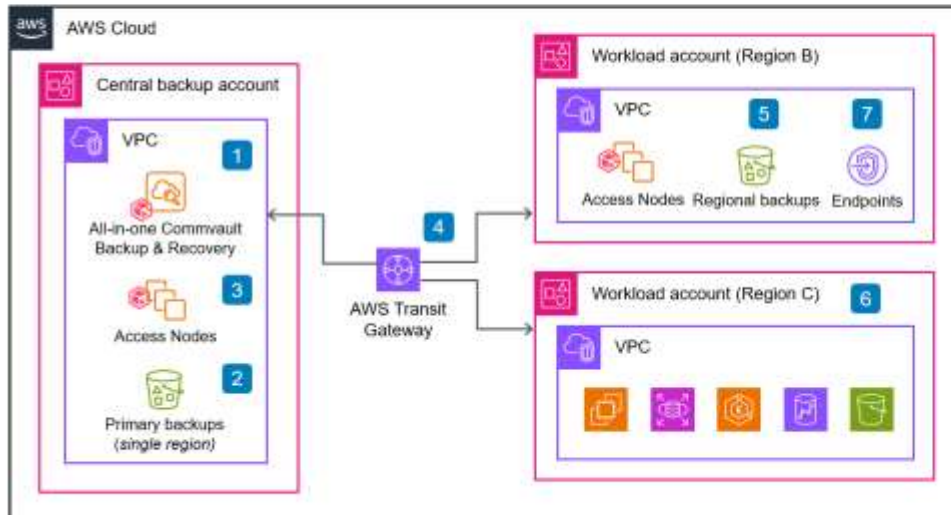
1. Begin day-one with an **all-in-one** Commvault CommServe® instance with MediaAgent and Access Nodes roles co-located on a single instance, deployed from AWS Marketplace.

2. Write in-region backups to a regional Amazon S3 bucket via the MediaAgent role on the **all-in-one** CommServe® instance.

3. When high-availability is required for the MediaAgent function, or all-in-one instance can no longer meet recovery point / recovery time objectives, scale data movement onto separate Access Node/MediaAgent grids (1-4 nodes)

4. Connect your central backup account VPC and workload VPCs using AWS PrivateLink, VPC Peering, or AWS Transit Gateway

5. Each additional **Region expansion** will require (at minimum) one MediaAgent + Access Node instance and an in-region backup storage location (owned by central backup account)

6. Each protected **Workload account** will be protected by regional shared backup resources, accessed by AWS Transit Gateway.

7. (Optionally) Regional expansions can consolidate **AWS Privatelink VPC Endpoints** in the central backup VPC.

Reviewed for technical accuracy Jul 13, 2023
© 2023, Commvault

AWS Reference Architecture

Protect on-premises and Cloud edge with Commvault Backup & Recovery

Protecting on-premises and Cloud edge with Commvault

This reference architecture describes how Commvault Backup and Recovery protects workloads located in hybrid cloud (cloud edge), and traditional owned, operated or leased data-centers.

1. Commvault CommServe (command and control) may be located in the AWS Region, On-premises and/or replicated for Disaster Recovery.

2. If cost-optimized rapid recovery is required for on-premises workloads, one or more MediaAgents and/or HyperScale X nodes are used for **Primary backup** copy.

3. Commvault can replicate some or all backups (synchronously, asynchronously) to Amazon S3 for offsite protection in **Secondary backup** copies.

4. Commvault intelligently selects the nearest backup during restore to minimize restore type and network impact.

5. Restores from the Secondary Copy may be performed if the Primary Copy is damaged or lost.

6. Alternatively, in the event of a total site loss, Amazon Snow Family devices may be used to seed a new primary copy (offline transfer).

Tip: Keep a short-term retention data copy on-premises to accelerate recovery and avoid egress restore fees.

Tip: Place MediaAgents in the AWS Region with their own Deduplication DataBase (DDB) to allow backup direct to Region in event of onsite storage failure or site footprint reduction.

Reviewed for technical accuracy Feb 17, 2023
© 2023, Commvault

AWS Reference Architecture

260

Ransomware protection

Ransomware and malware protection are key considerations in any cloud-based architecture. Commvault has the following **ransomware protection** capabilities that should be layered into your layered data protection plan.

Hardening your CommServe

Considered hardening your CommServe® instance by restricted access to authorized users, authorized hosts, and authorized ports.

Access to the centralized configuration MS SQL Server database should also be limited, see **Securing the CommServe Database**.

Account segregation

Utilize **multiple AWS accounts** to segregate users, departments, and backup copies. Commvault supports the protection and storage of data across multiple accounts, regions, and availability zones.

Air-gapped backup copies

An air gap back copy is a specialized type of backup copy that provides additional protection over the traditional day-to-day copy. An **air gap** is defined by the NIST Computer Security Resource Center (CSRC) as:

"An interface between two systems at which (a) they are not connected physically and (b) any logical connection is not automated (i.e., data is transferred through the interface only manually, under human control)."

Source(s): **CNSSI 4009-2015** from **IETF RFC 4949 Ver 2**

- Consider **air-gapping** at least <u>one copy</u> of your critical data by powering down the MediaAgent that provides access to the Data. This approach ensures an effective response to an unplanned and uncontrolled **ransomware propagation event**.
    - o Commvault **Cloud MediaAgent Power Management** can be used to automate the power-down/power-up of the MediaAgent.
    - o (optional) Consider utilizing **Amazon EC2 Scheduler** (**AWS Solutions**) to schedule the power-up/power-down of air-gapped MediaAgents.
    - o A *powered-down MediaAgent* is considered an 'air-gapped' copy from your primary operational recovery infrastructure.
- Always utilize **Amazon S3** cloud storage with dedicated credentials and authorized MediaAgents
    - o The use of *object-based storage* without end-user access is considered another **airgap** for your data.
    - o Traditional storage mediums (disk-based storage) can be directly accessed by malware and infected if security credentials are discovered/breached.

See **Offline Backup Copies** for additional information.

WORM backup copies

Consider utilizing **write-once-read-many (WORM)** storage mechanisms provided by AWS (S3 Object Lock, S3 Glacier Vault Lock, SnapLock on Amazon FSx for NetApp ONTAP) to prevent **ransomware propagation** into backup stores. Commvault provides the **Enable WORM on Cloud Storage** and the **Enable WORM on Disk Storage** workflow to configure cloud and disk libraries to Commvault's best practices in one easy step.

See the following information for more details:

- **Configuring the WORM Storage Mode on Cloud Storage**
- **Configuring NetApp® SnapLock® Compliance and SnapLock Enterprise on Amazon FSx for NetApp ONTAP (SnapLock)**
- **Setting Up Amazon Vault Lock for Amazon Glacier**
- **Workflow for Configuring WORM Storage Mode on Cloud Storage** (Amazon S3 Object Lock)

When enabling **Amazon S3 Object Lock**, Commvault makes the following changes:

- Enables Amazon S3 **Compliance Mode**
- Configures a default **retention period** to 2x the data retention period set on the Server Plan (or Storage Policy)
- Enables **default retention mode** (aka legal holds) on the bucket, for newly written objects.
- Configures **periodic DDB sealing** to match the Server Plan (or Storage policy).
- Micro-pruning is disabled on the Cloud Library

Example:

For a cloud library that has a configured Server plan with **180 days of retention**

The default retention period will be 360 days.

The DDB will be sealed every 180 days.

> ⓘ **Note**
>
> WORM should ideally be enabled on the backup library before writing any backup data, to ensure all data is immutable from initial creation.

Effects of DDB sealing

S3 best practices

When the DDB is sealed, the sealing process closes the DDB and starts a new DDB. When the new database is started, the next instance of each data block processed creates a **new signature tracking entry** and the data block is written to the disk again as a new initial baseline.

This will result in a **full copy** of the backup content being resent to the Cloud Library. This is <u>intentional</u> and provides multiple, segregated data copies to protect from corruption or other unforeseen data access issues.

See **Sealing the Deduplication Database**.

Detecting Ransomware activity

Ransomware section

Ransomware is continually evolving and changing its methods of infiltration and infection. One constant exists with all forms of known ransomware and malware, *they modify your critical data*. Commvault as your centralized data management system can observe file I/O that may represent a ransomware event, alert you, and take automated action. See:

- **Monitoring File Anomalies On Client Computers** (Honeypots, File-system changes)
- **Monitoring File Anomalies on the CommServe Computer** (Changes in normal activity – failed, pending, succeeded jobs, Job runtime, Events)

Protecting Mount Paths from Ransomware

In environments where disk-based backup copies are still held (e.g., edge-based locations, owned/operated data centers) preventing malware from accessing your **backup library mount paths** is critical. Commvault includes protection that prevents malware from writing to or manipulating your critical backup data. See:

- **Ransomware Protection for Disk Libraries on a Windows MediaAgent**
- **Ransomware Protection for Disk Libraries on a Linux MediaAgent**

Massively Parallel Recovery

Ransomware site-wide or environment-wide recovery

Be sure to consider the impact of a ransomware event on your organization. Having a copy of your data is problem #1, next you will want to recover *FAST.* Commvault optimizes recovery events by allowing multiple Cloud Access Nodes (MediaAgent, Access Node) to read your backup data in parallel. This information can be **temporary**, existing only for the period of the recovery event.
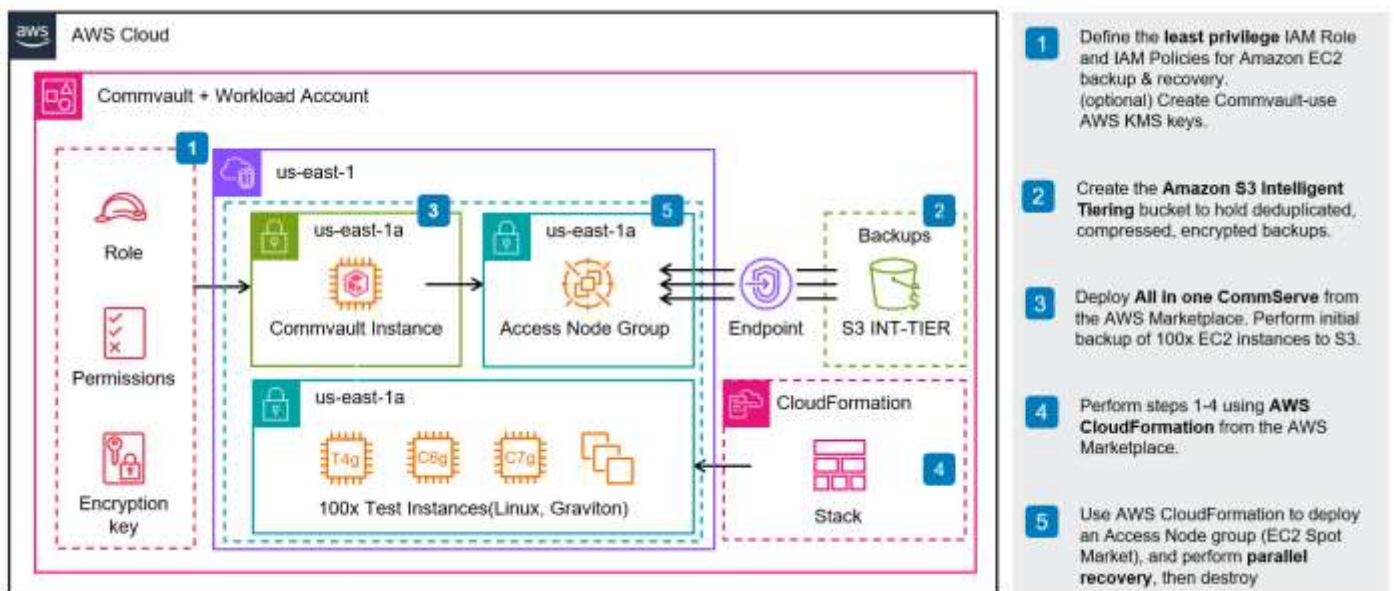
For more information, see **Ransomware protection**.

Commvault provides the AWS CloudFormation and HashiCorp Terraform to build, test, and productize massively parallel recovery in your own environment – see **GitHub – Commvault Backup & Recovery – Massively Parallel Recovery Solution**.

Additional resources

- **AWS Security Blog**.
- **Ransomware mitigation: Top 5 protections and recovery preparation actions**.
- **AWS Cloud Security**.
- **Financial Services Industry Lens – Protecting against ransomware and malware**.
- **Ransomware Risk Management on AWS Using the NIST Cyber Security Framework (CSF)**.

Design and best practices

In this section, we provide design principles and architecture principles that have been employed within the Commvault® platform to provide an optimal cloud experience for organizations planning to leverage the cloud as part of their data protection and management strategy.

Principles

As you design and build your single-region and multi-region distributed data management platform, you will be faced with several decisions. You will be deciding which workloads are protected, where to store backups, and how to optimize for cost and recovery objectives. Commvault recommends using the following **Design Principles** when making design decisions for your AWS and edge-based workload protection.

Right-sizing Over Forecasting

Commvault recommends that traditional infrastructure t-shirt sizing and forward-looking capacity provision practices are discarded in AWS. Always select the smallest recommended compute instance and scale to meet your business service-levels. Use **AWS Compute Optimizer** to receive recommendations when resources are being exhausted and review your achieved service-levels before increasing resource sizing.

Match protection to business impact

Use a *business-value data lifecycle* approach for backups that store backups using AWS snapshots, then service-independent copies in Amazon S3/S3-IA, and finally in **Amazon S3 Glacier storage classes**. Use your *data classification policy* to direct backups, based on the data value, risk, and impact if unavailable. Applications rarely require a single storage technology across their lifetime. Leverage technology enhancements like **S3 Batch Operations** to gain **faster restores from S3 Glacier archives**, without additional spend for expedited recovery.

Centrally managed AWS recovery points

Commvault recommends using a **single unified data management platform** to perform AWS snapshot creation, replication, and deletion following your business policy. Using point-solutions and scripted solutions can lead to *dark data* and *unmanaged recovery points* that drive unforeseen storage costs across your AWS resources.

Optimize At Rest

Ensure that all copies of data are stored in a least-cost optimized format to reduce storage fees and the overall sustainability of your data management landscape. As your recovery needs no longer require rapid-recovery AWS snapshots, take service-independent backup copies to Amazon S3. Use Commvault deduplication, compression, and encryption to optimize your Amazon S3-based backup copies,

Optimize On Wire

Ensure that replication of backup data for cross-region recovery services or disaster recovery occurs in an optimized format. Use incremental AWS-native snapshots (where supported) and Commvault deduplication enhanced replication to reduce the network transfer costs and transfer time. Consider if all data requires replication, a *selective replication* approach reduces the required network bandwidth as the data footprint grows.

Separation of Duty

Commvault recommends leveraging **fine-grained role-based access control (RBAC)** to provide users, admins, and business analysts the least privilege rights to protect, recover and report on business-wide data management.

Encrypt Everything

Encrypt your workloads, encrypt your workload backups stored in AWS-native snapshots and Commvault independently encrypted Amazon S3 buckets. Encryption can protect you from unintended data leakage.

Automate Over Runbooks

Automate operations to scale with speed and remove human error from daily operations. Commvault API, SDK, and CLI allow integration with Amazon CloudWatch, EventBridge, and Systems Manager to automate operations where required.

Guides

The following section guides planning and implementing Commvault data management in AWS with a focus on best practice guidance and technical limitations to consider during design.

Backup and recovery approaches on AWS and beyond

This guide describes a high-level process for assessing and implementing a consistent data protection and data management approach for your AWS and edge-based workloads.

Why data-protection?

The goal of *data protection* is two-fold and applies to AWS and edge-based (on-premises) workloads.

- Creating and storing copies of important business data to protect from data loss.
- Recover lost or deleted data to resume business services or comply with regulatory requests.

Hybrid protection

The availability of elastic compute, network and storage resources has changed how business services are delivered, and how *data protection* services can be designed and implemented. AWS and Commvault provide several services and technologies to protect your workloads running in the AWS Region, in edge locations like AWS Local Zones, AWS Outposts, and your traditional data center.

The approach does not differ based on location:

- Protect *all data* that is required to provide business services or regulatory compliance.
- Locate data for *required recovery performance*, which will differ in approach based on services available.
- Replicate data for *disaster recovery* when a device, site, or region experiences a systemic failure or outage.

Establish your business RPOs and RTOs

Your *data protection* approach starts with a definition of your business objectives for recovery. These are defined as:

- **Recovery Time Objective (RTO)** is defined by the organization. RTO is the maximum acceptable delay between the interruption of service and the restoration of service. This determines what is considered an acceptable time window when service is unavailable.
- **Recovery Point Objective (RPO)** is defined by the organization. RPO is the maximum acceptable amount of time since the last data recovery point. This determines what is considered an acceptable loss of data between the last recovery point and the interruption of service.

Disaster Recovery (DR) objectives.

You may choose to implement two tiers of recovery objectives:

- **Operational recovery objectives** refer to the RTO and RPO for recovering services back to the original Region or site.
- **Disaster Recovery objectives** refer to the RTO and RPO for recovery services to an alternate Region, Availability Zone, or site.

The costs and complexity of implementing cross-region failover and failback are often higher than cross-availability zone architectures. Consider reviewing and agreeing on business requirements with your cross-functional business leaders, before beginning an end-to-end data protection architecture.

Secure self-service access and restore

Start with the *recovery of business services* in mind when architecting and designing your data management solution. A key limitation in traditional implementations was the requirement for application owners to contact the IT helpdesk and wait for a recovery to occur.

Commvault Command Center™ console, API, command-line, and SDK empowers application owners and line-of-business owners to self-service their backup and recovery needs through a web-based interface. Access is granted securely using integration with organizational single sign-on (SSO), multi-factor authentication, and rich role-based access controls. Application owners can restore directly back to their existing AWS resources or opt to recreate and replace malfunctioning resources.

Optimize for elastic cloud resource

Leverage the elastic nature of AWS compute to minimize the amount of infrastructure used to perform hybrid data management. Commvault components should be deployed at a minimal size, then scaled when business recovery objectives can no longer be met.

Backup infrastructure approaches range from:

- **Zero cloud infrastructure** by performing the backup activity and cloud orchestration from on-premises (Commvault does not require infrastructure in the AWS region to write backups to Amazon S3) or reading data directly from cloud snapshot copies (see **Restoring Guest Files and Folders for Amazon**)
- **Power Managed infrastructure** that is powered on only when a data management activity is performed (see **Cloud MediaAgent Power Management**).
- **Ephemeral infrastructure** exists only during the backup operation and is automatically shut down and then terminated if no longer in use (see **Automatic Scaling for Amazon Access Nodes**).

The following table summarizes Commvault backup infrastructure options when deployed in the AWS cloud (region, Local Zones, and AWS Outposts), and the **recommended** approach in **raspberry**.

Commvault Component	Long-running	Power-managed	Ephemeral
CommServe® instance	✓		
MediaAgents		✓	
Access Nodes for Backup		✓	✓
Access Nodes for Restore		✓	

Protection approaches

The introduction of AWS global services to your application landscape has also introduced enterprise-grade protection technologies often available only to the largest organization. Protection in AWS and your edge-based locations should use a mix of:

- **API-based protection** by integrating natively with a service API to discover, protect, and recover data.
- **Snapshot protection** for rapid creation, replication, and rapid recovery of services using native snapshots that do not need to move data between your production application location and the backup location.
- **Streaming protection** for a service-independent or vendor-agnostic copy of your data that is copied onto Commvault-optimized storage for recovery, replication, and regulatory compliance. Streaming protection requires streaming application data from the production application location to Commvault storage used compute instances.

AWS service/product	API	Snapshot	Streaming	Comments
Amazon Aurora		✓	✓	**Export backup** is used for streaming copy.
Amazon DocumentDB		✓		
Amazon DynamoDB	✓			
Amazon EBS (inc. Outposts)	✓	✓	✓	Amazon EBS snapshots are stored in the AWS Region.
Amazon EC2 Amazon EC2 (inc. Outposts) VMware Cloud on AWS		✓	✓	Amazon EC2 on Outposts snapshots are stored in the AWS Region.
Amazon EFS			✓	Protection via NFSv4l. No snapshot API is available for Amazon EFS.
Amazon EKS (inc. Outposts) Red Hat OpenShift on AWS			✓	Cloud-native protection via Kubernetes API server (kube-apiserver).
Amazon FSx (Windows, Lustre, ONTAP)			✓	Protection via CIFS/SMB 2.0.
Amazon Redshift		✓		
Amazon RDS (inc. Outposts)		✓	✓	**Export backup** is used for streaming copy. Amazon RDS snapshots are stored in the AWS Region.
Amazon S3 (inc. Outposts)	✓			Direct API request for objects.
AWS WorkSpaces			✓	Protection via an agent installed with each Amazon WorkSpaces instance.

Backup consistency

Crash consistency refers to backups or *recovery points* that are taken without coordinating with the operating system or application writing the data. Crash consistency is often sufficient for simple file-based datasets and cloud-native applications (i.e., object storage) but may not be appropriate for traditional applications such as Microsoft SQL Server or Oracle Database. Database instances need to be quiesced to ensure the database is valid at the time of backup, and recoverable when required.

Commvault® software supports both crash-consistent and application-consistent backups, providing flexibility in your design while assuring application recoverability. Not only are the most common types of applications covered, but a wide variety of classic applications and cloud applications are supported. For a complete list of protected applications please review the online documentation: **Backup and Restore Agents**.

Identifying which data to protect

Not all workloads within the cloud need protection – for example, with micro-services architectures, or any architecture that involves worker nodes that write out the valued data to an alternate location, there is no value in protecting the worker nodes. Instead, the protection of the gold images and the output of those nodes provides the best value for the business. However, it is important to note that data stored in ephemeral locations may need to be protected before termination operations against those instances to ensure that any valuable data is not lost.

Commvault has broad support for persistent data stores outside of traditional block-based storage including Amazon EBS volumes, Amazon S3, Amazon EFS, Amazon FSx, and AWS cloud database protection.

Review your workloads to understand what data is generated and whether it is required to recover services if lost.

Always consider recovery time

The primary purpose of your *data protection* platform is to recover data that is lost or required by a regulatory request, following the business-established recovery objectives.

When selecting protection approaches and data locations, always validate you will be able to meet the business recovery objectives. Remember as data grows, your ability to restore in a reduced timeframe will become more challenging.

Likewise, as you expand from the AWS Region out to your edge processing facilities, consider the impact of network transfer times on your recovery times. When utilizing edge-based optimized infrastructure like AWS Outposts, be sure to locate a backup copy locally and remotely. This method will provide rapid recovery from localized events but also allow a region-based recovery should the edge location between unreachable. A good approach is to create low-RTO service-specific snapshots, then service-independent *backup copies* stored in-zone, then in remote-zones.

Selecting storage for performance and cost

Amazon S3 and Amazon S3 Glacier provide a broad selection of frequent access and infrequent access storage classes to meet your unique backup and archive storage needs. Be sure to consider the *first byte latency* of the storage class you are selecting, to ensure that recovery time objectives can be met with your chosen storage class.

Commvault stores your data in Amazon S3 in an optimized, deduplicated, and compressed format which reduces backup storage and replication costs. This approach creates multiple logical containers within your Amazon S3 storage and prevents the use of **Amazon S3 Storage Lifecycle policies** that perform transition or expiration actions.

Commvault manages your backup data as distinct copies that are **selectively copied** to lower storage classes as they age.

Commvault also provides **Commvault Combined Storage Tier** as an alternative to using Amazon S3 Glacier asynchronous access storage classes directly. By using Commvault Combined Storage Tier, Commvault can locate your archive files, restore your archived objects from S3 Glacier, then automatically restore your workload.

Finally, consider options to optimize your restore time from archival storage classes (i.e., S3 Glacier Flexible Retrieval, S3 Glacier Deep Archive, and S3 Intelligent-Tiering Archive Access and Deep Archive Access). You can leverage **faster restores** with the STANDARD retrieval tier by using S3 Batch Operations, or invest in using the EXPEDITED retrieval tier and **provisioned capacity** for planned/one-time data recall operations.

Auto-discovery and protection at scale

Commvault recommends using your **tagging strategy** to inform Commvault of the workloads that you need to be protected and what *data classification* or business value applies to each workload. Commvault uses *auto-discovery rules* dynamically discover new workloads at backup runtime for the following AWS resources:

- Amazon EC2
- Amazon RDS
- Amazon Redshift
- Amazon DynamoDB

- Amazon DocumentDB
- VMware Cloud™ on AWS
- Amazon EKS
- Amazon EKS Anywhere, EKS-D

- Red Hat OpenShift on AWS
- EC2 on AWS Outposts
- RDS on AWS Outposts
- EKS on AWS Outposts

Additionally, Commvault can **auto-detect** applications running inside your Amazon EC2 agents and push the software required to achieve application consistency automatically.

Auto-detection and auto-protection approaches remove the requirement for a backup or cloud administrator to continually update data protection configuration to protect newly created workloads. This results in improving your operational recovery excellence, improving resiliency within your cloud infrastructure, and ensuring new data is protected thereby ensuring your *data protection* Service Level Agreements (SLAs) are maintained.

Programmatic Data Management

Commvault® software provides a robust **Application Programming Interface (API)** that allows for automated control over deployment, configuration, and backup and restore activities within the solution.

Whether you are designing a continuous delivery model that requires the automated deployment of applications or automating the refresh of a disaster recovery copy, data warehouse, or development/testing environment that leverages data from a protection copy, Commvault® software provides the controls necessary to reduce administrative overhead and integrate with your toolset of choice.

Beyond API access, the most common use cases for data protection and management are built into the Commvault user interface. Simply enter the Cloud credentials and necessary permission and the Commvault platform will query the Cloud environment accounts and present wizards with the necessary attributes to create instances and populate with the data required to support the above uses discussed. Since format conversions are handled by the Amazon Access Node, the entire operation is orchestrated even if the source of data is an on-premises hypervisor. This reduces the operational overhead, human error, and unique IT skill sets required to adopt cloud technologies.

Modeling your environment with the Commvault Solution Design Tool (CSDT)

Commvault provides the **Commvault Solution Design Tool (CSDT)** for partners to assist customers in modeling initial day-one and future day-two expansion needs. CSDT uses a data-based approach to sizing that analyzes and recommends the MediaAgent infrastructure to best suit the data types, retention period, and the number of regions. Smart defaults for deduplication and compression benefits are built-in into the estimations.

It should be stressed that the CSDT tool was built to assist in traditional infrastructure purchases and does not consider the elasticity of Amazon EC2 instances which can scale on demand. Amazon EC2 instance recommendations are therefore typically over-provisioned and do not constitute a hard requirement.

Disclaimer: CSDT is used to model and estimate the amount of computing infrastructure required to service a particular dataset, it does not model the pricing impact of consumed AWS services (e.g., Amazon S3 GET/PUT API calls, network egress costs, VPC endpoint costs). Use the **AWS Pricing Calculator** exclusively for modelling your AWS costs.

Each AWS service to be protected is entered with the required retention and number of copies (in-region, cross-region replication), and an estimate of the required number of MediaAgents to receive, store, and replicate the data is provided.

Commvault Solution Design Tool (available to partners at **cloud.commvault.com**, contact your Commvault sales engineer to model your environment).

See the **CSDT User Guide** (pdf) for additional details on how to best utilize the CSDT tool.

Provisioning and scaling production-ready backup and recovery

This guide shows how to design a day-one Commvault data management platform, and then scale for additional performance or protect additional locations.

Seed All-in-One deployment for day one

Commvault recommends deploying a Commvault all-in-one configuration on day one. An all-in-one configuration combines the CommServe, MediaAgent, and Access Node components or roles in one instance. You can get started by finding, testing, purchasing, and deploying your all-in-one configuration from the **AWS Marketplace**.

Selecting compatible Amazon EC2 instance types

Availability of Amazon EC2 instance types varies by region.

Commvault does not limit deployment to a specific Amazon EC2 instance type.

Commvault publishes AWS Marketplace images for an all-in-one CommServe® instance and expansion-based Cloud Access Nodes (MediaAgent and Access Node roles combined). Commvault has limited the compatible instance types in AWS Marketplace to only instances that meet Commvault minimum requirements.

Commvault recommends the use of current generation, EBS-optimized Compute-optimized (C Class family), Memory-optimized (R Class family), and General Purpose (M-Class family) for Commvault components.

Commvault also recommends Compute-optimized burstable instances (T-Class family) for performing snapshot-only protection and orchestration only.

Commvault recommends instances with a maximum ceiling of 128GiB RAM, as this is the maximum memory consumption observed in Commvault customers at the time of writing.

See the **Cost Optimization Pillar – Selecting cost-effective resources** for more detail on recommended options.

See the **Performance Efficiency Pillar – Compute Architecture Selection** for more details on recommended options.

Use of compute-optimized burstable instances

Commvault does not support the use of **burstable** instance types (T2, T3, T4g family) for network-streamed workloads.

Commvault network-streamed backup drives sustained CPU and network consumption that is quickly exhausted on burstable instance types. As burstable instances are designed for very-low baseline performance, with limited bursting within a 24hr period – backup SLAs may not be able to be met.

Dev/test and Proof of Concept (POC) deployments may safely use the T-Class family with the understanding that performance will not be consistent if fully consuming burstable CPU credits on a given day.

Commvault support may request the recreation of logged support issues on a non-burstable instance type before support services can be provided.

This applies to <u>any</u> Commvault infrastructure or role, including but not limited to – The CommServe® instance, MediaAgents (including IntelliSnap®), and Access Nodes (Virtual Server Agent, CloudApps).

Use of network-optimized instances

Commvault does not recommend, nor require the use of network-optimized instances that offer networking performance at or exceeding 25Gbps. This includes M5n, M5zn, C6gn, C5n, R5b, R5n, X2gd, High memory, z1d, P4, P4, P2, Inf1, G4dn, G4ad, G3, F1, i3, i3en, d3, h1).

Usage of Amazon EC2 instances with networking bandwidth exceeding 12.5Gbps will result in significant underutilization of provisioned resources leading to waste, higher EC2 runtime costs, and reduced sustainability of the solution.

Use of instance-store-backed instances

Avoid using Amazon EC2 instance types with local NVMe SSD volumes (such as the R5d, M5d. C5d, i3, and i3en instance types), as these local NVMe volumes will not retain their data if the Amazon EC2 instance is powered off or terminated, which makes them unsuitable choices for the Index Cache or DDB.

EBS Optimized Instances

If choosing an Amazon EC2 size other than the sizes recommended, choose an Amazon EC2 instance type/size which is "**EBS optimized**" to ensure guaranteed baseline IOPS performance to support MediaAgent Index Cache and DDB workloads. Review both the 'burst' EBS bandwidth offered for thirty (30) minutes per day and the baseline EBS bandwidth guaranteed.

Commvault has two (2) primary index locations that drive high random I/O. The Deduplication Database (DDB) and the Index cache volume. Each of these volumes is formatted with different block sizes, impacting the maximum number of I/O operations. Be sure to factor in the size of an I/O when assessing the baseline performance of an EC2 instance:

- Deduplication DataBase (DDB) volumes = **32KB block size**
- Index cache volume = **4KB block size**

> ⓘ **Note**
>
> If deploying Commvault in the AWS Marketplace, Commvault will automatically provision and format separate Amazon EBS gp3 volumes for each high-I/O file system.

> 💡 **Pro-Tip**
>
> **EBS-optimized instances** are recommended as they provide dedicated network bandwidth for EBS volumes, improving deduplication and Index Cache performance and freeing up bandwidth to send/receive from clients, other MediaAgents, and Amazon S3 endpoints.

Use of Microsoft Windows Cloud Access Nodes

Commvault does not publish a Microsoft Windows Server **Cloud Access Node** AMI image.

As Linux-based Amazon EC2 infrastructure represents a 50% runtime cost saving over Microsoft Windows, Commvault recommends the use of Linux-based Cloud Access Node infrastructure exclusively.

You may build and deploy a self-built Windows-based Cloud Access Nodes(s) if required.

The only requirement for running a Microsoft Windows-based Access node is a need to protect the following workloads:

- Microsoft Office 365 (O365) – including Exchange Online, SharePoint Online, Teams, and OneDrive for Business.
- Microsoft Exchange.

Where to place your instances in a multi-region deployment

CommServe placement

Place your primary CommServe instance in the AWS Region where most of your end-users are located, to minimize the latency of using the Commvault Command Center™ console, API, command-line, or SDK.

MediaAgent placement

Place at least one MediaAgent with the Virtual Server Agent (VSA), CloudApps, and relevant Database packages installed, per region with workloads to protect. If replicating data to an alternate region for Disaster Recovery, deploy at least one MediaAgent in the alternate region.

Commvault recommends writing your *primary backup copy* within the primary region and optionally replicating mission-critical backups to an alternate region for protection from regional outages.

Deploy MediaAgents in a common **GridStor® configuration** within a single zone to avoid data transfer fees, or across zones for improved resilience.

You can deploy combined MediaAgent + Access Node instances using the **Commvault Cloud Access Node** products available in AWS Marketplace.

Access Node placement

Access Nodes performing agent-less protection for Amazon EC2 instances, Amazon RDS / Redshift / DynamoDB / DocumentDB databases, and Amazon EFS and Amazon FSx* file-level data. No Commvault software agents are required inside your Amazon EC2 instance (in-guest) to perform a block-level backup and provide full instance, volume, or item-level recovery.

Place at least one Access Node (installed on your MediaAgent) per region with workloads to protect.

Use Access Node auto-scaling for protecting Amazon EC2, which will auto-deploy zonal Access Nodes to protect workloads by RPO and total data volume.

Use your MediaAgent(s) for a regional baseline backup and recovery capacity, scale horizontally using Access Node groups as data volume and business RPOs require more network bandwidth to complete backup or restores within business objectives.

Deploy Access Nodes in Access Node groups to provide load-balancing and resilience to data management activities.

> ⓘ **Note**
>
> Commvault will attempt to utilize Access Nodes from the Availability Zone (AZ) matching the workload being protected or recovered, to reduce **Data Transfer within the same AWS Region** fees. If a zonal Access Node cannot be located, a regional Access Node will be used with relevant cross-AZ data transfer fees.

You can avoid Amazon EBS volume initialization from S3 (or lazy-loading) by using Commvault HotAdd transport mode for the restore. To perform a HotAdd restore you must have an Access Node within the AZ that the Amazon EC2 instance or EBS volumes are being restored to.

You can deploy combined MediaAgent + Access Node instances using the **Commvault Cloud Access Node** products available in AWS Marketplace.

See below for the supported backup and recovery use cases based on the location of the Commvault **Cloud Access Node.** Commvault recommends the use of **EBS direct APIs** for improved backup and recovery performance and simplification of infrastructure planning.

Placement of Access Node (relative to protected data)	EBS direct API (default)		HotAdd	
	Backup	Restore	Backup	Restore
Same Region	✔	✔	✔	✔
Different Region	✔[1]	✔[1]	✘	✘
Same Availability Zone (AZ)	✔	✔	✔	✔
Different Availability Zone (same region)	✔	✔	✔	✔
Different Availability Zone (different regions)	✘	✘	✔	✔
Same Account	✔	✔	✔	✔
Different Account	✔	✔	✔	✔
On-premises	✔[1]	✔[1]	✘	✘

[1] Supported but the most cost-effective protection method is in-region access to ebs.{region}.amazonaws.com endpoint

Remote Office Branch Office (ROBO) backup with Storage Accelerator

For remote office locations, small cloud environments, roaming devices such as laptops, and any architecture that proves unfeasible or cost-prohibitive to implement a traditional or cloud-based MediaAgent, backups can be done directly from the source to a cloud target such as Amazon S3., completely bypassing the MediaAgent. This is achieved by installing and enabling the Storage Accelerator feature on the client for direct communication to a storage target and will speed up the backup and reduce costs in these situations.

See **Accelerating Backups to Cloud Storage Libraries** for more details.

Storage Accelerator can be used to backup data from all **Commvault backup and restore** agents.

Distributing instances for HA/DR

Multi-availability zone placement

Commvault supports deploying multiple MediaAgents in high-availability *MediaAgent Grids* by creating **Cloud Network Storage Pools** with multiple MediaAgents. Commvault will failover deduplication processing or cloud library read/write requests between healthy MediaAgents.

Commvault supports deploying multiple Access Nodes in high-availability groups by creating Access Node groups.

Components distributed across availability zones protect backup and recovery services during a zonal failure. It should be noted that cross-AZ data transfer fees are incurred in normal operations as Commvault distributes I/O for availability in multi-AZ deployments.

Multi-region placement

Commvault does not support high-availability *MediaAgent Grids* or Access Node groups that spawn AWS regions.

Commvault CommServe LiveSync for High Availability Disaster Recovery may be used to place passive CommServe replica instances in remote regions. Any backup data required during DR failover must be replicated to the DR region to provide uninterrupted recovery services during a DR event.

Protecting multi-account landing zones

For AWS environments with multiple accounts, you can deploy your Commvault data management resources (CommServe, MediaAgent, Access Nodes, and Cloud storage) in your central backup account. When you configure protection of your Amazon EC2, and AWS Cloud database protection you configure Commvault use **Use resources in admin account** (see **Using Resources from an Admin Account**).

Commvault creates backups (i.e., resource snapshots) within the workload account and then shares them with your central backup account for backup. Commvault recommends taking a *backup copy* outside the scope of the workload account to prevent *accidental or intentional recovery point deletion*.

See **Processing for Using Resources from an Admin Account**.

Improving backup and recovery performance

There are several approaches for improving the performance of backup and recovery operations. Consider the following approaches based on your specific application and *recovery time objectives (RTOs).*

- Leverage service-level snapshots (Amazon EC2, RDS, Redshift, DocumentDB) to perform rapid backup and recovery operations.
- Use Amazon EBS direct APIs (default) to create and restore *service-independent backup copies* of EBS volumes.
- Use Commvault HotAdd to create and restore *service-independent* backup copies of EBS volumes, when the per-region, per-account **EBS service quotas** are being reached.
- Scale the number of Access Nodes being used to perform the restore for additional concurrency (1 reader per volume/workload).
- Scale the number of MediaAgents in your MediaAgent Grid to support more parallel read streams from Amazon S3.
- Increase the number of **data readers** and **device streams** to improve concurrency.

Right-sizing your data management resources

Commvault recommends using **AWS Compute Optimizer** to review when resources require right-sizing to larger or smaller instance sizes.

CommServe scaling

Monitor the consumption of CPU, memory, and network to determine when to scale your CommServe instance. As the volume of streaming backup data increases, consider migrating backup, recovery, and replication activities to a separate MediaAgent grid.

MediaAgent scaling

MediaAgents are scaled for additional performance or additional availability. Consider the following when planning to scale your regional MediaAgents:

- Deploy as many as four MediaAgents in a grid for protection from availability zone outages (see **Single-Region scenarios** for 2 9s, 3 9s, and 4 9s architectures).
- Create a *deduplication partition* for the total number of nodes and DDB volumes you want to support at maximum scale (maximum is four nodes, two partitions per node).
- Co-locate DDB partitions on the available nodes and volumes.
- Create each DDB partition in a dedicated directory to simplify DDB relocation in the future.
- When a single DDB volume MediaAgent begins to exceed the recommended Q&I times of two (2) milliseconds, redistribute DDBs on the node across two (2) discrete DDB volumes.
- When a multiple DDB volume MediaAgent begins to exceed the recommended Q&I times of two (2) milliseconds, add additional nodes and redistribute the remaining DDBs to newly created MediaAgent nodes.
- When the network interface for a MediaAgent node exceeds 70% sustained usage during backup, add another MediaAgent host to the grid (and migrate the DDBs intended for that node).

Refer to the **Snapshot and Streaming MediaAgent Grid Specifications** for the guidelines on the maximum amount of deduplicated storage each configuration can support. These are guidelines for paper-based planning only and will differ based on data types, protection frequency, retention, and data change rates.

Access node scaling

Commvault will automatically scale the required number of Access Nodes to perform Amazon EC2 backup.

For non-Amazon EC2 protection and recovery, add Access Nodes to meet your required *recovery point objective (RPO)* and *recovery time objective (RTO)*. You can provide a baseline performance level by ensuring your MediaAgents are installed with VSA, CloudApps, and Database packages to perform the role of Access Node.

Reference - CommServe Storage Layout

The following table shows the required volume layout for an all-in-one CommServe® instance. Volumes and contained file-system may be expanded online to respond to growing disk usage or performance demands (see **Extend a Windows file system after resizing a volume**).

If deploying within AWS Marketplace, Commvault will auto-create and configure these volumes within Commvault at first boot.

AWS Commvault Backup & Recovery - Drive Layout				
Drive Letter [Label]	Initial Capacity (GiB)	File-system Block size	Type / Performance	Purpose
C:\ [WINOS]	35 (default)	4096 (default)	gp3 / 3000 IOPS at 4KB / 125 Mbps	Microsoft Operating System.
E:\ [CVLT] E:\SoftwareCache	60	4096 (default)	gp3 / 3000 IOPS at 4KB / 125 Mbps	Commvault binaries, MS SQL Server binaries, Commvault log files, & Commvault software cache
F:\ [MSSQL]	40	65536 (64K)	gp3 / 3000 IOPS at 64KB / 125 Mbps	MS SQL database files + tempdb
G:\ [TLOGS]	10	65536 (64K)	gp3 / 3000 IOPS at 64KB / 125 Mbps	MS SQL transaction logs
H:\ [DDB1]	100	32768 (32K)	gp3 / 3000 IOPS at 32KB / 125 Mbps	Commvault Deduplication DataBase (DDB) #1 of 2
I:\ [INDEXC] I:\IndexCache	50	32768 (32K) (default)	gp3 / 3000 IOPS at 32KB / 125 Mbps	Commvault Index Cache
J:\ [JOBS] J:\JobResults J:\DR	100	32768 (32K)	gp3 / 3000 IOPS at 32KB / 125 Mbps	Commvault Job Results, 3DFS Cache, DR backup location, temporary upgrade location.

Each listed drive must be a separate dedicated volume.

Volumes must use **gpt** partitioning and must be formatted as **NTFS** with the specified block size.

① **Note**: Stated IOPS are a *day one* configuration and may be increased as indicated by AWS Compute Optimizer – EBS recommendations.

Separating Commvault data from the Operating System volume (C:\) complies with Amazon's best practices, see below:

> *Use separate Amazon EBS volumes for the operating system versus your data. Ensure that the volume with your data persists after instance termination. For more information, see Preserve Amazon EBS volumes on instance termination.* **Best practices for Amazon EC2**.

Reference – Cloud Access Node Storage Layout

The following is the default storage layout for an arm64 or x86_64 Cloud Access Node deployed from the **AWS Marketplace.** These nodes may be used as MediaAgents, Access Nodes, or perform both roles.

AWS Commvault Cloud Access Node - Drive Layout				
Volume group [mountpath]	Capacity (GB)	Block size	Type / Performance	Purpose
nvme0n1 /	10	4096	gp3 / 3000 IOPS at 4KB / 125 Mbps	Linux Operating System.
nvme1n1 [vg_commvault]				
vg_commvault-lv1 /opt/commvault	10	4096	gp3 / 3000 IOPS at 4KB / 125 Mbps	Commvault binaries, Commvault software cache
vg_commvault-lv2 /var/log/commvault	4.9	4096	gp3 / 3000 IOPS at 64KB / 125 Mbps	Commvault log files
vg_commvault-lv3 /mnt/commvault_jobresults	40	4096	gp3 / 3000 IOPS at 32KB / 125 Mbps	Commvault Job Results, 3DFS Cache, temporary upgrade location.
vg_commvault-lv4 /opt/commvault_indexcache	25	4096	gp3 / 3000 IOPS at 32KB / 125 Mbps	Commvault Index Cache
nvme2n1 [vg_commvault2]				
vg_commvault2-lv_ddb /mnt/commvault_ddb	20	4096	gp3 / 3000 IOPS at 32KB / 125 Mbps	Commvault Deduplication DataBase (DDB) #1 of 2

Reference - Scalability limits

The following are the scalability limits for each Commvault component. See linked Amazon resource limit documentation for details on which limits may be adjusted via AWS support request.

Commvault Scalability Limits	
Number of CommServe® instances per environment	One (1)
Number of CommServe® instances per Private Metrics centralized reporting environment	No limit.
Maximum number of MediaAgents (nodes) per CommServe	No limit.
Maximum number of MediaAgent Grids per environment	No limit.
Maximum number of MediaAgent nodes per high-availability MediaAgent grid	Four (4)
Maximum frontend storage protected by a single MediaAgent grid	No limit.

Maximum backend storage managed by a single MediaAgent grid	4PB
Maximum number of Commvault-protected workloads using AWS snapshots *Amazon EC2, Amazon RDS, Amazon Redshift, Amazon DocumentDB*	No limit
Maximum number of Commvault-protected workloads using network-streamed backup copies (per CommServe® instance)	25,000
Maximum network throughput (GB/hr.) for a single standalone MediaAgent node performing network-streamed backup *Benchmark test perform on c7g.12xlarge with **ReadAhead**=256, **WriteBehind**=256 optimizations applied for optimal transfer speed from EBS direct service endpoint.*	284 (backup) 190 (restore)
Maximum number of Amazon EBC snapshots per Region, per account	1,000 per second
Maximum number of concurrent Amazon EBS snapshot backups per AWS account, per region	5 See **Amazon EBS – Service quotas** • Concurrent snapshots per General Purpose SSD (gp2) volume • Concurrent snapshots per General Purpose SSD (gp3) volume • Concurrent snapshots per Magnetic (standard) volume • Concurrent snapshots per Provisioned IOPS SSD (io1) volume • Concurrent snapshots per Provisioned IOPS SSD (io2) volume • Concurrent snapshots per Throughput Optimized HDD (st1) volume
Maximum number of Amazon EBS direct API GetSnapshotBlock requests per account, per region (used for Amazon EC2 backup)	1,000 per second
Maximum number of Amazon EBS direct API GetSnapshotBlock requests per snapshot, account, per region (used for Amazon EC2 backup)	1,000 per second
Maximum number of ListChangedBlocks requests per account, per region (used for Amazon EC2 incremental backup)	50 per second

Maximum number of ListSnapshotBlocks requests per account, per region (used for Amazon EC2 full backup)	50 per second
Maximum number of Pending snapshots per account, per region (used for Amazon EC2 restores)	100
Maximum number of StartSnapshot requests per account, per region (used for Amazon EC2 restores)	10 per second
Maximum number of PutSnapshotBlock requests per account, per region (used for Amazon EC2 restores)	1,000 per second
Maximum number of PutSnapshotBlock requests per snapshot per account, per region (used for Amazon EC2 restores)	1,000 per second
Maximum number of concurrent Amazon RDS user snapshots per AWS account, per region	100 See **Amazon RDS – Quotas in Amazon RDS** • Manual DB cluster snapshots • Manual DB instance snapshots
Maximum number of concurrent Amazon EBS user snapshot copies (cross-account, cross-region)	20 per destination region See **Amazon EBS – Service Quotas** • Concurrent snapshot copies per destination Region See **Amazon EBS increases concurrent snapshot copy limits to 20 snapshots per destination Region**
Maximum number of concurrent Amazon RDS user snapshot copies (cross-account, cross-region)	20 per destination region See **Copying a DB snapshot** See **Amazon RDS increases concurrent copy limit to 20 snapshots per destination region**
Maximum number of Amazon Redshift user snapshots for an account in a single region	20 See **Quotas for Amazon Redshift objects** • Snapshots

Maximum number of Amazon DocumentDB user snapshots, per account, per region	100 See **Amazon DocumentDB – Service quotas** • Manual cluster snapshots

Selecting Amazon S3 storage for backup and archive

This guide details how to select and optimize the usage of Amazon S3 storage classes for backup and archive data over the storage lifecycle of an application.

Understanding Amazon S3 storage classes

Amazon S3 provides a broad selection of **storage classes** with differing performance and cost characteristics.

Amazon S3 storage is written to at least three availability zones (AZs) to deliver 99.999999999& (11 9s) of *data durability**

* S3 One-Zone IA and S3 on AWS Outposts reside in a single region.

S3 Intelligent-Tiering	S3 Standard	S3 Standard-IA	S3 Glacier Instant Retrieval	S3 Glacier Flexible Retrieval	S3 Glacier Deep Archive	S3 One Zone-IA	S3 Outposts
Changing access patterns	Frequently accessed data	Infrequently access data	Rarely accessed data	Archive data	Long-term archive data	Re-creatable infrequently accessed data	On-premises data
Milliseconds access	Milliseconds access	Milliseconds access	Milliseconds access	Minutes to hours	Hours	Milliseconds access	Milliseconds access
No retrieval charge	No retrieval charge	Per-GB retrieval charge	Per-GB retrieval charge Free bulk retrievals Free **faster restore**	Per-GB retrieval charge Free bulk retrievals Free **faster restore**	Per-GB retrieval charge Free bulk retrievals	Per-GB retrieval charge	

Selecting Amazon S3 storage for backup and archive

Command recommends a **tiered approach** to consuming Cloud storage from your Primary backup, Secondary backup copies, and finally your Tertiary (archive) data vaults.

Frequent access storage classes (S3 Standard, S3 Intelligent-Tiering) should be avoided for backup data, except for backup data accessed in a repeatable and frequent pattern. One example of frequent access backup data would be a weekly backup that is stored and then replicated to one or many regions for disaster recovery.

Infrequent access storage classes (S3 Standard-Infrequent Access, One Zone-IA) should be used for backup data exclusively to provide low first byte latency at reduced storage cost. Commvault recommends **Amazon S3 Infrequent-Access (S3-IA)** as the default backup storage class when you know your backup data is infrequently accessed. When using One Zone-IA, consider keeping a Commvault **storage copy** of critical backup data in another location.

Archive storage classes (S3 Glacier Instant Retrieval, S3 Glacier Flexible Retrieval, S3 Glacier Deep Archive) should be used for archive data exclusively as the first-byte latency is decreased and data retrieval costs are higher than frequent or infrequent access storage classes. Commvault provides **Commvault Combined Storage Tier** to enhance the use of S3 Glacier to support simplified self-service recovery without the need to manage retrieval delay incurred with asynchronous Glacier recalls.

Amazon S3 Intelligent-Tiering storage class provides automated monitoring and transitioning of S3 objects between frequent and infrequent storage classes based on observed access patterns. As Commvault deduplicates data before writing into Amazon S3, Amazon S3 Intelligent-Tiering will result in backup data occupying the S3 Standard and S3 Infrequent-Access storage classes only. Use S3 Intelligent-Tiering if you are migrating a largely inactive dataset to Amazon S3, consider **sealing the deduplication database** after the migration is complete to further reduce any ongoing write activity.

S3 Intelligent-Tiering	S3 Standard	S3 Standard-IA	S3 Glacier Instant Retrieval	S3 Glacier Flexible Retrieval	S3 Glacier Deep Archive	S3 One Zone-IA	S3 Outposts
Changing access patterns	Frequently accessed data	Infrequently access data	Rarely accessed data	Archive data	Long-term archive data	Re-creatable infrequently accessed data	On-premises data
Milliseconds access	Milliseconds access	Milliseconds access	Milliseconds access	Minutes to hours	Hours	Milliseconds access	Milliseconds access
Backup ⚠	Backup Archive	Backup ★ Archive	Archive	Archive	Archive	Backup ★ Archive	Backup Archive
0-90+ days retention	less than 30 days retention	30 days or more retention	90 days or more retention	90 days or more retention	180 days or more retention	30 days or more retention	Any data retention required on-premises

★ Recommend default for backup data.

⚠ Recommended for frequently accessed backup data with unknown access patterns.

Using Commvault deduplication for reduced storage costs

Commvault recommends using deduplication and compression for <u>all</u> backup and archive data stored in Amazon S3.

Commvault uses your MediaAgent and optionally your Access Nodes or clients to identify and discard duplicate data during the backup process. Removing duplicate or redundant data is a well-architected best practice and leads to reduced Amazon S3 storage costs and data transfer costs when replicating backups between Regions.

> ⓘ **Note**
>
> Commvault does support the creation of Cloud storage locations with deduplication and/or compression disabled, however, this is not commonly recommended as most data types will experience storage reduction benefits from Commvault storage optimization.

Leveraging Commvault Combined Storage Tier for archives

Commvault discourages the use of Amazon S3 Glacier storage classes directly as a **Cloud storage** location.

When storing indexed, deduplicated backups in S3 Glacier directly, Commvault software must perform an **index recall** (to determine where your data is stored) and then a <u>data recall</u> (to restore your required workload data), this can lead to undesirable end-to-end restore delays.

When creating your Cloud storage location, if you select an Amazon S3 storage class Commvault software will automatically configure and enable a <u>Combined Storage Tier</u>. By default, Commvault will place backup indexes in S3 Standard-Infrequent Access avoiding the multi-hour delay of recalling indexes before performing an actual data recall.

User-initiated workload recovery — index lookup 1 → Indexes in frequent access storage class — restore 2 → Backup Data in infrequent access storage classes

1 Lookup Commvault indexes in Amazon S3, Amazon S3 Standard-IA (minutes)

2 Restore backup data from Amazon S3 Glacier Flexible Retrieval (3-5 hours) or S3 Glacier Deep Archive (12 hours)
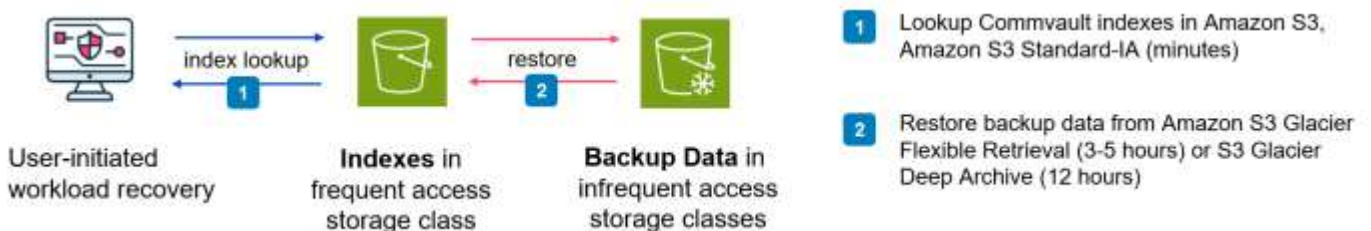
Consider enabling the use of **S3 Batch Operations** when performing restores from S3 Glacier Flexible Retrieval, S3 Glacier Deep Archive, and S3 Intelligent-Tiering Archive Access and Deep Archive Access tiers. You can achieve up to an 85% reduction in restore times from S3 Glacier by using **faster restores for Amazon S3 Glacier** recalls

Managing backup copies through the application lifecycle

Amazon S3 Storage Lifecycle refers to the ability to transition objects to lower storage classes and to expire or delete objects when they are no longer required.

Commvault *Server Plans* to provide the same functionality as Amazon S3 Storage Lifecycle but manage the transitioning or *copying* of backups and *data aging* (expiring or deleting) based on your configured **storage copies** and **data retention** settings.

Commvault does not support the use of *Amazon S3 Lifecycle policies* that transition data to a storage class with a different first byte latency.

Commvault does not support the use of *Amazon S3 Lifecycle policies* to expiry or deletion of objects from Commvault-managed Cloud storage locations.

> ⓘ **Note**
>
> Commvault cloud storage has a default storage class for all newly created objects, if using *Amazon S3 Lifecycle* to transition objects between S3 Standard, S3 Infrequent-Access, and S3 Glacier Instant Retrieval, be aware Commvault will write only new objects to the default storage class configured on the Cloud storage location.

S3 Cross-Region Replication vs. Discrete Independent Copies

Amazon S3 Cross-Region Replication (CRR) supports replication at the object storage layer from one region to another. However, replication is not synchronized with Commvault data management activities and may result in a remote replica that does not contain all objects to perform a specific instance or application point-in-time recovery (PiTR).

Commvault® software provides the ability to replicate your entire cloud storage library or perform selective copies of only the data you require in the remote location to reduce transfer costs and storage fees (see **Configuring Replication for Cloud Storage**).

Commvault recommends using Commvault auxiliary copies and DASH copies to maintain **independent Commvault-consistent copies** of your data across regions. Additionally, Commvault auxiliary copy replication may be paused, disabled, and initiated both interactively and programmatically when designing air-gapped storage solutions.

Creating immutable backup stores with Amazon S3 Object Lock

Commvault can create immutable data vaults by activating **Amazon S3 Object Lock** on your Cloud storage. Immutable backup copies are invaluable for recovery from organization-wide events that target your Primary data and Secondary (backup) copies (i.e., ransomware or malware infections).

See **Configuring WORM Storage Mode on Cloud Storage** for the instructions to enable S3 Object Lock on an existing Cloud storage location.
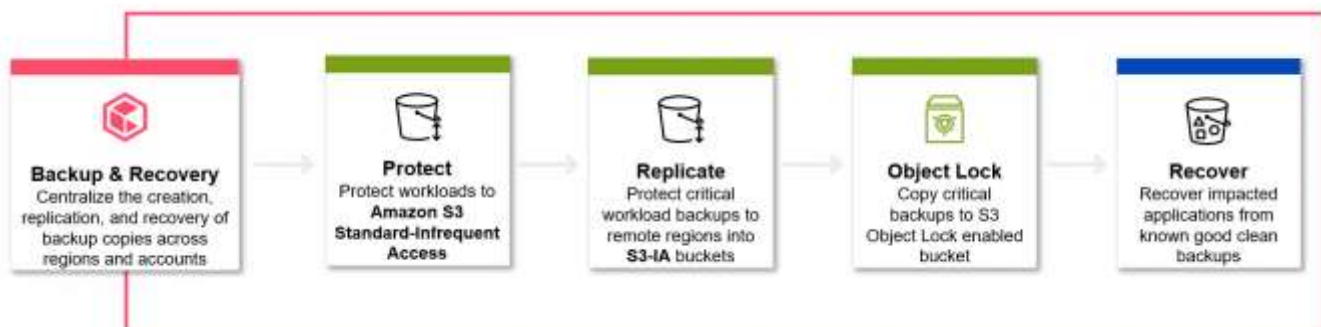
Important considerations for *immutable cloud storage* include:

- Data within the Cloud storage location is set to write-once-read-many (WORM) or compliance lock mode.
- Data retention for the copy should be set to half of the total desired retention age (e.g., a 90-day vault requires 2 x 45-day retention buckets)
- Commvault will utilize **macro pruning** to age out (expire) vault data once the retention period has been reached for an entire vault vs. micro-managing storage in ultra-low-cost archival storage classes.

- Deduplication is supported and recommended in data vaults.
- Deduplication Databases (DDBs) should be sealed upon the creation of a new vault store.
- Retention will always incur at least one additional vault store to handle data aging (see below).

An example of a **180-day retention** data vault with deduplication seals at 3 months. There will always be three (3) copies of the dataset at any time. Each vault represents a full backup of the archival content, however, the ultra-low cost of S3-IA + Glacier/Deep Archive means the storage cost is negligible.
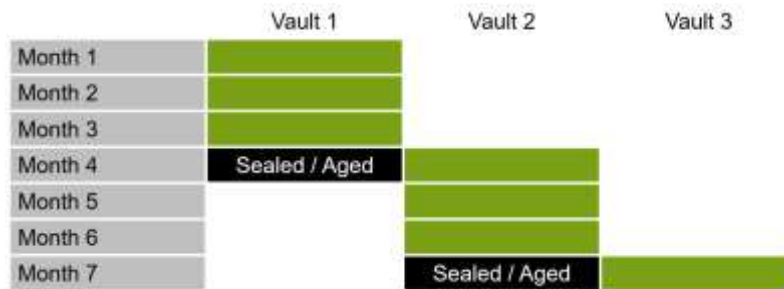
Separation of storage copies

Commvault requires an Amazon S3 bucket per **Cloud storage** location, do not co-locate Commvault backup data in existing or shared buckets due to the risk of accidental deletion.

Consider separating backup and archive data into separate AWS accounts to allow for additional logging and security controls for each data type and *data sensitivity.* For example, access to S3 buckets containing *personally identifiable information (PII)* or *company financials* may be restricted to a specific AWS account, specific MediaAgent, and specific VPC or subnet via **S3 bucket policies**.

Bulk migrating data from Amazon Glacier to Amazon S3

Amazon Glacier direct considerations

Amazon recommends migration from Amazon Glacier direct to the new **Amazon S3 Glacier** service. While Commvault supports the use of Amazon Glacier (direct) for **Commvault Cloud libraries**, the creation of a new Glacier (direct) library is now prevented via Commvault Command Center.

The creation of an Amazon glacier (direct) Cloud Storage Library will also warn the user with the following prompt:

"You can use Amazon S3 with Glacier storage class instead of Amazon Glacier. Are you sure you still want to use Amazon Glacier?"

Commvault recommends contacting your Amazon sales representative before creating new Amazon Glacier Cloud storage libraries.

Migrating Amazon Glacier data with Amazon S3 Glacier Re:Freezer

Commvault supports the mass recall of **Amazon S3 Glacier** vaults back to Amazon S3 for ongoing usage, auxiliary copy to an alternative storage solution or location (Commvault HyperScale™, disk, tape, alternate cloud provider) via the **Amazon S3 Glacier Re:Freezer solution**.

Migration (**mass recall**) of data from Amazon Glacier (direct) is subject to the recall pricing available on the Amazon Glacier website. Large-scale recall of Amazon Glacier data is very costly (in comparison to daily storage costs), Commvault often recommends sealing archival cloud libraries and letting data age out (in-place) vs. large-scale recalls onto a new storage class. New data can then be written to the new location by updating existing archival **Plans** within Commvault.

Commvault recommends the use of the **Amazon S3 Glacier Re:Freezer utility** to perform large-scale migration of Amazon Glacier (direct) vaults into an Amazon S3 bucket. Commvault has a supported procedure using Amazon S3 Glacier Re:Freezer to recall data to Amazon S3 and then inform Commvault of the new location within Amazon S3.
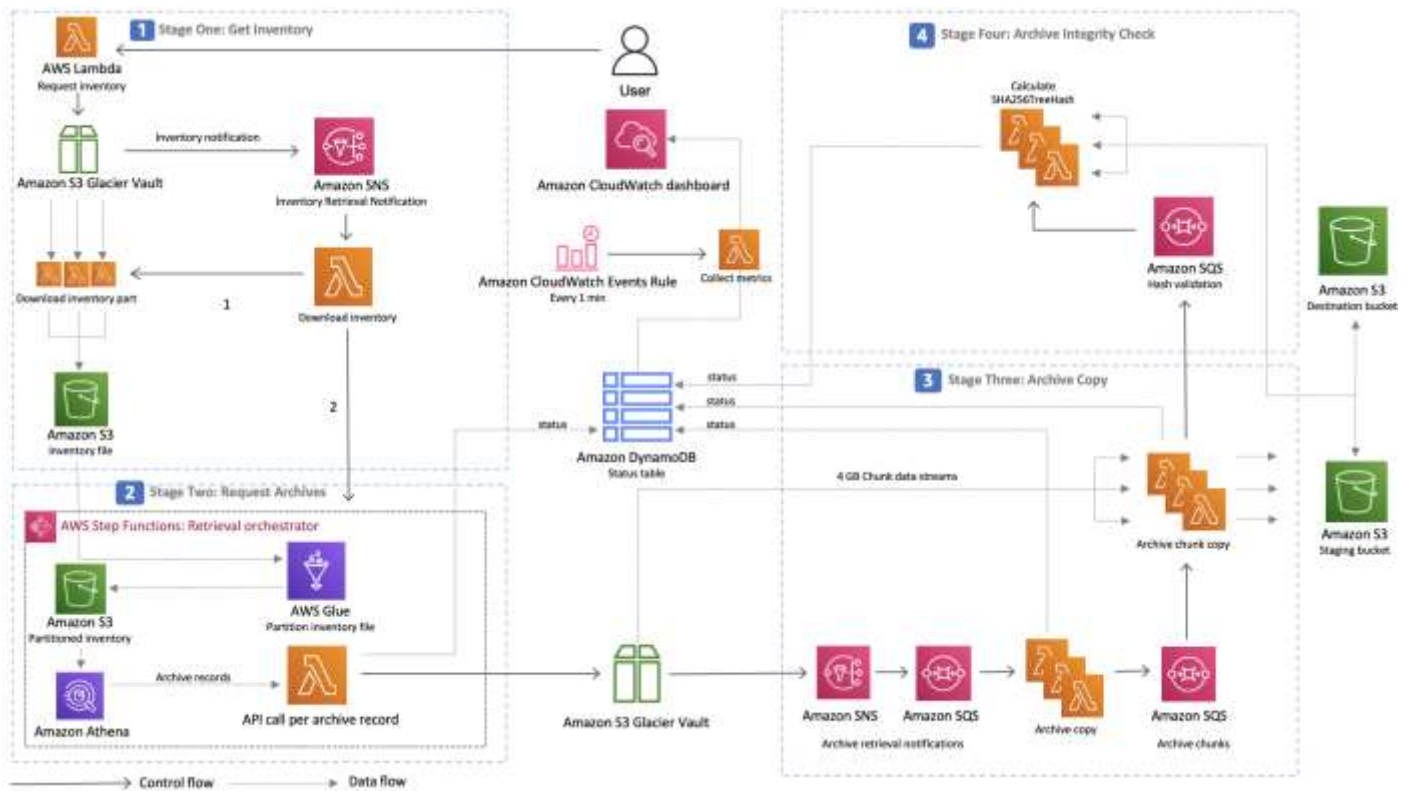
Source: Amazon S3 Glacier Re: Freezer Solutions Implementation Architecture (**URL**)

Please contact your Commvault sales representative to assess and validate your migration scenario and provide guidance on your migration scenario. **Commvault Technology Consulting** services provide Commvault Subject Matter Experts (SMEs) that can assist with assessing, recommending, and implementing your migration scenario.

See the **Amazon S3 Glacier Re:Freezer implementation guide** for more information.

> ⓘ **Note**
>
> **Amazon S3 Glacier Re: Freezer** is an AWS Solutions library solution that is not supported by Commvault directly. Amazon provides best-effort support for the AWS Solutions library via your standard Amazon support channels.

Performing a cost optimization review

AWS Billing and Cost Management

Before any **cost optimization** activities, it is important to set a baseline of costs within your Amazon environment. Commvault recommends utilizing one or many of the following tools to understand your protection cost(s) before any optimization activities.

- Apply **Cost Allocation Tags** for all Commvault infrastructure – Amazon EC2 instances, Amazon S3 buckets, Amazon VPC Endpoints, and Amazon CloudWatch alarms.
  - Commvault applies the following tags during protection operations, enabling these tags as **Cost Allocation Tags** will allow visibility into the dynamic resources that Commvault creates.

285

- o Tags: _GX_BACKUP_, _CV_Retain_Snap, _CV_Integrity_Snap, _GX_AMI_
- Utilize **Amazon S3 Storage Lens** to better understand the volume of backup data residing within the Amazon S3 service and the relative age of stored objects.
  - o Consider enabling **Amazon S3 Storage Class Analysis** (additional cost) to identify if particular Commvault Cloud Storage Libraries are under-utilized.
  - o Storage Class Access Pattern Analysis will not provide an accurate representation of data access and when deduplication is utilized. But if Amazon S3 Object age is skewed to incredibly old periods (exceeding 90 days), a tiered backup solution may be beneficial.

Amazon EBS block storage costs

Commvault recommends the use of **Amazon EBS General Purpose Volumes (gp3)** for Commvault block volume needs. Amazon EBS gp3 volumes allow tuning of capacity, IOPS, and throughput independently. Additionally, gp3 volumes represent a 20% cost reduction on gp2 volumes for all use cases.

Review your high-IOPS block volumes with the following Commvault reports:

- **Health Report: DDB Performance and Status Tile**.

> ⓘ **Note**
>
> AWS Compute Optimizer will only provide EBS recommendations on gp3, io1, and io2 volumes – this provides another reason to upgrade or use Amazon EBS gp3 volumes.

Amazon S3 backup storage costs

Commvault recommends employing Commvault software-based deduplication for all data stored within Amazon S3. Deduplication applies an inline hashing algorithm to all data blocks and only stores new, unique data (see **Optimize Storage Space Using Deduplication**). Deduplication may be used to reduce data uploads, data written, and data transferred between regions.

Amazon VPC network egress costs

Transferring data between availability zones and AWS Regions incurs a data transfer fee. Ensure that any backup data being replicated is deduplicated and compressed to reduce data transfer fees. Consider only replicating a subset of backup data to further reduce data transfer.

Amazon S3 transaction costs

Amazon S3 storage incurs costs for **Storage, Requests & data retrievals**, **Data transfer, Management & analytics, Replication,** and **Amazon S3 Object Lambda** (see **Amazon S3 Pricing**). When planning for Amazon S3 storage usage using the **AWS Pricing Calculator**, the following questions are asked:

- S3 Standard storage.
- PUT, COPY, POST, LIST requests to S3 Standard.
- GET, SELECT, and all other requests from S3 Standard.
- Data returned by S3 Select.
- Data scanned by S3 Select.

> ⓘ **Note**
>
> Contact your Commvault sales representative to help predict your expected S3 storage consumption based on the type of data, data change rate, and frequency and retention of backup copies.

How Commvault writes data to Amazon S3

Commvault stores data in Amazon S3 storage classes across multiple folders (referred to as **Chunks**), a single chunk folder holds multiple **Subfiles** (SFILE), which are represented as folders. A subfile consists of multiple **subfile containers** (SFILE_CONTAINER) files. The SFILE_CONTAINERs represent the size of S3 objects that contribute to the PUT/COPY/POST/LIST and GET/SELECT/HEAD requests.

Bucket / CVFOLDER / CV_MAGNETIC / V_nnn / **CHUNK_NNNN** / **SFILE_CONTAINER_nnn.FOLDER** / nnn
/ **CHUNK_NNNN** / **SFILE_CONTAINER_nnn.FOLDER** / nnn
/ **CHUNK_NNNN** / **SFILE_CONTAINER_nnn.FOLDER** / nnn

In the screenshot (below) showing a deduplicated Amazon S3 cloud library, you can observe:

- **CHUNK_26** folder representing an individual chunk.
- **SFILE_CONTAINER_001.FOLDER** represents an individual **Subfile** (SFILE).
- **0 .. 6** represent individual **Subfile containers** that contribute to the Subfile.

Amazon S3 GET/PUT activity occurs at the **Subfile container** level.
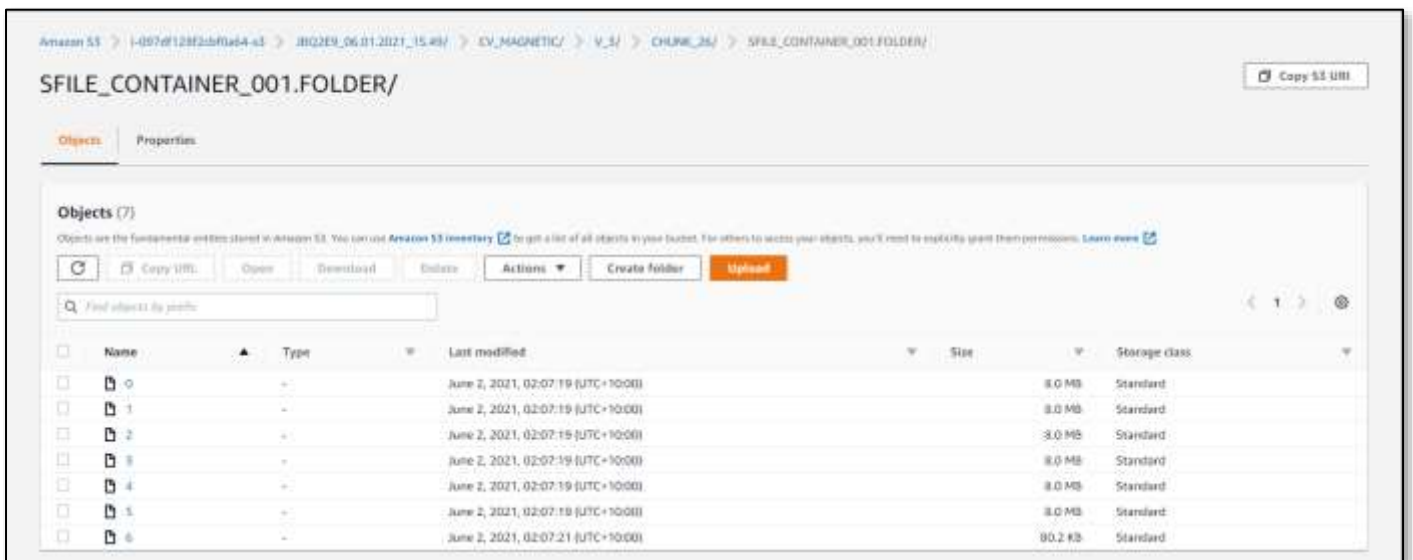
A **subfile container** represents the smallest unit requested or written by Commvault.

A **subfile container** represents the Amazon S3 Object size to model with the Amazon Pricing Calculator.

Commvault performs **Byte-Range Fetches** for all Amazon S3 restores or **GET Object** requests.

Commvault does not support modification of the **subfile container** or **Amazon S3 object size**, see details below on how to reduce Amazon S3 costs with Commvault.

Amazon S3 Object Size (Deduplicated Data)

Amazon S3 Storage Class	Commvault Chunk size	Amazon S3 Object Size
Amazon S3 Standard	4GB	8MB
S3 Intelligent-Tiering	4GB	8MB
S3 Standard-IA	4GB	8MB
S3 One Zone-IA	4GB	8MB
Glacier	4GB	32MB
S3 Glacier	4GB	32MB

Source: **Increasing chunk size**.

Amazon S3 Object Size (Non-deduplicated Data)

Amazon S3 Storage Class	Commvault Chunk size	Amazon S3 Object Size
Amazon S3 Standard	4GB	32MB
S3 Intelligent-Tiering	4GB	32MB
S3 Standard-IA	4GB	32MB
S3 One Zone-IA	4GB	32MB
Glacier	4GB	32MB
S3 Glacier (all storage classes)	4GB	32MB

It should be noted that when using **Commvault Combined Storage Tier**, subfiles are stored within the Amazon S3 Glacier Flexible Retrieval, and Amazon S3 Glacier Deep Archive only.

To perform some baseline planning, see an example backup of a single 8GB Amazon Linux EC2 instance to several types of Commvault storage:

Example Amazon EC2 Backup + Amazon S3 API Usage

Amazon S3 Storage Class	Deduplication	Size of EC2 instance	Data written (Commvault)	S3 Standard Storage (GB)	PUT, COPY, POST, LIST requests to S3 Standard	GET, SELECT, and all other requests from S3 Standard	Amazon S3 Number of Objects	Amazon S3 Size of Objects
Amazon S3 Standard	**Yes**	1.75GB	825.67 MB	828.7 MB	300	-	144	828.7MB
Amazon S3 Standard	**No**	1.75GB	824.86 MB	828.7 MB	33	-	71	1.6GB

In the example above, modeling with an **Amazon S3 Standard Infrequent Account** storage class:

- Using a **deduplication-enabled** library would result in a monthly fee of USD$0.01
  - *0.8287 S3 IA Storage x 0.0125 USD = 0.0104 USD (S3 IA storage cost)*
  - *300 PUT requests for S3 IA Storage x 0.00001 USD per request = 0.003 USD (S3 Standard-IA PUT requests cost)*
- Using a **non-deduplicated** library would result in a monthly fee of USD$0.02
  - *1.60 S3 IA Storage x 0.0125 USD = 0.02 USD (S3 IA storage cost)*
  - *33 PUT requests for S3 IA Storage x 0.00001 USD per request = 0.0003 USD (S3 Standard-IA PUT requests cost)*

Commvault recommends utilizing **deduplication** to reduce the cost of data stored and transferred between Regions. While there are specific data types that do not deduplicate, these are often a small percentage of the overall data footprint within the business.

Data recall costs

Low-cost cloud storage solutions typically have an increased cost associated with recalling data or deleting data earlier than an agreed time. Storing infrequently accessed data on a low-cost cloud storage solution may be attractive upfront, however, Commvault recommends modeling realistic data recall scenarios. In some cases, the data recall charges may be more than the potential cost savings vs. using a frequent-access storage class.

As a best practice, Commvault recommends developing realistic use case scenarios and modeling cost against the identified scenarios to ensure the Cloud solution meets your organization's SLAs, as well as cost objectives, by leveraging the **AWS cost calculator**.

A common approach is:

- Amazon S3-IA for operational backups held for at least 30 days.
- Amazon S3 Glacier for yearly long-term-retention backups.
- Amazon S3 Glacier Deep Archive for multi-year regulatory archival backups.

Commvault does not recommend the use of Amazon S3 Glacier or Amazon S3 Glacier Deep Archive as the **primary backup copy**. Primary copies are intended for operational recovery of business services, and as such should be placed on Amazon S3 services intended for backup data usage.

Sizing Guidelines

Commvault software provides the ability to build a distributed data management platform servicing a small single-region single-account environment, to an exceptionally large multi-region multi-account data landscape. Commvault is typically deployed on day 1 as a **seed deployment,** aimed at conserving cost, and then scaled with **scale-out components** which add high availability and additional data management concurrency.
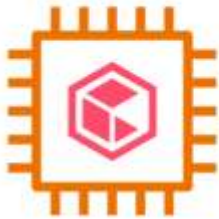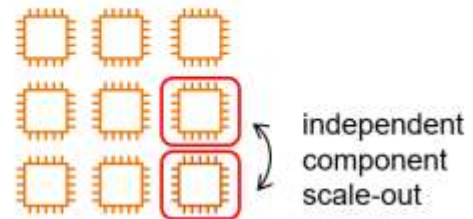
The following section provides the Commvault-recommended Amazon EC2 instances for both deployment models. As your data management needs expand, you should refer to this section to determine the most appropriate scaling steps based on your availability and performance needs.

Commvault publishes recommended sizing for:

- **Seed all-in-one CommServe® Instance**
- **Seed MediaAgent – Snapshot** (Snapshot only protection)
- **Seed MediaAgent – Snapshot and Streaming** (Snapshot and streamed protection)
- **Scale-out all-in-one CommServe® Instance**
- **Scale-out MediaAgents**

The following diagram shows the logical components of a typical Commvault data management platform.
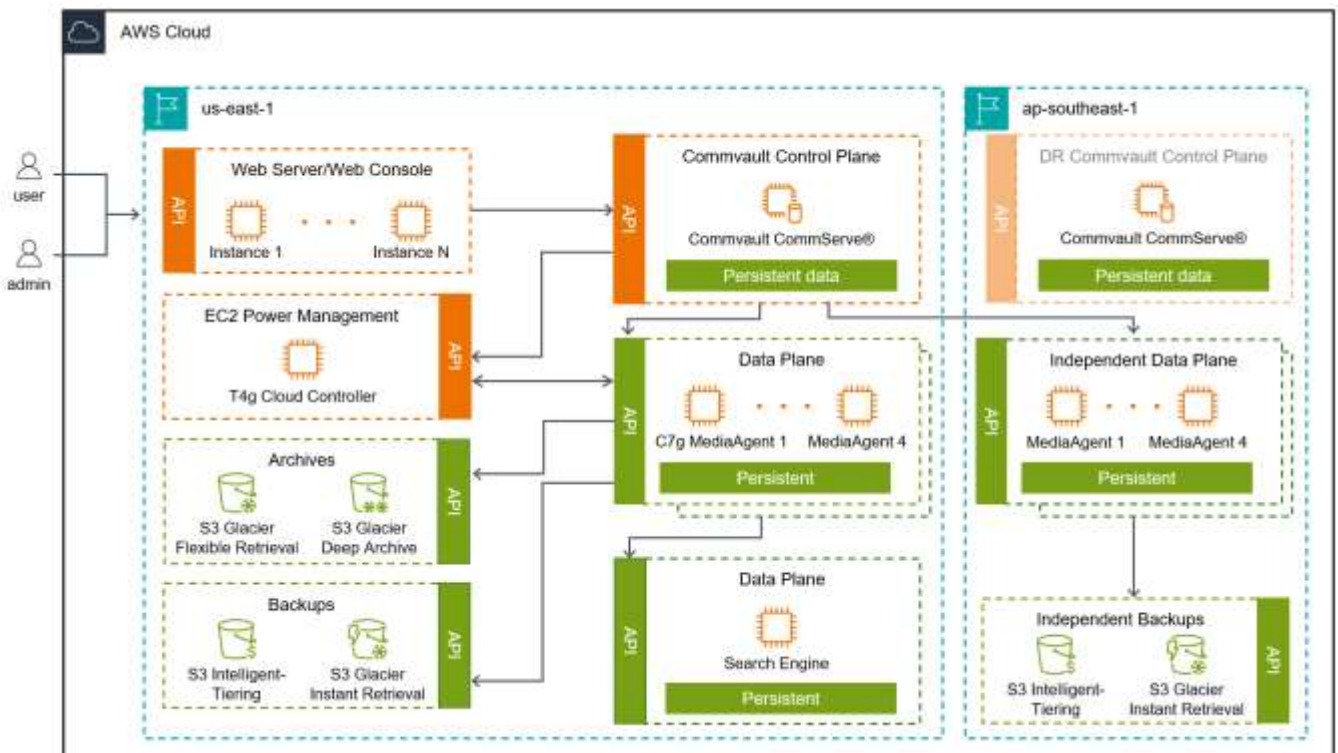
On day 1, you should deploy a **seed architecture** that deploys all components on a single Amazon EC2 instance, located within a single region and availability zone. As your data and protected workloads grow, you will add one or more of the following:

- **Web Servers** that provide secure access to the Commvault Command Center™ and Commvault REST API for authorized users, administrators, and system-to-system automation.
- **Commvault CommServe®** which provides the centralized backup & recovery orchestration, discovering workloads to protect by tag, scheduling automated protection, and replicating critical state information to an *optional* DR Commvault CommServe® in an alternate AZ or Region. CommServe servers are deployed in active/passive high-availability architectures.
- **MediaAgent grids** that are responsible for collecting workload data and writing it to backup and archive storage within Region. Grids may span availability zones for automated load-balancing and failover. Additional grids may be deployed as data volume grows, or expansion into a new Region occurs.
- **Backup data** that is written to Amazon S3 bucket(s) and resides on frequent or infrequent access storage classes. Once data is no longer needed for day-to-day operations, a long-term retention or archival copy is made to Amazon S3 Glacier Flexible Retrieval or S3 Glacier Deep Archive.
- **Search engines** are responsible for methodically indexing and searching across your live and protected data for personally identifiable information (PII), sensitive data, and data involved in legal discovery actions.
- **Cloud controllers** provide the ability to automatically power MediaAgent grids down and up in response to data management demands. *Cloud controllers are only required when the CommServe® is in an edge location outside the Region.*

> ⓘ **Important**
>
> The sizing recommendations in this document have been tailored specifically for deployment within Amazon Web Services (AWS). These recommendations replace the hardware requirements found at **docs.commvault.com** and are a recommendation only. Commvault recommends always starting small and using a data-driven approach to scaling. Use **AWS Compute Optimizer** to receive recommendations on the optimal AWS resources for your specific environment as quickly as 14 days after deployment.

Commvault recommends the following Amazon EC2 instance types for initial **seed deployments** and subsequent **scale-out expansion** to protect additional Regions and additional data volume. Commvault has identified instances that meet Commvault software <u>minimum requirements</u>. Instances are listed in priority order. Options for x86_64 (Intel, AMD) and arm64 (AWS Graviton) architecture are provided, where supported.

> 💡 **Pro-Tip**
>
> Commvault recommends selecting the instance with the **least cost** to get started and using **AWS Compute Optimizer** to guide when to right-size your instance type.

	Commvault Recommended Amazon EC2 Instance Types	
	Least Cost[1]	Most Performant[2]
Seed all-in-one CommServe® instance	r5a.xlarge r5.xlarge	r6i.xlarge r6a.xlarge
Seed MediaAgent (*snapshot protection*)	t4g.small t3a.small	*Not applicable*[3]
Seed MediaAgent (*snapshot + streaming protection*)	c6g.xlarge c5a.xlarge	c7g.xlarge c6i.xlarge
Scale-out all-in-one CommServe® instance	r5a.xlarge-4xlarge (snapshot) c6a.4xlarge-8xlarge (streamed) m6a.2xlarge-8xlarge (mixed workload)	r6i.xlarge-4xlarge (snapshot) c6i.4xlarge-8xlarge (streamed) m7i.4xlarge-8xlarge (mixed workload)
Scale-out MediaAgent Grids (*snapshot and streamed protection*)[4]	c6g,xlarge-8xlarge (snapshot and stream) c5a.xlarge-8xlarge (snapshot and stream)	c7g.xlarge-8xlarge c6i.xlarge-8xlarge

*1 See **AWS Well-Architected - Cost-effective resources** for an estimate of costs for seed and scale-out configurations.*

*2 See **AWS Well-Architected - Compute Selection** for per instance network and EBS performance details.*

3 Network performance is not critical for snapshot protection, burstable instances are recommended for reduced cost.

*4 See **AWS Well-Architected – Workload Architecture** for horizontal scaling for resilience and performance.*

Seed – Commvault CommServe® Instance

The following is the recommended day-one <u>minimum</u> configuration for getting started with Commvault software.

> 💡 **Pro-Tip**
>
> Get started in AWS Marketplace with **Commvault Backup & Recovery** or **Commvault Backup & Recovery BYOL** products.

AWS Quick Start Specifications – Seed Commvault CommServe® Instance		
Configuration	An all-in-one deployment with Commvault CommServe®, MediaAgent, Access Nodes, and Cloud Controller in one Amazon EC2 instance. Hosts deduplication cloud libraries for backups	
Instance type	Least cost *see **cost-effective resources*** r5a.xlarge* r5.xlarge*	Best price/performance see **compute selection** r6i.xlarge* r6a.xlarge*
Operating systems	Red Hat Enterprise Linux 8 $^{x86\_64}$ (recommended) Microsoft Windows 2019 $^{x86\_64}$	
In AWS Marketplace	**Yes** (Microsoft Windows 2019) $^{x86\_64}$	
Required Amazon EBS storage	All volumes are Amazon EBS General Purpose (gp3) SSD storage with initial baseline IOPS configuration. - 60GiB Commvault binaries and logs, 3000 IOPS @ 4K - 50GiB Commvault deduplication database, 3000 IOPS @ 32K - 50GiB Commvault Index Cache, 3000 IOPS @ 32K - 50GiB Commvault Job Results, Job Cache, tempdir, 3000 IOPS @ 4K - 40GiB Microsoft SQL Server 2019 datafiles, 3000 IOPS @ 64K - 10GiB Microsoft SQL Server 2019 transaction logs, 3000 IOPS @ 64K	
Use for	Snapshot and streaming protection of AWS services in the region, AWS Local Zones, AWS Wavelength, AWS Outposts, and edge-based hybrid locations. Protects Amazon EC2, Amazon EBS, Amazon EKS, Amazon RDS, Amazon Redshift, Amazon DynamoDB, Amazon DocumentDB, Amazon EFS, Amazon FSx, and Amazon S3.	
Protects	Protect up to 3.6K EBS vols per account, per supported region, per day, see **service quotas**** Avg. throughput 504GB/hr. (Backup)/ 1424GB/hr.(Restore) observed on an r5a.xlarge instance*** Plan for up to 100TiB of written streamed backup data, 5% streamed per day.	

* **Amazon EBS-optimized instances** can support maximum performance for 30 minutes at least once every 24 hours.

** Commvault software creates 5 snapshots at a time by default, to avoid exceeding **Amazon EBS resource quotas**.

*** Amazon EC2 instance has a **baseline and burst network performance**, test in your VPC for achievable throughput.

*** Avg. throughput used tuned configuration with **ReadAhead**=256, **WriteBehind**=256 for optimal transfer speed.

ⓘ **Note**: Due to Commvault's dependence on Microsoft SQL Server, the Commvault CommServe® instance is supported on Red Hat Enterprise Linux (RHEL) only (see **Installation guidance for SQL Server on Linux**).

Seed – Snapshot-only MediaAgent Grids

The following is the recommended day-one <u>minimum</u> configuration for initial MediaAgent deployment (single node) responsible for performing a snapshot-only backup architecture. MediaAgents may be combined in resilience grids of one to four nodes. MediaAgent grids must consist of identical operating systems and CPU architecture (arm64, x86_64).

> ☀ **Pro-Tip**
>
> Get started in AWS Marketplace with **Commvault Backup & Recovery** or **Commvault Backup & Recovery BYOL** products.

AWS Quick Start Specifications – Seed Commvault MediaAgent Instance (**single node, snapshot only**)		
Configuration	A single node all-in-one MediaAgent, Access Node (Virtual Server Agent), CloudApps, and IntelliSnap® instance that hosts non-deduplicated cloud libraries for snapshot backup indexes.	
Instance type	<u>Least cost</u> see *cost-effective resources* t4g.small* t3a.small* t3.small*	<u>Best price/performance</u> see **compute selection** Not applicable, all instances have an identical network and EBS performance profile.
Operating systems	Amazon Linux 2 ^{arm64} (recommended) Red Hat Enterprise Linux 8 $^{x86\_64}$ Microsoft Windows 2019 $^{x86\_64}$	
In AWS Marketplace	**Yes** (Amazon Linux 2 ^{arm64}, Red Hat Enterprise Linux $^{x86\_64}$)	
Required Amazon EBS storage	All volumes are Amazon EBS General Purpose (gp3) SSD storage. - 80GiB Commvault binaries / Logs / Job Results / Index Cache, 3000 IOPS @ 4K - 25GiB Commvault deduplication database, 3000 IOPS @ 4K *Your backup metadata is stored in scalable, durable, and secure Amazon S3 storage.*	
Used for	Performing AWS snapshot creation, sharing, and copying between regions and accounts for Amazon EBS, Amazon RedShift, and Amazon DocumentDB resources. Writing backup activity for snapshot-only jobs to a hosted non-deduplicated cloud library. (Optional) Sending backup activity for snapshot-only jobs to a remote cloud library.	
Protects	Protect up to 3.6K EBS volumes per account, per supported region, per day, see **service quotas****	

* **Amazon EBS-optimized instances** can support maximum performance for 30 minutes at least once every 24 hours.

** Commvault software creates 5 snapshots at a time by default, to avoid exceeding **Amazon EBS resource quotas**.

Seed – Snapshot and Streaming MediaAgent Grids

The following is the recommended day-one <u>minimum</u> configuration for an initial MediaAgent grid responsible for performing snapshot and streaming backup. MediaAgents may be combined in resilience grids of one to four nodes. Snapshot-only grids cannot be mixed with Snapshot and Streaming MediaAgent grids. MediaAgent grids must consist of identical operating systems and CPU architecture (arm64, x86_64).

> 💡 **Pro-Tip**
>
> Get started in AWS Marketplace with **Commvault Backup & Recovery** or **Commvault Backup & Recovery BYOL** products.

AWS Quick Start Specifications – Seed Commvault MediaAgent Instance (**single node, snapshot & streaming**)	
Configuration	A single node all-in-one MediaAgent, Access Node (Virtual Server Agent), CloudApps, and IntelliSnap® instance that hosts deduplicated cloud libraries for snapshot and streamed backups.
Instance type	<u>Least cost</u> *see **cost-effective resources*** c7g.xlarge* c5a.xlarge* <u>Best price/performance</u> see **compute selection** c6i.xlarge* c6a.xlarge*
Operating systems	Amazon Linux 2 ^{arm64} (recommended) Red Hat Enterprise Linux 8 $^{x86\_64}$ Microsoft Windows 2019 $^{x86\_64}$
In AWS Marketplace	**Yes** (Amazon Linux 2 ^{arm64}, Red Hat Enterprise Linux 8 $^{x86\_64}$)
Required Amazon EBS storage	All volumes are Amazon EBS General Purpose (gp3) SSD storage. - 80GiB Commvault binaries / Logs / Job Results / Index Cache, 3000 IOPS @ 4K - 25GiB Commvault deduplication database, 3000 IOPS @ 4K
Used for	Performing AWS snapshot creation, sharing, and copying between regions and accounts for Amazon EBS, Amazon Redshift, and Amazon DynamoDB. Creating service-independent streamed backup copies of AWS workloads (Amazon EC2, Amazon RDS, Amazon DynamoDB, Amazon EFS, Amazon FSx, Amazon S3) to Amazon S3 deduplicated storage and replicating to alternate regions for disaster recovery purposes. Writing backup data for snapshot and streaming jobs to a hosted deduplicated cloud library. (Optional) Sending backup data for snapshot and streaming jobs to a remote cloud library.
Protects	Protect up to 3.6K EBS volumes per region and account, per day, see **service quotas**** Avg. throughput of 684GB/hr. (backup) / 781GB/hr. (restore) observed on a c7g.xlarge instance*** Plan for up to 100TiB of written streamed backup data, 5% streamed per day.

* **Amazon EBS-optimized instances** can support maximum performance for 30 minutes at least once every 24 hours.

** Commvault software creates 5 snapshots at a time by default, to avoid exceeding **Amazon EBS resource quotas**.

*** Amazon EC2 instance has a **baseline and burst network performance**, test in your VPC for achievable throughput.

*** Avg. throughput used tuned configuration with **ReadAhead**=256, **WriteBehind**=256 for optimal transfer speed.

Seed – Cloud Controller

Commvault can deploy a minimal EC2-based instance to perform power management of Amazon EC2 MediaAgents. This is referred to as a **Cloud Controller**. The Cloud Controller is an Amazon EC2 or virtual machine with the Commvault Virtual Server Agent (VSA) software installed.

Use the **Seed Snapshot-only MediaAgent** for sizing a cloud controller, it requires only minimal resources to orchestrate the creation of snapshot-based backups and write backup metadata to Commvault Amazon S3 storage.

> ⚙ **Pro-Tip**
>
> Commvault recommends enabling your cloud controllers in **server groups**, distributed across availability zones or regions for improved resilience of **cloud power management**. Deployment on Amazon EC2 instances allows the use of STS:AssumeRole or IAM Role authentication, which aligns with AWS's well-architected best practices for authentication and authorization.

Scale-out – Commvault CommServe® Instance

The following are the options to scale out the CommServe® instance to provide additional resilience or support additional concurrency of data management activities (backup, recovery, replication).

These configurations are recommended guidance only, use **AWS Compute Optimizer** to receive recommendations and tune your instance to meet your needs. There are two scale-out options for the CommServe instance:

- Horizontal scaling by adding a second identical CommServe Instance in an alternate availability zone or region to provide a passive failover instance for the CommServe component.
- Vertical scaling by increasing the instance size of the CommServe to add CPU, RAM, network, and I/O performance to handle additional concurrent data management workload.

> ⚙ **Pro-Tip**
>
> Get started in AWS Marketplace with **Commvault Backup & Recovery** or **Commvault Backup & Recovery BYOL** products.

AWS Quick Start Specifications – Scale-out Commvault CommServe® Instance	
Configuration	An all-in-one deployment with Commvault CommServe®, MediaAgent, Access Nodes, and Cloud Controller in one Amazon EC2 instance. May host deduplication cloud libraries for backups.
Instance type	Least cost *see cost-effective resources* r5a.xlarge-4xlarge (snapshot)* c6a.4xlarge-8xlarge (streamed)* m6a.2xlarge-8xlarge (mixed workload)* Best price/performance see **compute selection** r6i.xlarge-4xlarge (snapshot)* c6i.4xlarge-8xlarge (streamed)* m7i.4xlarge-8xlarge (mixed workload)*
Operating systems	Red Hat Enterprise Linux 8 $^{x86\_64}$ (recommended) Microsoft Windows 2019 $^{x86\_64}$
In AWS Marketplace	**Yes** (Microsoft Windows 2019) $^{x86\_64}$
Required Amazon EBS storage	All volumes are Amazon EBS General Purpose (gp3) SSD storage with initial baseline IOPS configuration. The following volumes are expected to grow in capacity and IOPS as managed data volume grows, utilize Amazon EBS gp3 IOPS, capacity, and throughput tuning to match demand. - Plan for Deduplication Database capacity of 0.002% of total stored data in Amazon S3. - Plan for Index Cache capacity of 0.0025% of total stored data in Amazon S3. - Use the **Job Results Directory Disk Space Calculation** to estimate Job Results size. *Your backups are stored in separate scalable, durable, and secure Amazon S3 buckets.*
Use for	Snapshot and streaming protection of AWS services in the region, AWS Local Zones, AWS Wavelength, AWS Outposts, and edge-based hybrid locations. Write backup data to local or remote deduplicated cloud libraries. Protects Amazon EC2, Amazon EBS, Amazon EKS, Amazon RDS, Amazon Redshift, Amazon DynamoDB, Amazon DocumentDB, Amazon EFS, Amazon FSx, and Amazon S3.
Protects	Protect up to 3.6K EBS vols per account, per supported region, per day, see **service quotas**** Avg. throughput 504GB/hr. (Backup) / 1424GB/hr.(Restore) observed on an r5a.xlarge instance*** Plan for up to ~43TiB/core of managed streamed backup data (using two DDB volumes). Plan for expected concurrent backup, restore, and replication throughput (GB/hr.)

* **Amazon EBS-optimized instances** can support maximum performance for 30 minutes at least once every 24 hours.

** Commvault software creates 5 snapshots at a time by default, to avoid exceeding **Amazon EBS resource quotas**.

*** Amazon EC2 instance has a **baseline and burst network performance**, test in your VPC for achievable throughput.

*** Avg. throughput used tuned configuration with **ReadAhead**=256, **WriteBehind**=256 for optimal transfer speed.

Scale-out – MediaAgent Grids

The following are the options to scale out an existing Commvault environment to support additional concurrent protection (backup, restore, replication). Commvault is typically deployed in a fan-out architecture with a single active CommServe controlling multiple MediaAgent grids and optionally multiple Access Node groups.

MediaAgents may be combined in resilience grids of one to four nodes. Snapshot-only grids cannot be mixed with Snapshot and Streaming MediaAgent grids. MediaAgent grids must consist of identical operating systems and CPU architecture (arm64, x86_64).

Snapshot-only MediaAgents

Snapshot-only MediaAgent grids are intended to be used to perform snapshot management for Amazon EC2, Amazon Redshift, and Amazon DocumentDB resources. Combine two to four **Seed – Snapshot-only MediaAgent** instances to create highly available **MediaAgent GridStor® grids** responsible for persisting snapshot backup indexes. Additionally, configure MediaAgents in an Access Node group to provide resilience for backup and recovery activities.

These configurations are recommended guidance only, use **AWS Compute Optimizer** to receive recommendations and tune to the best resources to meet your needs. Commvault does not publish a CommServe-only specification, always consume your compute investment in the CommServe before scaling out to additional MediaAgents.

Best Practices

The following section provides best practices to implement, anti-patterns to avoid, and known limitations for implementing your unified Commvault Data Management Platform in AWS. These best practices build on the **Amazon Well-Architected Framework** and **Commvault Well-Architected recommendations** with a focus on Commvault data management components and configuration.

Compute

Best Practices

- CV-COMP-BP01 – Always use the smallest EC2 instance size recommended by Commvault, then scale to meet business RPO and RTO requirements (see **Sizing Guidelines**).

- CV-COMP-BP02 – Use Linux-based resources by default for reduced cost, unless prevented by workload protection requirements.

- CV-COMP-BP03 – Use **AWS Graviton**-based Access Nodes and/or MediaAgents for best price-performance for cloud-based data management

- CV-COMP-BP04 – Use **AMD EPYC**-based resources as a least-cost x86_64 alternative to AWS Graviton instances.

- CV-COMP-BP05 – Use **Intel Xeon**-based resources as the best price-performance x86_64 alternative to Graviton instances, for workloads requiring an x86 chipset (e.g., Microsoft O365 protection).

- CV-COMP-BP06 – Use **Auto-Scaling Access Nodes for Amazon EC2** backup to avoid ongoing management of Access Node infrastructure.

- CV-COMP-BP07 – Enable **Cloud Power Management** on all MediaAgents to avoid EC2 runtime costs when not being used.

- CV-COMP-BP08 – Select **Amazon c7g.xlarge instances** for the best price-performance auto-scaling access nodes, if available in your region.

- CV-COMP-BP09 – Consolidate data management resources until availability or RPO/RTO needs require additional resources (i.e., all-in-one CommServe).

- CV-COMP-BP10 – Scale data management resources vertically to complete one-time data movement tasks, then right-size back to baseline sizes (e.g., migration of on-premises archives to AWS).

- CV-COMP-BP11 – Scale data management and movement resources horizontally (MediaAgent, Access Nodes, Search Engines) for more cost-effective network bandwidth and resilience.

- CV-COMP-BP12 – Always deploy Virtual Server Agent (VSA) and Cloud Apps resources to MediaAgents to provide self-contained recovery grids.

- CV-COMP-BP13 – Expand into new regions with a minimum of at least one MediaAgent + Access Node + Cloud Library to create a physically separated and isolated *regional recovery* capability.

- CV-COMP-BP14 – Use AWS Savings Plans for Commvault infrastructure that sup

Anti-patterns

- Deploying Commvault infrastructure based on static t-shirt sizes that do not reflect your workload or recovery service levels.

- Selecting Amazon EC2 instances that do not have the <u>minimum</u> specification required by Commvault software.

- Selecting Amazon EC2 burstable instances for Commvault deduplication or network streaming compute-intensive workloads.

- Selecting Amazon EC2 instances with resources that Commvault cannot use (25Gbe+ networking, GPUs, Machine-learning chipsets, Instance storage).

- Using Windows-based instances (CommServe, MediaAgent, Access Nodes) when Linux instances can be used for the same outcome.

- Using **Placement groups** for Commvault infrastructure with an expected performance increase. Commvault does not require Placement groups for distributed MediaAgent grids or Access Node groups.

- Mixing EC2 instance sizes or architectures in MediaAgent grids or Access Node groups, which results in indeterministic backup and restore performance based on the node that handles the data management activity.

Limitations

- Auto-scaling Access Nodes may only be used for backup of Amazon EC2 workloads.

- AWS Graviton-based Access Nodes cannot be used to protect Amazon EKS workloads.

- AWS Graviton-based Access Nodes cannot be used to protect Microsoft O365.

- AWS Graviton-based Access Nodes cannot perform a Live Browse of Amazon EBS snapshots, use an x86_64-based instance instead.

Additional Resources

- **Best practices for Amazon EC2**.
- **Right Sizing: Provisioning Instances to Match Workloads**.

- **AWS re:Invent 2021 - Selecting and optimizing Amazon EC2 instances**.

Containers

Best Practices

- CV-K8S-BP01 – Use **Amazon EKS**, **EKS Anywhere**, or **EKS-D** to host containerized apps requiring protection.

- CV-K8S-BP02 – Use the **Amazon EBS CSI driver** to provision, present, and protect EBS block storage to applications.

- CV-K8S-BP03 – Use the **Amazon EFS CSI driver** to provision, present, and protect EFS file storage to applications.

- CV-K8S-BP04 – Use the **Amazon FSx for NetApp ONTAP CSI** driver (NetApp Astra Trident) to provision, present, and protect high-performance NetApp-based FSxN storage to applications.

- CV-K8S-BP05 – Group EKS resources requiring protection in **namespaces** to automate discovery & protection.

- CV-K8S-BP06 – **Label** all *namespaces*, *applications*, or *persistent volumes* to automate discovery & protection.

- CV-K8S-BP07 – Implement application-consistent backups for **supported applications** or custom applications.

- CV-K8S-BP08 – Check **supported EKS releases** and stay current with Commvault updates for the latest fixes.

- CV-K8S-BP09 – Use scalable, secure, and high-performance cloud-native storage like Amazon S3 for your EKS apps. Protect **Amazon S3** data stores with Commvault.

- CV-K8S-BP10 – Use cloud-native databases like Amazon Aurora, Amazon RDS, Amazon DynamoDB, Amazon DocumentDB, and Amazon Redshift for your EKS apps. Protect your **AWS Cloud Databases** with Commvault.

- CV-K8S-BP11 – Consider implementing a **pull-through cache** on Amazon EKS for your Elastic Container Registry (ECR) hosted container images. Commvault can protect your EKS-local pull-through cache.

- CV-K8S-BP12 – Consider **copying** your critical application backups to another region to allow on-demand recovery of your containerized applications during a disaster event to a new Amazon EKS cluster.

Anti-patterns

- Locating containerized workloads requiring data protection on unsupported Amazon Elastic Container Service (ECS) or AWS Fargate.

- Locating persistent volumes (PVs) requiring data protection on non-CSI-based storage without the ability to dynamically provision or snapshot for data protection.

- Performing containerized application re-factoring and protection without considering modernization of underlying application storage or databases.

- Performing containerized application protection without considering adjacent resources (Amazon EC2 compute, VMware Cloud on AWS compute, AWS cloud databases).

Limitations

Kubernetes-based application development and modernization is a fast-moving discipline that is continually evolving. Check **Restrictions and Known Issues for Kubernetes** protection for the latest list of limitations.

Additional Resources

- **Amazon EKS Best Practices**.
- **Security best practices for Amazon EKS**.
- **Amazon FSx for NetApp ONTAP CSI driver**.

Database

Best Practices

- CV-DB-BP01 – Utilize AWS Cloud database snapshots and cross-region snapshot copies as the *primary recovery point* within and across regions.

- CV-DB-BP02 – Utilize Commvault **dump-based** database full backups to provide the ability to perform recovery in-place to the original RDS instance.

- CV-DB-BP03 – Utilize Commvault **dump-based** database full backups to create cost-optimized long-term and regulatory retention backups, and to provide a service-independent copy for database mobility (i.e., migrating from Amazon RDS to Amazon RDS Custom)

- CV-DB-BP04 – Copy **Amazon Aurora** and **Amazon RDS** snapshots to a central backup account to prevent accidental deletion by authorized workload or member account users.

- CV-DB-BP05 – Copy **Amazon Aurora** and **Amazon RDS** snapshots to alternate regions to protect from regional outage events.

- CV-DB-BP06 – Use Commvault Amazon DynamoDB **Adjust read capacity** and **Adjust write capacity** to tune provisioned throughput during data management operations to meet business SLAs.

- CV-DB-BP07 – Use STS:AssumeRole <u>exclusively</u> for snapshot-based backup and recovery for AWS Cloud databases.

- CV-DB-BP08 – Use AWS Config or equivalent configuration management and monitoring tool to ensure that Amazon RDS **Option Groups** and/or Parameter groups are synchronized across regions that require them.

- CV-DB-BP09 – Use Commvault to **migrate your Amazon RDS instances** between the AWS Region and AWS Outposts, while staying protected.

- CV-DB-BP10 – Review and implement **Amazon RDS best practices** for your specific database engines, and consider **Enhanced monitoring** and **metrics** to identify backup impacts on database performance.

- CV-DB-BP11 – Use **Amazon RDS Custom** and Commvault database agents where granular database, transaction logs, data, and configuration files recovery is required (**Oracle**, **SQL Server**).

- CV-DB-BP12 – Review and design to avoid resource quotas for **Amazon RDS**, **Amazon Redshift**. **Amazon DynamoDB**, **Amazon DocumentDB**, and **Amazon Aurora**.

- CV-DB-BP13 – Ensure AWS data backup is performed by AWS Graviton-based instances for best price-performance and sustainability of snapshot and export-based database backups (see **Back Up and Recover Amazon Aurora, Amazon RDS, and Salesforce by Using AWS Graviton**).

Anti-Patterns

- Using a snapshot-only approach to protecting AWS cloud databases, regardless of cost or recovery needs.

- Migrating to AWS Cloud databases and expecting the same data protection practices used on-premises to work without modification.

- Migrating to AWS Cloud databases and expecting the ability to download (backup) and apply transaction logs to your database for roll-forward and roll-back activities.

- Attempting to restore Amazon RDS databases across regions without first staging the Amazon RDS backups within that region.

Limitations

- Commvault **PassKey password-protected recovery** is not supported for Amazon Redshift and Amazon DocumentDB resources.

- Commvault dump-based database protection does not support IAM database authentication or Kerberos authentication for the database being protected (dump, export).

- Commvault requires that SQL Server backups from on-premises being restored to Amazon RDS are taken as a single stream backup

- Commvault RDS snapshot-based backups must configure a dedicated AWS hypervisor to discover and orchestrate snapshot-based protection, then cannot re-use an existing AWS hypervisor in Commvault.

Additional Resources

- **Best practices with Amazon Aurora**.
- **Best practices for Amazon RDS**.
- **Best practices for designing and architecting with DynamoDB**.
- **Best Practices for Amazon DocumentDB**.
- **Amazon Redshift best practices**.
- **AWS re:Invent 2021 Breakout Sessions – Databases**.

Management & Governance

Best practices

- CV-MGMT-BP01 – Centrally manage your environments with **AWS Organizations** to simplify permission management to ensure workload account resources are protected with Commvault.

- CV-MGMT-BP02 – Automate the creation of your **multi-account landing zone** that separates your Production, Pre-Production, and optional Sandbox environments using **AWS Control Tower**. Build and operate separate Commvault Backup & Recovery instances in each landing zone.
  (ⓘ **Note**: *Development has stopped on AWS Landing Zone, use AWS Control Tower for landing zone deployment*).

- CV-MGMT-BP03 – Setup an **Infrastructure Organizational Unit (OU)** to hold your Commvault shared backup services infrastructure (CommServes, MediaAgents, Access Nodes, Cloud storage.

- CV-MGMT-BP04 – Setup a **Workloads OU** for accounts that will workloads that require backup and recovery services from Commvault.

- CV-MGMT-BP05 – Implement **automated account provisioning** that adds the Commvault required **IAM policies** to workload accounts, and established **STS:AssumeRole** trust to your shared Commvault administrative account.

- CV-MGMT-BP06 – Configure **AWS Budget alerts** for all AWS accounts that are part of your AWS Organization, provide the ability for your users to configure their own alerts (**Allow IAM users to create budgets**).

- CV-MGMT-BP07 – Enable Commvault-create tags (`_GX_BACKUP_`, `CV_Retain_Snap`, `CV_Integrity_Snap`, `_GX_AMI_`, `CV_Subclient`) as **AWS cost allocation tags** to gain insight into resources created by Commvault as part of protecting your AWS resources (see **Tags Created and Used by Commvault Software**).

- CV-MGMT-BP08 – Use **AWS CloudFormation** to automate the deployment and configuration of your Commvault all-in-one CommServes from **AWS Marketplace**.

- CV-MGMT-BP09 – Use Commvault-published **AWS Marketplace** AMI Machine Images (AMIs) and CloudFormation Template (CFT) delivered resources to automate deployment using Commvault and Amazon best practices.

- CV-MGMT-BP10 – Publish **Commvault application logs**, metrics, and events to Amazon CloudWatch for centralized **visualization**, **alarming**, and automated actions using **EventBridge**. ⓘ **Note**: Don't forget to protect your **Amazon S3** stored CloudWatch logs.

- CV-MGMT-BP11 – Enable **CloudTrails** in your **Organization**, all **AWS accounts,** and all Regions, written into S3 buckets in a **separate AWS account**. ⓘ **Note**: Don't forget to protect your **Amazon S3** stored Cloudtrails.

- CV-MGMT-BP12 – Configure dedicated CloudTrails for Commvault Backup & Recovery initiated activities to aid in troubleshooting and tracing of Commvault data management across your Organizations. Commvault CloudTrails will need to include **management events** (default) and **data events** (additional cost).

- CV-MGMT-BP13 – Ensure AWS Systems Manager (**SSM Agent**) is installed and its use is permitted by **Commvault IAM policy** to allow agentless file and folder recovery to Amazon EC2 instances.

- CV-MGMT-BP14 – Use **Savings Plans** to reduce the cost of baseline Commvault data management resources by up to 72% compared to On-Demand prices. Use **AWS Cost Explorer Savings Plans recommendations** to identify opportunities to convert from on-demand to AWS Savings Plans.
(ⓘ **Note**: Commvault reduces your backup compute runtime costs by auto-scaling resources and then terminating then when no longer required. You will have a **baseline usage** and **burst usage** for your Commvault compute infrastructure, Savings Plans are ideal for your **baseline usage**).

- CV-MGMT-BP15 – Tag early, Tag often, Use too many tags, Not too few (see **Automatically tag new AWS resources based on identity of role**), and use tags to identify and protect workloads following your business policy, risk-classification, and data-classification (e.g., **EC2**, **RDS**, **DynamoDB**).

- CV-MGMT-BP16 – Review tag assignment during AWS workload recovery or cloning to ensure that tags still match the data classification of the workload when it was protected. Data may change in sensitivity and value to the business over its lifetime.

Anti-patterns

- Operating your organizations with a single AWS account for shared services and protected workload accounts.

- Placing Production, Pre-Production, and Sandbox workloads in the same AWS account and/or networks.

- Placing AWS CloudTrail and Amazon CloudWatch logs in the workload account that generates them, allows for workload owner log modification or deletion.

- Attempting to troubleshoot backup & recovery performance without associated Amazon CloudWatch metrics and baseline performance.

- Attempting to perform backup and recovery with Commvault without enabling all the required IAM actions.

Limitations

- Commvault requires that in multi-account landing zones, an AWS hypervisor is created for all protected accounts (see Adding an **Amazon Web Services Hypervisor**).

- Commvault requires that one or multiple **IAM policies** are associated with your workload accounts, to protect your AWS resources, IAM policies cannot be modified or reduced outside of removing an entire use-case detailed at **Amazon Web Services Permission Usage**.

Additional Resources

- **Best Practices for Organizational Units with AWS Organizations**.
- **Best practices for AWS Organizations**.
- **Organizing Your AWS Environment Using Multiple Accounts**.
- **Best practices for AWS Budgets**.
- **AWS CloudFormation best practices**.
- **Designing and implementing logging and monitoring with Amazon CloudWatch**.
- **AWS CloudTrail Best Practices**.
- **AWS Systems Manager – Use cases and best practices**.
- **Getting Started with AWS Savings Plans**.

Networking & Content Delivery

Commvault is an enterprise-grade data management platform capable of deployment into any network architecture that your organization utilizes in and across your hybrid data locations.

Commvault supports deployment and protection across multiple Regions, accounts, and Amazon VPCs.

Commvault can be deployed in environments utilizing VPC Peering and AWS Transit Gateway (recommended).

Commvault can protect hybrid or edge locations using AWS Direct Connect (recommended) or AWS VPN services.

Commvault can mimic your internal network security using **encrypted network topologies** providing the ability to tunnel and direct traffic over specific ports, protocols, and flow directionality.

Commvault recommends the use of VPC Endpoints to ensure data access and transfer occur within the VPC (where supported. A VPC endpoint is described as:

'A VPC endpoint enables private connections between your VPC and supported AWS services and VPC endpoint services powered by AWS PrivateLink. AWS PrivateLink is a technology that enables you to privately access services by using private IP addresses. Traffic between your VPC and the other service does not leave the Amazon network.'
AWS Privatelink

Commvault can protect **Amazon VPC resources** and recreate **supported resources** during the recovery of full EC2 instances to the original Region and account. Commvault recommends validating that the required **IAM permissions** are granted to Commvault-utilized IAM roles to discover Amazon VPC and related resources, and record network configuration as part of Amazon EC2 protection.

Best Practices

- CV-NET-BP01 – Isolate Commvault backup infrastructure in an isolated Virtual Private Cloud (VPC) with added network controls and inspection.

- CV-NET-BP02 – Isolate Commvault backup infrastructure in isolated VPCs separated by environment type (dev/test, production, mission-critical) for the ability to provide granular network flow control and inspection.

- CV-NEV-BP03 – Force encryption of <u>all</u> network traffic using **mandateEncryption** and **nCLNT_FORCE_TUNNEL** and **nAUTO_TUNNEL_PROTO** settings.

- CV-NET-BP04 – Utilize **Amazon S3 Gateway endpoints** to secure and reduce the cost of transfer of backup and restore data between Commvault MediaAgents and the Amazon S3 service within the region.

- CV-NET-BP05 – Utilize **AWS PrivateLink - VPC Endpoints** (where supported) to secure and optimize the transfer of backup and restore data between protected resources and Commvault resources.

- CV-NET-BP06 – Consider consolidating VPC endpoints in a shared services network so that multiple accounts can use the same VPC endpoints to save cost (see **Reduce Cost and Increase Security with Amazon VPC Endpoints**).

- CV-NET-BP07 – Leverage **Amazon Privatelink for S3** (S3 Interface endpoints) to ensure backup data from on-premises over AWS Direct Connect is kept within your VPC.

- CV-NET-BP08 – Use Commvault client-side deduplication (**source-side deduplication**) to reduce data egress from compute instances and conserve **network I/O credits** for network-streamed backups and replication.

- CV-NET-BP09 – Use Amazon EC2 instances with **enhanced networking** for Commvault MediaAgents and Access nodes performing high-performance, low-latency data transfers.

- CV-NET-BP10 – Monitor **Enhanced Network Adapter (ENA) metrics** with Amazon CloudWatch to be aware of network packet drops caused by exceeding your instance **network baseline bandwidth** guarantees.

- CV-NET-BP11 – Restrict allowed incoming network flows to Commvault resources with **security groups** that limit incoming ports/protocols to remote management and Commvault cvd **ports** (tcp/8400,8403) only.

- CV-NET-BP12 – Use Amazon Route 53 DNS failover via health checks (**public**, **private**) to failover Commvault web-service DNS addresses between regions or availability zones during DR events.

- CV-NET-BP13 – Use **Application Load Balancers (ALBs)** to balance and failover incoming HTTP requests between healthy **Web Console and Command Center** instances.

- CV-NET-BP14 – Use **VPC peering** (small environments) or **AWS Transit Gateway** to access applications and resources in workload VPCs within and across regions.

- CV-NET-BP15 – Perform Amazon VPC and related **network resource protection** to allow for automated VPC resource recovery and forensic investigations of failed network changes (see **forensic network configuration restores**).

Anti-Patterns

- Placing Commvault data management resources and application workloads in the same VPC.

- Performing backup and recovery to Amazon S3 without using a free gateway endpoint to keep traffic internal to your VPC.

- Performing Amazon EBS direct API backup and recovery, without creating VPC endpoints within the availability zones that backup is occurring within.

- Using Amazon EB2 instances with baseline and burst bandwidth allocations without detailed observability on network utilization.

Limitations

- Commvault Cloud libraries can only contain one service host entry. Accessing a Commvault cloud library using a Gateway endpoint -or- Interface endpoint is not supported. A single access method must be used.

Additional Resources

- **Security best practices for your VPC**.
- **AWS re:Invent 2021 – Advanced Amazon VPC design and new capabilities**.
- **AWS re:Invent 2021 – Securing your data perimeter with VPC endpoints**.
- **AWS re:Invent 2021 Breakout Sessions – Networking and Content Delivery**.

Security, Identity, & Compliance

Commvault recommends that all credentials provided to Commvault utilize the Security Token Service (STS) AssumeRole capability to obtain temporary credentials in multi-account environments.

AssumeRole is the **recommended** approach for providing the least privileged, time-bound credentials for protection operations. This allows the removal of the use of access keys and secret keys use within Commvault, dramatically simplifying ongoing management.

Commvault supports most use-cases with IAM Roles:

Service	STS AssumeRole with IAM policy	IAM role	Access and secret key
Amazon EC2	✓	✓	✓
Amazon DocumentDB Amazon DynamoDB Amazon RDS Amazon RedShift	✓	✓	✓
Amazon S3 Cloud Library (target)	✓	✓	✓
Amazon S3 Object Storage (source)	✓	✓	✓
AWS KMS			✓

See the **Amazon Security Token Service (STS) AssumeRole Activation Guide** for more information.

Best Practices

- CV-SEC-BP01 – Use **STS:AssumeRole** to obtain temporary credentials with IAM roles to protect workload resources and read/write backups to centralized shared storage (see **STS Role Authentication**).

- CV-SEC-BP02 – Enable **Multi-Factor Authentication (MFA)** for your Commvault administrators and end-users as an extra level of security from unauthorized access to your backup data.

- CV-SEC-BP03 – Use **IAM roles for Amazon EC2** to attach an *IAM role* to your Commvault EC2 infrastructure which allows Commvault to securely make API requests without the need to store or manage credentials.

- CV-SEC-BP04 – Where Access Key / Secret key credentials must be used, **rotate access keys** regularly to protect from unintended access, and update credentials stored in Commvault **Credential Manager** (see **api.commvault.com** Credentials Manager). Alternatively, you can use **Commvault Software with CyberArk Password Security Platform** to simplify and centralize rotation.

- CV-SEC-BP05 – Use **Commvault Roles** to implement a least-privilege approach to users granted self-service backup, recovery, and data management using Commvault Command Center™, API, command-line, or SDK.

- CV-SEC-BP06 – Integrate Commvault web services with a centralized identity store using **SAML 2.0**, **OpenID Connect (OICD)**, or **Active Directory (AD)** to centralize the removal and management of user access.

- CV-SEC-BP07 – Add **IAM policy conditions** to your Commvault-supplied **IAM policies** to create logical boundaries that restrict Commvault access to specific resources, consider **using tags** on resources that require Commvault protection.

- CV-SEC-BP08 – Implement **AWS Organizations – Service Control Policies (SCPs)** to establish permission guardrails to control access for all IAM users and roles across your accounts. Ensure your SCPs include Commvault IAM policies that you want to enable across your organization.

- CV-SEC-BP09 – Consider using **AWS IAM Identity Center** (successor to AWS Single Sign-On) as your centralized identity store for users accessing Commvault web services. You use AWS IAM Identity Center SAML application wizard to add Commvault as a **Custom SAML 2.0 application** to enable **single sign-on**.

- CV-SEC-BP10 – Consider using **AWS Directory Server** fully-managed Microsoft Active Directory to provide a centralized identity store for your windows workloads and users accessing Commvault **web services** (Commvault supports **AWS Managed Microsoft AD** and **AD Connector**).

- CV-SEC-BP11 – Centrally govern and manage the provisioning and security of your multi-account AWS environment using **AWS Organizations** (see *Management & Governance best practices*).

- CV-SEC-BP12 – Ensure human and machine identity activity is tracked and monitored across your AWS accounts with **AWS CloudTrail** and **Amazon CloudWatch** to allow forensic inspection during and post-security incidents. Consider forwarding **Commvault Audit Trail** events to CloudWatch for centralized security incident & event management.

- CV-SEC-BP13 – Consider using **AWS Web Application Firewall (AWS WAF)** to block common attack patterns such as SQL injection or cross-site scripting (XSS) attacks on Commvault web services (Web console, **Web server**).

- CV-SEC-BP14 – Use AWS KMS to create and control cryptographic keys used to encrypt your AWS services and Commvault backup data copies (see **Adding an AWS Key Management Service Server**).

- CV-SEC-BP15 – Optionally, supply your own KMIP-compliant Key Management Server (KMS) to provide an additional level of control/separation over your cryptographic encryption keys used to encrypt Commvault backup data (see **Adding a Key Management Interoperability Protocol Server**). Commvault supports the use of **AWS CloudHSM** as a FIPS-140-3 Level 3 single-tenant HSM instance for your Commvault data.

- CV-SEC-BP16 – Consider using **AWS Certificate Manager** for the provision and management of your SSL/TLS certificates associated with public or organization-facing web services (i.e. **Commvault Command Center™**, **Compliance Search**)

- CV-SEC-BP17 – Ensure all communication (control plane, data plane) is encrypted. Ensure communication to **AWS service endpoints** is encrypted. Additionally, review per-application communication to ensure that application-level communication is encryption (i.e., Use TLS 1.1/1.2 for AWS Cloud database export backups).

Anti-Patterns

- Not managing your AWS accounts and data centrally, requiring per-account, per-workload security policies, and management.

- Not managing your users centrally, requiring account provisioning and management by application, region, or organizational unit (OU).

- Not rotating your Access Keys and Secret Keys when users leave the organization, which leads to unintended access to sensitive data and services.

- Not centrally governing and automating account provisioning and security policy assignment, leading to unprotected workloads due to missing permissions.

Limitations

- Commvault does <u>not</u> support **AWS KMS automatic key rotation**, rotation must be **disabled** (see **Key Rotation Guidelines for AWS Key Management Service Server**).
  - o As a workaround, Commvault supports API-based rotation of storage policy copies using the Commvault RESTful API (see **Rotating Master Key for a Storage Policy Copy**).
  - o Alternatively, you can use the CyberArk Password Security Platform to perform account password rotation (see **Integrating the Commvault Software with CyberArk Password Security Platform**).

Additional Resources

- **Security best practices in IAM**.
- **Security best practices for AWS Key Management Service**.
- **Best practices for AWS Managed Microsoft AD**.
- **AWS Security Hub**.
- **AWS re:Invent 2021 Breakout Sessions – Security, Compliance, and Identity**.

Storage

Design Best Practices

- CV-STG-BP01 – Use **Amazon EBS gp3** volumes <u>exclusively</u> for Commvault block volumes, to allow for tuning of capacity, IOPS, and throughput independently, while saving up to 20% on storage costs.
  (ⓘ **Note**: This best practice does <u>not</u> apply to Commvault backup data, which uses Amazon S3 exclusively.).

- CV-STG-BP02 – Use **AWS Compute Optimizer – EBS volume recommendations** to obtain EBS performance recommendations for right-sizing for Commvault block volumes.

- CV-STG-BP03 – Use Amazon EBS snapshots as the *primary recovery point* for low-RTO, rapid recovery of Amazon EC2 workloads. Enable snapshot management by **enabling IntelliSnap®** (not default) on your AWS **VM Groups** in the Commvault Command Center console, via API, CLI, or SDK.

- CV-STG-BP04 – Create Commvault **service-independent backup copies** of your Amazon EC2 instances and Amazon EBS volumes. Backup copies reduce backup costs, for example, EBS Snapshots are charged at

$0.05/GB-month, whereas S3 Infrequent-Access is charged at $0.01 per GB. As backups age, they are less likely to be used for recovery and can be copied to cost-effective S3 storage.

- CV-STG-BP05 – Use **Amazon EBS direct APIs** to create backup copies of data residing in Amazon EBS snapshots. Commvault **auto-scales zonal access nodes** to access the snapshots, and copied data to Commvault-optimized Amazon S3 storage within and across regions. EBS direct APIs reduce backup time by up to 80%, reducing backup infrastructure runtime costs.

- CV-STG-BP06 – Use **Amazon EBS direct APIs** to restore Amazon EC2 instances by creating EBS snapshots from data stored in Commvault backup copies. Commvault access nodes create an empty EBS snapshot, populate it with backup data, then create a new EC2 instance with a volume created from the snapshot.

- CV-STG-BP07 – Create an AWS PrivateLink **interface VPC endpoint** for the EBS service in each subnet that you will perform backup or recovery. AWS Privatelink provides connectivity to AWS services without exposing network traffic to the internet. Using EBS interface endpoints ensures that backup and recovery control and data plane traffic stays within your VPC and does not traverse Internet Gateway (IGW) or NAT Gateway (NGW).

- CV-STG-BP08 – Use the **Amazon S3 Infrequent Access (S3 Standard-IA)** storage class for **Commvault cloud storage** storing backup copies with a data retention period of 30 days or more. S3-IA has a minimum storage duration of 30 days which matches most primary backup retention windows (*S3-IA is the default when creating new Amazon S3 cloud storage within Commvault software*).

- CV-STG-BP09 – Use the **Amazon S3 Standard** frequent access storage class for **Commvault cloud storage** storing backup copies with a data retention period of fewer than 30 days, due to no minimum storage duration limit (*this is not a common use-case and is reserved for data known for frequent access during the retention period).*

- CV-STG-BP10 – Use **Amazon S3 Intelligent-Tiering (S3-INT)** storage class for **Commvault cloud storage** storing backup copies with a data retention period of up to 180 days, using Automatic Access tiers only (Disclaimer: Commvault deduplication of storage data will impact access pattern analysis and may prevent data movement to infrequent access storage classes – see CV-STG-BP11.)

- CV-STG-BP11 – Create a **Commvault Combined Storage Tier** cloud storage location to selectively copy backups and archives with a data retention period of greater than 90 days to S3-IA + **S3 Glacier Instant Retrieval**, S3-IA + S3 **Glacier Flexible Retrieval**, or S3-IA + **S3 Glacier Deep Archive**. Combined Storage Tier stores minimal indexing data in the frequent access tier (S3-IA) to perform **automatic recall** from the archive access tier.

- CV-STG-BP12 – Select the S3 Glacier-based storage class of your **Commvault Combined Storage Tier** based on your business service-level for *recovery time objective (RTO),* which dictates how long the business is willing to wait for the recall of archival data. Refer to the **first byte latency** offered by each storage class. Consider *faster restores for Amazon S3 Glacier* restore optimizations when assessing achievable RTO.

- CV-STG-BP13 – Use Commvault HotAdd backup and/or **recovery** to perform multiple concurrent backups or restore operations that would exceed the **EBS service quotas** per region, per account (i.e., `ebs:PutSnapshotBlock` requests per snapshot account are limited to 1,000 per second).

- CV-STG-BP14 – Use Commvault HotAdd **recovery** when production-level performance is required immediately after booting a restored EC2 instance. HotAdd recovery removes the requirement to **initialize or pre-warm EBS volumes** as the volume is not provisioned from an EBS snapshot stored in Amazon S3.
(ⓘ **Note**: Requires a Commvault Access Node in restore availability zone).

- CV-STG-BP15 – Use **Automatic Synthetic Full schedules**, to minimize the amount of S3 GET retrieval activity from Cloud storage.

- CV-STG-BP16 – Enable Amazon S3 Server-Side Encryption (**SSE-S3**, **SSE-KMS**, **SSE-C**, or **DSSE-KMS**) on all Amazon S3 buckets used as Commvault cloud storage. Commvault will <u>default</u> to using SSE-S3 when creating a cloud library from within the Commvault Command Center UI.

- CV-STG-BP17 – Enable **S3 Bucket Key** when using Server-Side Encryption with KMS keys stored in AWS KMS (SSE-KMS, DSSE-KMS). S3 Bucket Keys decrease the request traffic from Amazon S3 to AWS KMS and reduce the cost of SSE-KMS.

- CV-STG-BP18 – Optionally (to CV-STG-BP16) Enable Server-Side Encryption with customer-provided keys (**SSE-C**). Use SSE-C encryption sparingly, as the key must be stored in base64 format in a Commvault additional setting to enable reading and writing from the SSE-C encrypted Cloud storage (see **sCloudS3ServerSideEncryptionBase64CustomerKey**).

- CV-STG-BP19 – Optionally (to CV-STG-BP16, BP18) Enable Commvault FIPS-140-3 compliant ^{pending CMVP review} **software-encryption** of Cloud storage using Commvault or **Key Management Server** supplied encryption keys. Be aware that software encryption occurs on the Commvault MediaAgent and has a significant CPU load impact.

- CV-STG-BP20 – If use Commvault software encryption, optionally add an **AWS Key Management Service (AWS KMS)** server to store your encryption keys for Commvault cloud storage.

- CV-STG-BP21 – Ensure public access is **blocked** on all Commvault S3 buckets (<u>Warning:</u> Commvault-created buckets will allow public objects by default).

- CV-STG-BP22 - Enforce **encrypted connections** over HTTPS (TLS) using the **aws:SecureTransport** condition via Amazon S3 bucket policies.

- CV-STG-BP23 – Use **S3 Access Points** to restrict access to Commvault Cloud storage buckets to specific VPC and/or AWS account IDs. This allows increased protection of extremely sensitive backups.

There are several storage technology options available within Amazon S3 service that should <u>not</u> be used with Commvault due to the way Commvault stores and retrieves data. The following remaining best practices guide the features and functions are <u>should not be used</u> with your Commvault data management platform.

- CV-STG-BP24 – Avoid the use of Amazon EBS Data Lifecycle Manager and Amazon S3 Storage Lifecycle policies that relocate storage to infrequent access archive tiers.

- CV-STG-BP25 – Avoid the use of Amazon S3 Glacier Flexible Retrieval and Amazon S3 Glacier Deep Archive directly for backup or archive data as restores will require **multiple recalls** (Index, then Data recall). See CV-STG-BP11 for an alternative.

- CV-STG-BP26 – Avoid the use of S3 Intelligent-Tiering Opt-in asynchronous Deep Archive Access tier and Archive Access tiers, see CV-STG-BP11 for an alternative.

- CV-STG-BP27 – Consider the cost and latency of large-scale recalls from Amazon S3 Glacier Archive and Deep Archive storage classes. Often it is more cost-effective to **seal an archive** cloud storage location and redirect backups or archive plans to a new Commvault Combined Storage Tier location.

- CV-STG-BP28 – Use the **Amazon S3 Glacier Re:Freezer** serverless solution for copying an entire Amazon S3 Glacier vault archives to an Amazon S3 bucket and associated storage class. See **Migrating Data From Amazon S3 Glacier Vault to Amazon S3** for details on updating your Commvault data management platform after migration.

- CV-STG-BP29 – Avoid the use of **S3 Storage Lifecycle policies** with transition actions that move data to a storage class with a less frequent first byte latency. Instead, use **Commvault selective storage copies** to copy a

subset of data to the target storage class.

(*Use of storage lifecycle transitions within same first byte latency storage classes is permitted).*

- CV-STG-BP30 – Do not create or add additional Commvault mount paths to an existing Commvault cloud storage location to attempt increasing backup or recovery performance. There are no performance benefits to writing backups across multiple mount paths, you can safely ignore the following **Amazon S3 – Performance** guidance.

- CV-STG-BP31 – Do not enable **Multi-Factor Authentication (MFA) Delete** on Commvault Amazon S3 buckets, Commvault needs permission to delete objects when backup data retention expiry is reached.

> ⓘ **Note**
>
> Be aware that **S3 versioning** can be enabled on your Commvault buckets but it is not utilized by Commvault to perform rollback or recovery of Commvault cloud storage. Commvault *data aging* processes ensure that deletion requests remove all versions of an object.

Anti-patterns

- Using Amazon S3 archive tiers to store primary backup data with a high likelihood of recall.
- Using Commvault **extended retention rules** on Cloud Copies, which store long-term retention data alongside near-term backup data.
- Using an on-premises or edge-based MediaAgent to write *Primary backup copies* to Amazon S3. Any restores will incur Amazon S3 data egress costs.
- Using Amazon Glacier (Glacier direct, Glacier v1) instead of the current alternative, Amazon S3 Glacier.

Limitations

- Commvault writes data to Amazon S3 Intelligent-Tiering as a **Cold / Archive Tier** even though it primarily consists of frequent access storage classes with millisecond first-byte latency.
  - Object size for write operations is 32MB vs. 8MB for frequent access storage classes.
  - **Micro-pruning** is supported and enabled for Amazon S3 Intelligent-Tiering storage.
- Commvault will delete all versions of an Amazon S3 object when **Amazon S3 versioning** is enabled and data reaches an expiry age. Amazon S3 versioning cannot be used to recover or revert your Commvault cloud storage.
- When using Amazon S3 Object Lock, an authorized user can delete objects that are under compliance locks. Commvault has enhanced the **CloudTestTool** to assist in removing the **delete marker** on compliance-locked objects, to allow the removal of S3 delete markers, and to repair access to object-locked storage.

Additional Resources

- **Best practices for Amazon EC2** (includes Amazon EBS).
- **Amazon EBS direct APIs – Optimize performance**.
- **eBook: Security best practices and guidelines for Amazon S3**.
- **AWS re:Invent 2021 Breakout Sessions - Storage**.

Patterns

The following section provides common patterns employed when building your Commvault data management platform in AWS. Review patterns along with **Reference Architectures**, **Well-Architected** recommendations, and **Best practices**.

Backup on-premises directly to Amazon S3

As businesses adopt more public cloud services, there may be a desire to reduce the amount of owned and operated infrastructure. Removal of on-premises backup copies (secondary storage) allows the business to reclaim valuable data center space while leveraging the elasticity and low cost of cloud storage.
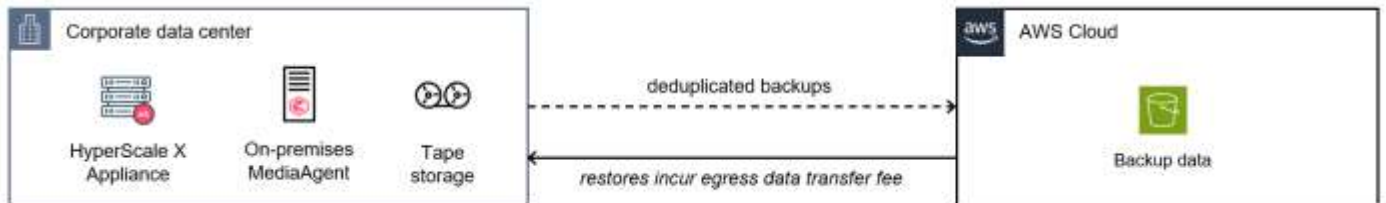
- Short-term operational recovery copies are held offsite in the Amazon S3 service.

- Leverages on-premises compute infrastructure (MediaAgent, MA) or Commvault HyperScale X appliances to optimize (deduplication, compress) data before the transfer.

- Allows businesses to completely remove secondary storage infrastructure from on-premises locations.

- All restores, synthetic full backups, and maintenance activities will incur minor egress charges.

This solution minimizes on-premises infrastructure by storing all backup copies offsite.

Backup on-premises to Amazon S3 with local backup

Cloud consumers often begin in the public cloud by extending their on-premises data center into the Cloud. In instances where the business has critical workloads still on-premises it is crucial to maintain rapid recoverability on-premises while leveraging the elasticity and low cost of cloud storage.
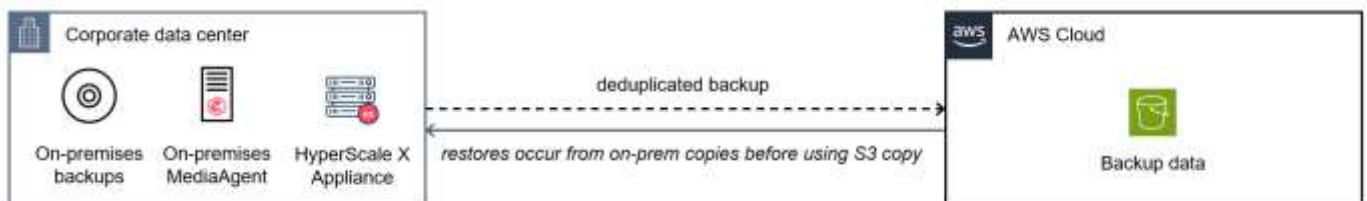
- Short-term operational recovery copies are held onsite disk/object/tape library (e.g., 7 - 30 days).

- Long-term operational and disaster recovery copies are held in the Amazon S3 service.

- Leverages on-premises compute infrastructure (MediaAgent, MA).

- No S3 egress cost unless restoring from Secondary.

This solution minimizes in-cloud infrastructure by leveraging on-premises compute resources.

Setup HA/DR for Access Nodes in a single availability zone

When designing your data management platform for resiliency from unplanned outages, you can create groups of redundant components called **Access Node groups**. Commvault will load-balancing and failover between healthy components in the event of one or more nodes becoming unavailable.
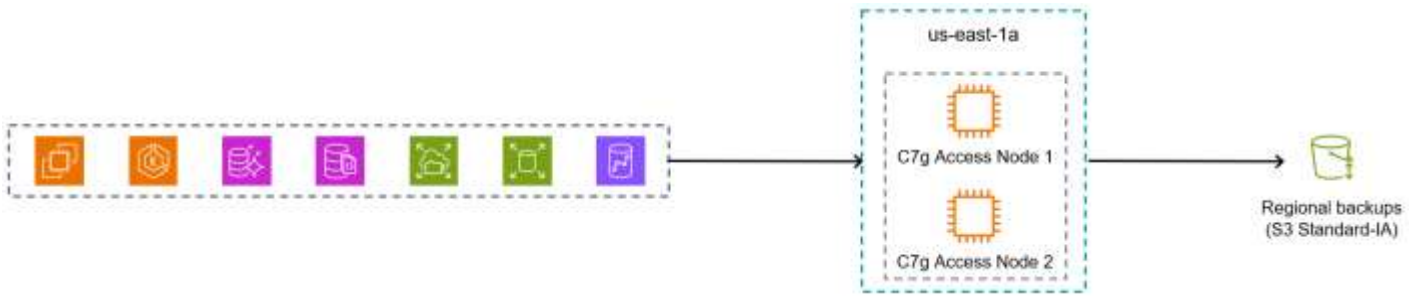
- Access Nodes (AN) are grouped within an Availability Zone (AZ) to provide load-balancing and high availability.

- Co-location of Access Nodes avoids in-region cross-availability zone data transfer fees.

- This solution protects workloads within the AZ at the least cost and may optionally protect other AZs with cross-AZ data transfer fees incurred.

This solution provides highly available, least cost, and least latency protection for a single availability zone.

Setup HA/DR for Access Nodes across multiple availability zones

Cloud services are distributed across availability zones to minimize the impact of planned or unplanned outages. High availability of data protection operations assists in ensuring data protection activities occur as scheduled and are unaffected by temporary planned/unplanned outages. The placement of data management infrastructure must be carefully considered to achieve the highest levels of availability.
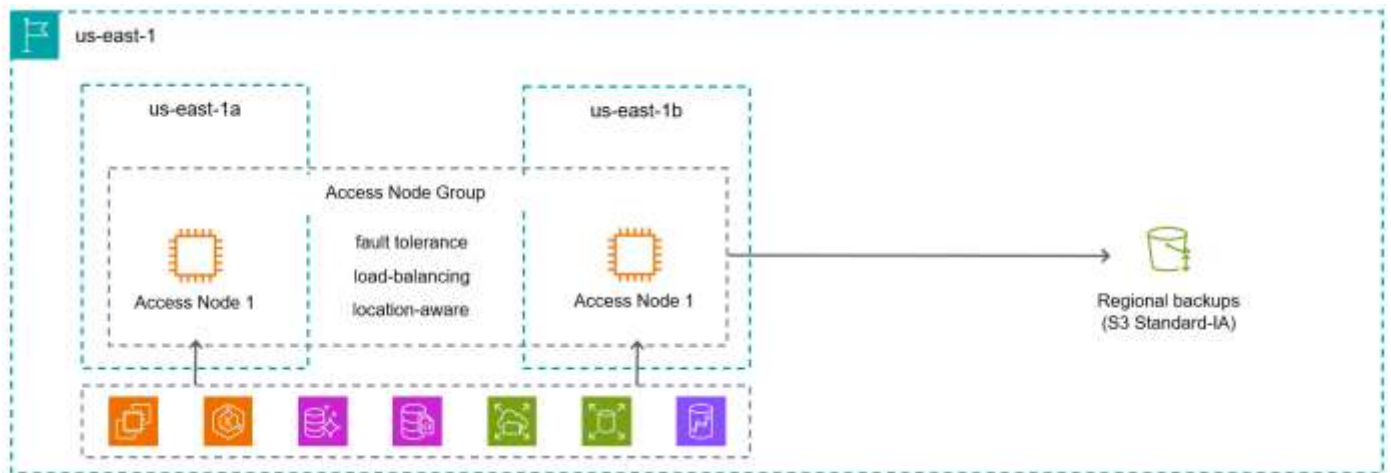
- Access Nodes (AN) are grouped across Availability Zones (AZs) to provide maximum availability from an AZ outage.

- Distribution of ANs across AZs will incur ongoing inter-AZ transfer fees for all data management activities.

- This solution protects workloads across all AZs within a region, with the acceptance that inter-AZ traffic will occur.

- This solution provides fault-tolerant protection for all availability zones within a region.

Setup CommServe HA/DR with Amazon Route 53 DNS failover

Commvault CommServe® LiveSync provides the ability to create active:passive deployments that distribute one or more passive CommServe instances across availability zones (depicted below) or Regions. Commvault implements health checks and automated failover when the **Active CommServe** becomes unavailable or enters Maintenance Mode. Use **Amazon Route 53 DNS failover** to automatically failover the Commvault customer-facing DNS entries

313

after a planned or unplanned failover. Note: Private resources require another pattern to failover private DNS entries (**Performing Route 53 health checks on private resources in a VPC with AWS Lambda and Amazon CloudWatch**).
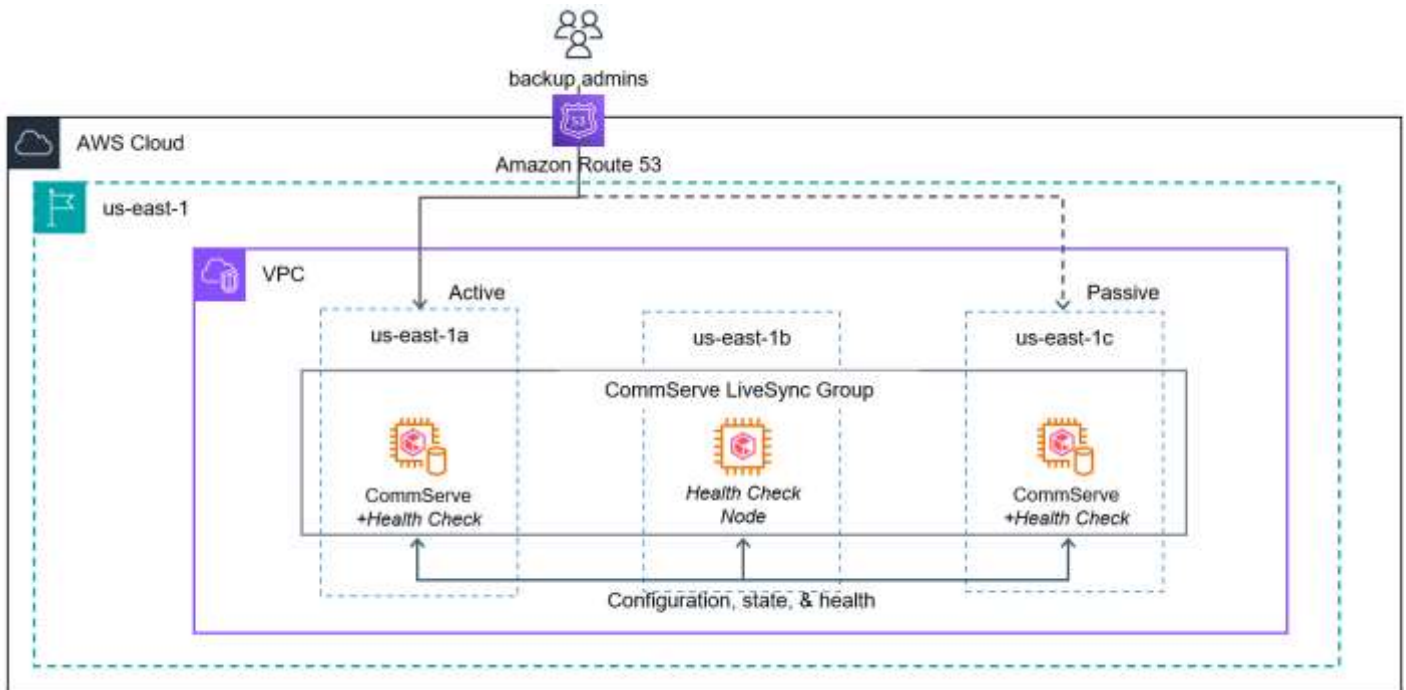
Archive and deduplicate data to Amazon S3

As businesses look to store more data for future **data analytics**, **visualization,** and **action** – a cost-effective data storage approach is required. Utilizing deduplication and compression to remove duplicate data before placing data in long-term storage can provide cost-optimal storage with minimal additional data handling.
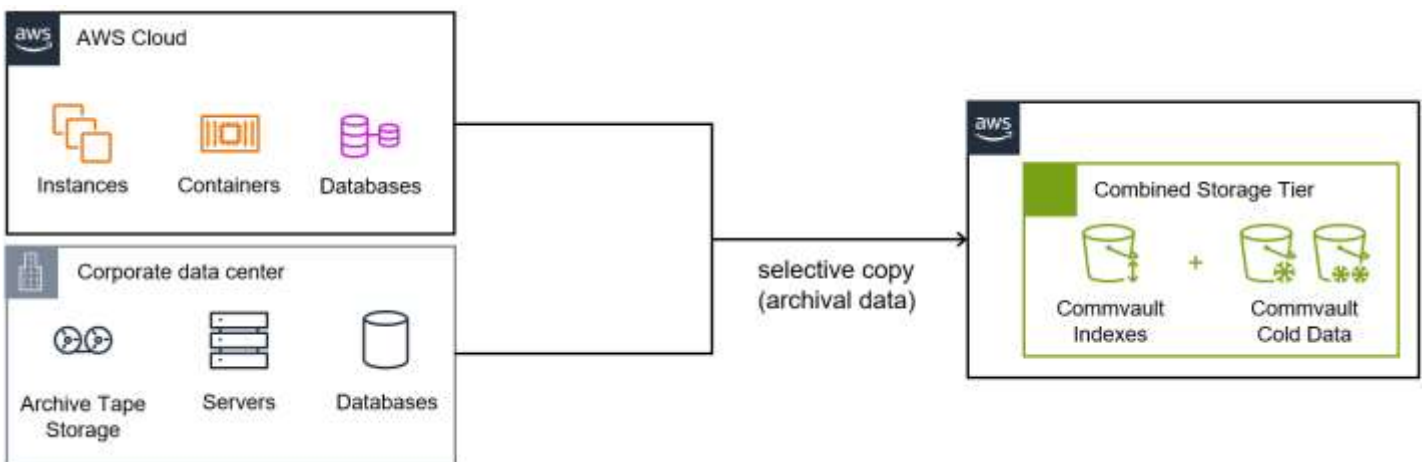
- Archival data represents a subset of backup data residing on-premises or in cloud storage services.

- Archival data is suitable for deduplication and compression (virtual machines, office documents)

- Archives are typically restored to their original operational location or a temporary file system for simplified search and retrieval.

- Indexes are stored in frequent access storage classes (S3 Standard, S3 Standard-Infrequent Access), and data is stored in an archive access storage class (S3 Glacier Flexible Retrieval, S3 Glacier Deep Archive, S3 Intelligent-Tiering Archive Access, and Deep Archive Access tiers).

- Data may be recalled as a simple <u>restore</u>, by leveraging accessible indexes in frequent access storage classes.

- This Solution provides long-term archive retention with an optimized recall process.

Archive data to Amazon S3

Businesses are retaining more data for historical **business insight** and **analytics**. Historically this meant expensive tape libraries, data preparation, and handling activities. Data archival to ultra-low-cost cloud archival services can now provide long-term retention without tape handling.
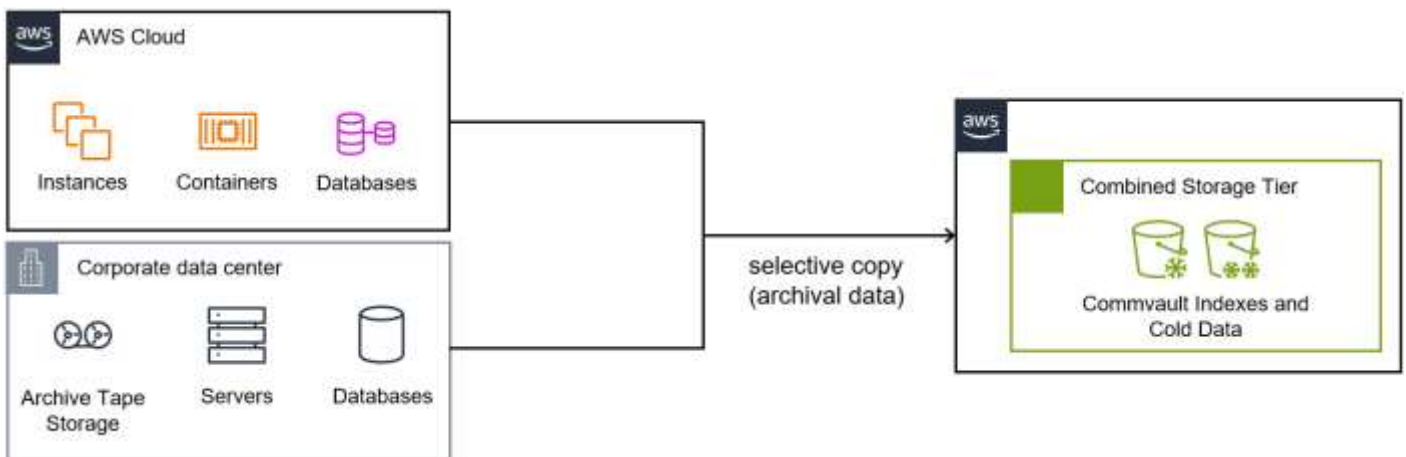
- Archival data represents a subset of backup data residing on-premises or in cloud storage services.

- Archival data is not suitable for deduplication or compression (lossless, x-rays, CAD, EDF archives)

- Very long-term retention data is required for regulatory compliance (i.e., age of patient + 10 years)

- Data is stored in its original unaltered format, recovery requires index recall, followed by data subset recall.

- This solution provides long-term data retention in the original application format.

Anti-Patterns

Performing data management for your traditional and modern workloads takes a thorough assessment of your application architecture and resiliency capabilities and needs. There are several fundamental changes that the durability, elasticity, and security of AWS cloud provide in your modern data management design. The **anti-patterns** below identify practices that were common on-premises but are no longer required or recommended in the AWS cloud.

- **Do not perform periodic media data verification**
  Amazon S3 storage provides eleven nines of durability by storing your data across multiple independent facilities. While Commvault provides the ability to perform **data verification**, this automated verification is disabled for Amazon S3 libraries, due to the API and data transfer costs that would be incurred performing the data verification on a scheduled basis.
- **Do not attempt to micro-manage cloud storage like random-access disk**
  Reclaiming storage space on-premises was critical to ensure the cost-effective utilization of limited resources. Amazon S3 has re-invented backup and archival storage with infinitely scalable and ultra-low-cost cloud storage.

Commvault will automatically manage and reclaim S3 storage using **micro-pruning** but based on the amount of data to be reclaimed and **Amazon S3 API throttling**, data aging may be delayed if your bucket is very active.

- **Do not store your primary backup for on-premises in the cloud, without analysis**
  Commvault can write your backup and archival data directly to Amazon S3 without the need to install, configure and maintain localized gateways or appliances. Be aware that storing your *primary backup copy* in Cloud will incur a **Data Transfer OUT From Amazon EC2 To Internet** or Data transfer over **AWS Direct Connect** data transfer fee.

- **Do not stretch infrastructure resiliency across regions**
  Amazon operates each AWS Region as a physically distinct and independent location with a minimum of three, isolated, and physically separate Availability Zones (AZs). Workloads can be deployed within a Region using the three AZs for high availability and resilience from a single facility outage, typically matching an organization's *disaster recovery* failure scenario. Cross-region resiliency is less common and incurs the latency of long-distance data transfers. Commvault recommends regional MediaAgent grids distributed across availability zones for high availability, do not span MediaAgents or Access Nodes across regions.

- **Do not scale vertically when scaling horizontally will deliver the same outcome**
  Historically scaling a workload in the data center meant upgrading or life-cycling the compute, storage, and network infrastructure every 3-5-8 years. This often led to an approach of vertically scaling applications within the budget of the individual line-of-business (LOB) and exceptionally large vertically scaled compute instances. Commvault recommends right-sizing data management compute to RPO/RTO demands, and scaling horizontally with smaller instances for improved resilience, load-balancing, and performance.

Use-Cases

Backup & Recovery

Commvault Backup Solutions lets you **rapidly recover** data **cost-effectively** and at scale, in AWS or edge-based data centers. Commvault unifies your data protection by protecting your cloud instances, containers, SaaS services, databases, storage, and traditional applications to Amazon S3.

How it works
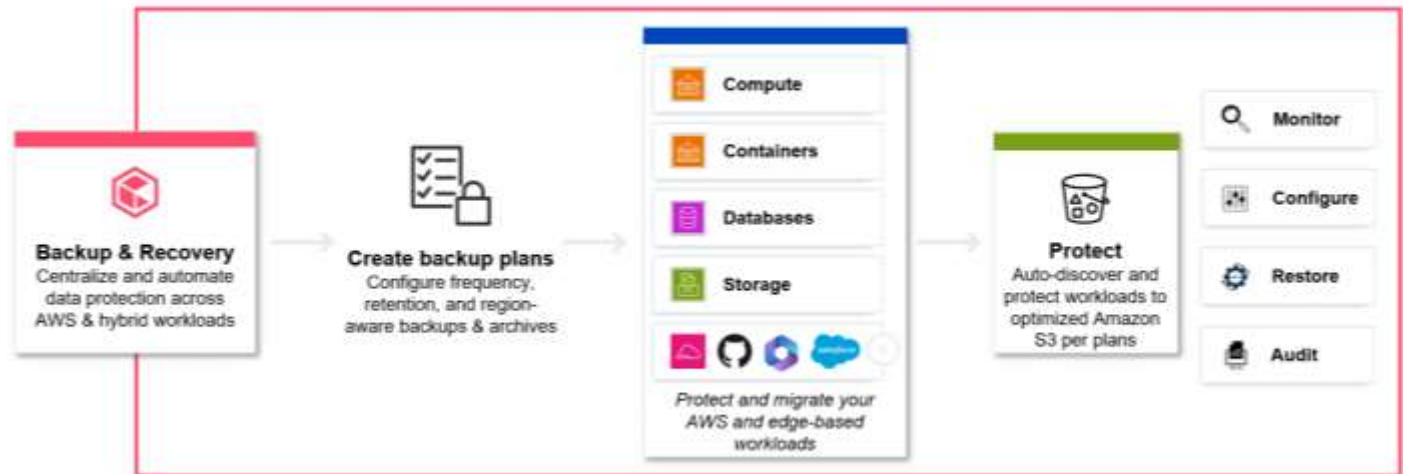
Use Cases

- **Complete unified backup & recovery**
  Back up <u>all</u> business data, including Cloud and edge-based Compute. Containers, Databases, Storage, SaaS, and traditional Applications. Recover across accounts and AWS Regions, AWS Local Zones, and AWS Outposts.
- **Anywhere to AWS disaster recovery**
  Quickly recover mission-critical operations by restoring virtual machines, databases, file systems, and object stores to AWS with configurable RTOs of minutes to hours. Delay resource creation for cost-reduced DR.
- **Cost-optimized cloud backups**
  Replace tape-based backup and archive stores with Commvault-optimized elastic, durable, and secure Amazon S3 cloud storage. Migrate large datasets offline with AWS Snow Family devices in network-constrained edges.

How to get started

You will require the following to get started with Commvault Data Protection to AWS:

- At least one AWS IAM machine-identity configured with an **IAM policy** with access to resources to protect.
- At least one MediaAgent/Access Node (in AWS or on-prem) to optimize and read/write data to Amazon S3.
- A network connection between your workload location and Amazon S3 (**AWS Direct Connect**, **AWS VPN**).
- Refer to **docs.commvault.com/** for a list of all protected workloads, including **AWS resources**.
- Refer to **Cloud-native backup with Commvault Backup & Recovery** reference architecture.
- Refer to **On-demand Disaster Recovery to AWS** reference architecture.
- ⓘ **Note**: Remote offices can write directly to Amazon S3 without a requirement for storage or tape gateways.

Cybersecurity

Commvault Data Security helps you **secure**, **defend**, and ensures ability to **recover** your production workloads and backup environment from ransomware and other threats. Commvault software is built on and enables you to adopt the best practices detailed in the NIST Cyber Security Framework.

How it works
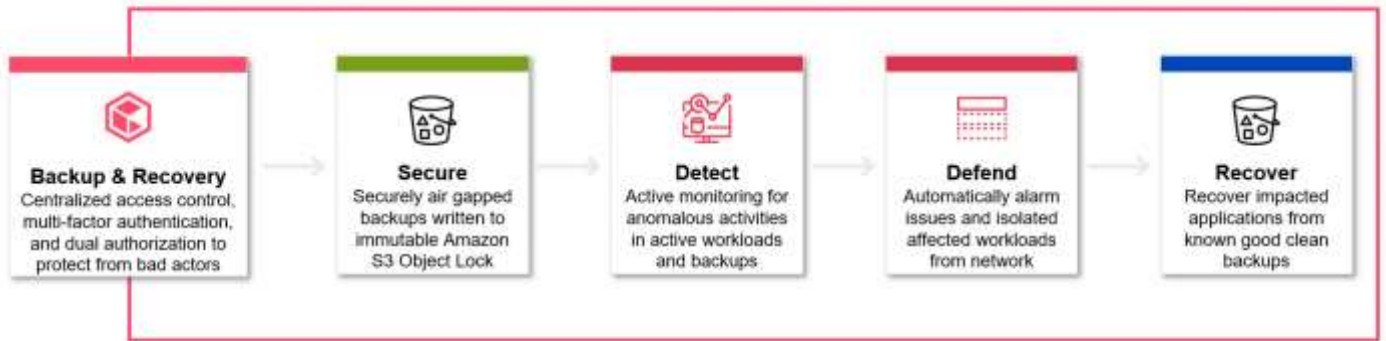
Use cases

- **Foundational security and data controls**
  Security of your backup data is job zero. Be confident that your most critical business data is kept secure through relentless pursuit of data security. Commvault employs a secure software development lifecycle (SSDLC) with static code analysis, static and dynamic vulnerability scanning, third-party penetration testing, and industry **certifications and compliance** validation. Commvault integrates natively with AWS IAM, AWS CloudHSM, and AWS KMS to allow secure authentication, authorization, and the ability to create and manage encryption keys. Additionally, Commvault keeps you in control of your data with **region-aware backup plans** that keep your data only where you allow it.

- **Automated and continuous vulnerability management**
  Reduce your mean time to remediate (MTTR) active vulnerabilities by automatically scanning, identifying, and quarantining malware in your backups with **Threat Analysis**. Once identified, malware can no longer propagate through recovery actions.

- **Intelligent threat detection and response**
  Expose threats quickly using anomaly detection, deep file extraction, file entropy, significant change variance, and signature-based threat detection with **Commvault Threat Scan**. Be notified in your centralized **SIEM** and SOAR platforms, or automatically respond to alarms by isolating potential threats and locking backups for forensic SecOps investigation.

- **Rapidly recover to a known good state**
  Accelerate your mean time to resolution by recovering of 'known good' securely air gapped backups that mitigate lateral moving threats. Good backups prevent modification by writing to immutable, encrypted, and proprietary format **Amazon S3 Object Lock protected** buckets.

How to get started

- Harden your CommServe with **CIS Level 1 benchmarks**, and enable **multi-factor authentication (MFA)** for all users.
- Harden your workloads by <u>encrypting everything</u> which slows attackers who do not have access to your KMS keys.

- Review your Security Posture in the **Security IQ Dashboard** and remediate non-compliant controls by implementing the recommendations and actions.
- Activate **Commvault Threat Scan** (including **Threat Analysis**, **File Data Analysis**) to proactively monitor, alert, and act on vulnerabilities and threats within your backup data and end-to-end data landscape.
- Activate **Cloud Command** to obtain your Data Protection Score, Platform Score, and Security Posture Score across your data protection activities, infrastructure health, and security control compliance.
- Enable business-logic workflows to require **dual-authorization** for high-risk changes or insider threats.
- Enable **S3 Object Lock** for business-critical backups requiring added protection from unintended modification.
- (optional) Implement an **Instance Scheduler on AWS** or a **blackout window** to limit access to backups.
- (optional) Consider alarming into Amazon CloudWatch, use **CloudWatch alarms** to automate response runbooks
- (optional) Implement Commvault Auto Recovery for rapid automated failover for mission-critical workloads.
- Review and implement the **Ransomware Protection Best Practices** from Commvault.
- Review the **Ransomware protection with Commvault Backup & Recovery** reference architecture.
- Review the **AWS Cloud Security – Protecting against ransomware** guidance for securing your AWS resources.

Works with

- Amazon EC2
- Amazon RDS Custom
- Amazon EFS
- Amazon FSx
- Amazon S3
- Amazon WorkSpaces

Requires file-system based backups using Windows, Linux or Macintosh-based file-system agent.

Cut IT Costs

Commvault software helps you **Cut IT Costs** by simplifying the choice of optimal AWS services and resources, accelerating build activities with Infrastructure as Code (IaC), and then automating ongoing data management with A.I. and ML for true automation at scale. Coupled with AWS Cloud Financial Management services, you can continually monitor, manage, and save in AWS with Commvault.
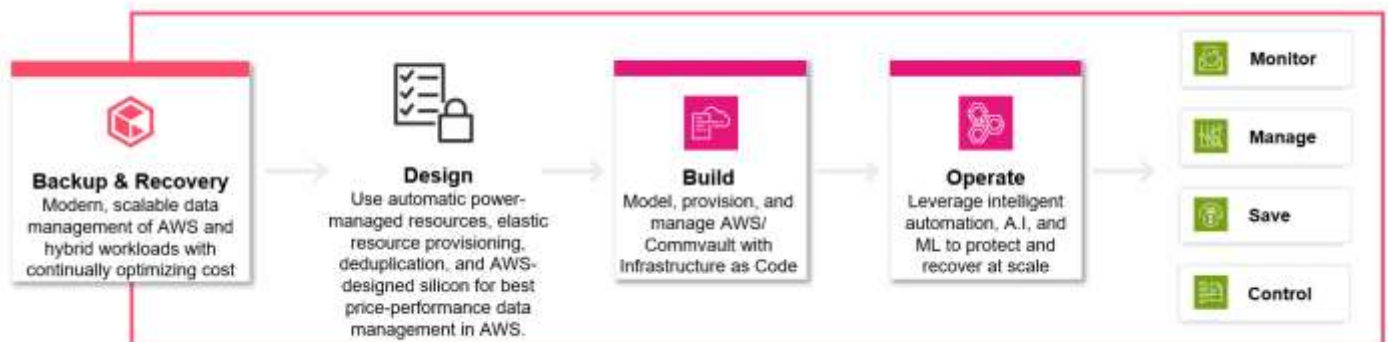
Use-cases

- **Design with cost optimal focus** by leveraging well-architected best practices from AWS and Commvault, like power-managed MediaAgents, auto-scaled Access Nodes, deduplicated storage, and the adoption of AWS Graviton for best price-performance compute in AWS.
- **Test, Deploy and Manage** your AWS infrastructure and Commvault using Infrastructure as Code (IaC) using AWS CloudFormation and the **Commvault + HashiCorp Terraform Module**.
- **Operate** your data protection environment with intelligent automation driven by A.I. and ML tuning scheduling, identifying vulnerabilities, threats, and anomalies, and acting.

Works with

All AWS services and hybrid workloads.

How to get started

Read the *AWS Cloud Architecture Guide – 2023e Edition* (this document) and start implementing the well-architected best practices for your data management environment.

Containers

Commvault software protects your re-factored and modern <u>containerized applications</u> with native Kubernetes integration to auto-discover and protect, replicate, and recover across regional **Amazon EKS** clusters and hybrid cloud locations (including **Amazon EKS on AWS Outposts**). Commvault leverages the Container Storage Interface (CSI) drivers to seamlessly provision, snapshot, and recover your persistent Amazon EBS block volumes.

Use-cases

- **Build and run modern applications** with complete protection for Kubernetes clusters, namespaces, applications, and persistent storage (block, file).
- **Protect your containerized applications** with cloud-native automatic discovery and protection of Amazon EKS applications and adjacent cloud analytics, database, and storage services.
- **Migrate and recover your applications** to any location where your Cloud Native Computing Foundation (CNCF) certified Kubernetes cluster is running.

Works with

- Amazon EKS
- Amazon EKS Anywhere
- Amazon EKS Distro
- Amazon EBS
- Amazon EFS*

- Amazon S3*
- Amazon Aurora*
- Amazon Redshift*
- Amazon DocumentDB*
- Amazon DynamoDB*

- All **CNCF-certified Kubernetes platforms**
- All **Production CSI drivers**

* Cloud-native adjacent persistent data services that are separately protected by Commvault Backup & Recovery.

How to get started

- Add your first **Amazon EKS cluster** and start protecting your entire cluster, select namespaces, or select applications.
- Configure application-consistent backup and recovery with **Commvault CvTasks Custom Resource Definition (CRD)**.
- Create **Commvault ResourceModifier Custom Resources** to control how application recovery and cloning is handled on and between your Kubernetes clusters.

Quick Recovery

Commvault helps you recover quickly with **Auto Recovery** (formerly Disaster Recovery) from any disaster (natural disaster, hardware failure, data breach, malware infection) that impacts your business. With automated replication for key compute, database, file, and applications – you will be ready to automatically recover with a single click.

Use Cases

- **AWS Region to AWS Region** data protection, workload-specific replication, and recovery for your AWS compute, storage, and database workloads.
- **On-premises to AWS** disaster recovery that leverages elastic nature of Amazon EC2 to auto-provision required compute resources only at failover time (i.e., **warm site failover**).
- **Cost optimized** disaster recovery preparedness with continuous replication, alerting, and automated failover runbooks that control application start-up order and priority.

Works with

- Amazon EC2
- Amazon EBS
- Amazon EFS
- Amazon RDS Custom
- Amazon S3

- AWS Local Zones*
- AWS Outposts*
- AWS Wavelength*
- File-systems (Windows, Linux)

- Oracle databases
- SAP HANA
- SQL Server databases
- VMWare Cloud on AWS
- Other hypervisors

* Includes Amazon EC2 and VMware Cloud compute protection at supported edge locations.

Governance, Risk, and Compliance

Commvault Data Compliance & Governance manages your data access to drive **regulatory compliance** and **mitigate** data **privacy risks**.

How it works
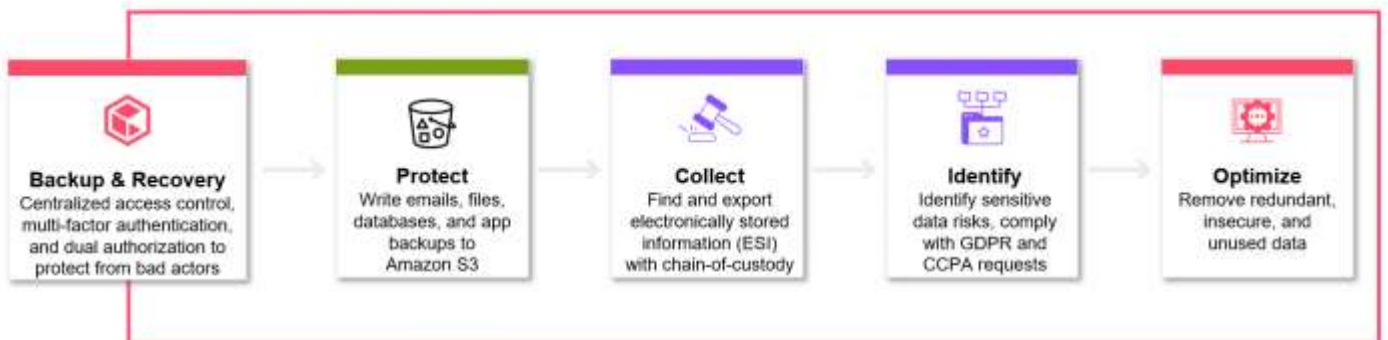
Use cases

- **Accelerated eDiscovery and compliance responses**
  Simplify and accelerate the search, tagging, export, and legal hold for eDiscovery and Compliance requests.
- **Respond to privacy regulatory requests**
  Respond to GDPR and CCPA Subject Access Requests (SARs) and Right to Access/Right to be Forgotten requests by identifying, extracting, or removing subject personally identifiable information (PII)
- **Optimize and secure unstructured data**
  Reduce sensitive data leakage risk and ongoing operational burden by identifying redundant, obsolete, and insecure files across live systems and historical backups, and remediate them by archiving, moving, or deleting them.

How to get started

- Review the **eDiscovery and Compliance** Getting Started guide.
- Review the Risk Analysis (**Data Governance**, **Storage Optimization)** Getting Start guides.

Works with

Data Source	eDiscovery & Compliance	Risk Analysis (Sensitive Data Identification)	Risk Analysis (Storage Optimization)
Database (Amazon RDS Custom)		●	
Exchange	●	●	
File Servers (Amazon EFS, Amazon FSx)	●	●	●
Endpoints (Amazon Workspaces)	●	●	●
Object Storage (Amazon S3)		●	●
OneDrive		●	
SharePoint Online		●	

Software Defined Storage

You can dramatically simplify your storage management and services by adopting **Software-defined storage** from Commvault. Software-defined storage powers **Commvault HyperScale X** backup appliances that provide scalable compute and storage for protection at the hybrid edge.

Use Cases

- **Protect all your workloads** in one location that scales to meet your data needs of today and tomorrow.
- **Safeguard your edge-based data** with built-in encryption, immutability, and ransomware protection.
- **Scale your data strategy from on-premise to Cloud** by leveraging optimized compute and storage to replicate critical data in Amazon S3 for cloud-based recovery and archival.

Works with

- Amazon EC2
- Amazon EBS
- Amazon EKS
- Amazon RDS
- Amazon S3
- Hybrid applications at the edge

How to get started

Reach out to your Commvault sales representative or partner and ask how you can modernize storage management on-premises with **Commvault HyperScale X™**.

Hybrid Cloud Services

Commvault Hybrid Cloud Services allows you to **seamlessly move data** across environments for **app modernization** & flexible data usage. Commvault enables a **seven R's migration strategy** including relocate, rehost, replatform, repurchase, and refactoring.

How it works
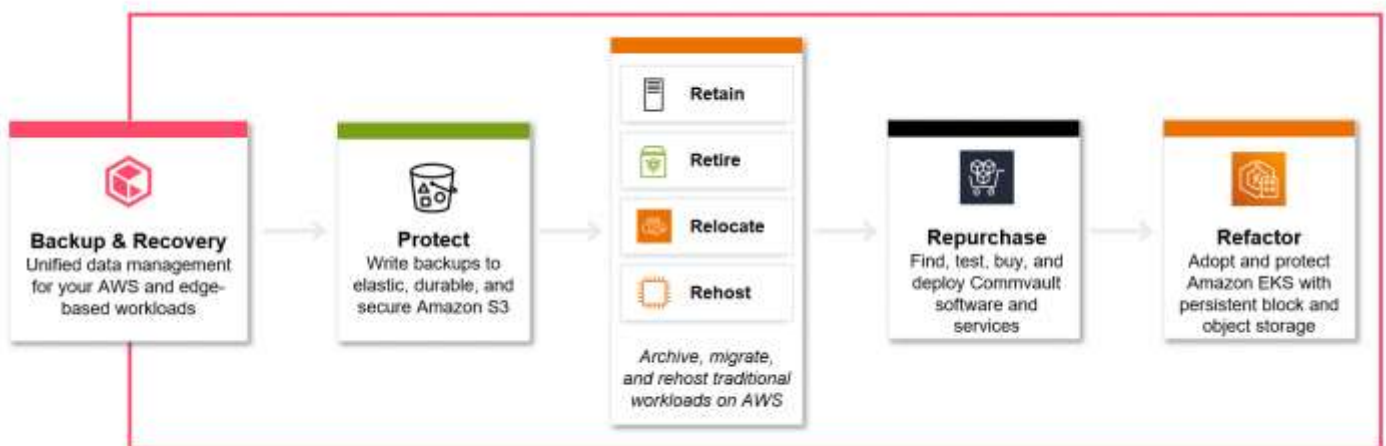
Use cases

- **Lift and shift workloads into AWS Cloud**
  Securely migrate and replicate on-premises virtual machines, databases, and storage to AWS fully-managed services in the AWS Region, AWS Local Zones, and AWS Outposts.
- **Drop and shop backup and recovery**
  Replace multiple point solutions and time-consuming scripting by testing, purchasing, and deploying your Commvault Backup & Recovery solution from the AWS Marketplace.
- **Protect and migrate cloud-native Kubernetes applications**
  Protect, migrate, and recover your containerized applications and persistent storage running on Amazon EKS.

How to get started

- Get started with a **Commvault Backup & Recovery BYOL** 60-day trial on AWS Marketplace.
- Migrate VMware VMs to VMware Cloud on AWS (VMC) using on-prem to AWS **Out of Place recovery**.
- Review compatibility for **Conversion to AWS** to rehost your existing VMs into fully managed Amazon EC2.
- Review options for migration of **Oracle**, **Cross-platform Oracle**, and **SQL Server** migration to Amazon RDS.
- Configure cloud-native **Kubernetes backup & recovery**, including **CSI driver** integration for containerized apps.

Works with

- Amazon EC2 (inc. EBS)
- Amazon EKS (anywhere)
- Amazon Aurora
- Amazon RDS

- Amazon DocumentDB
- Amazon EFS
- Amazon FSx
- Amazon FSx for NetApp

- Amazon S3
- VMware Cloud on AWS
- AWS Outposts
- AWS Marketplace

Risk Analysis

Commvault Risk Analysis applies A.I and machine-learning to your backups to help optimize and automate your cloud operations.

How it works
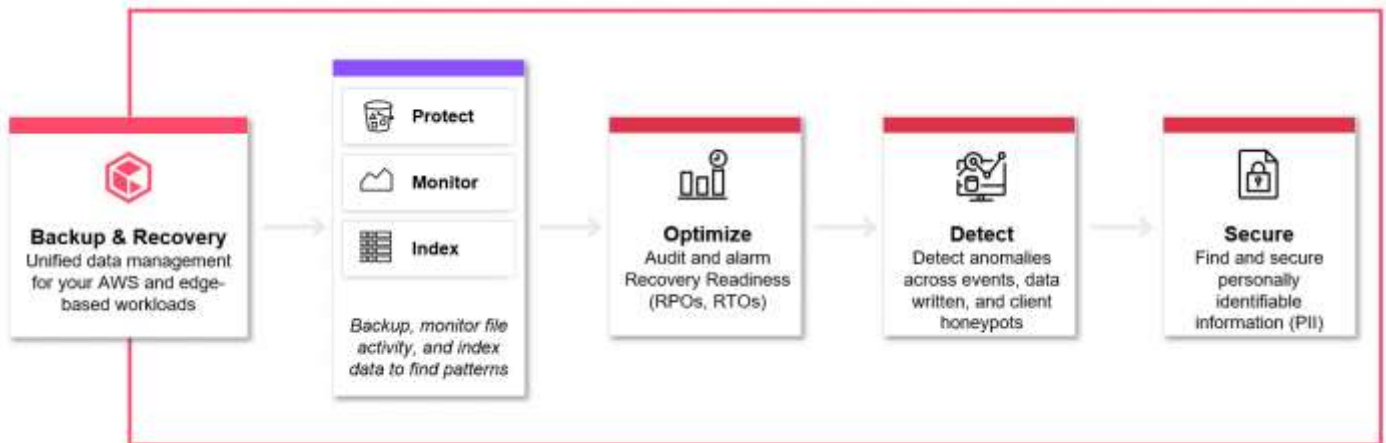
Use cases

- **Identify and optimize Recovery Readiness**
  Measure, monitor, and alarm your recovery readiness by measuring your achieved *recovery point objectives (RPOs)* and *recovery time objectives (RTOs)* against business needs.
- **Be alerted to anomalies to mitigate ransomware**
  Be alerted to anomalies across your AWS accounts and regions to proactively detect, alarm, and automatically respond to potential ransomware and malware activity.
- **Identify sensitive data and security misconfigurations**
  Identify personally identifiable data (PII) and ensure configurations before they become a data breach.

How to get started

- Enable protection for your AWS and edge-based workloads.
- Enable the **CommServe Anomaly Alert**, and forward it to your Ops, Service desk, or Amazon CloudWatch.
- Enable **honeypot monitoring** on clients to allow SecOps teams to respond quickly to malware events.
- Configure Content Analyzers & Data Sources, and access your **Data Governance dashboard** to view sensitive files.

Works with

- Amazon EC2
- Amazon EFS

- Amazon EKS
- Amazon FSx

- Amazon FSx for NetApp
- Amazon RDS

324

- Amazon S3
- Amazon Workspace

Virtualization or VM Backup

Commvault protects your virtualized compute workloads or **Virtual Machines (VMs)** from Amazon EC2 and VMware Cloud instances to hybrid Cloud VMs. Protect and recover your VMs on-premises, or replicate into Amazon S3 11 nine's durable storage for application recovery or whole data-center recovery to AWS.

Use cases
- **Holistic protection and recovery** of your virtualized compute workloads both on-premises and in cloud.
- **Self-service secure application recovery** for authorized administrators and application owners to recover their applications using Commvault Command Center, API, or command-line.
- **Lower the barrier to cloud migration** by allowing application owners to convert on-premises VMs and applications to native Amazon EC2 instances with a simple restore to AWS.

Works with:
- Amazon EC2
- Azure
- Google Cloud
- Oracle Cloud
- VMware (including VMware Cloud on AWS)
- Nutanix

How to get started

- Add your **hypervisor**j for protection and review your options for **automatic conversion to Amazon EC2** instances during recovery.
- Install **AWS Nitro drivers** on your on-premises VMs for seamless conversion to Amazon EC2.
- Prepare your VM **network configuration** and **file-system mount configuration** for conversion to Amazon EC2.

Backup and Recovery of AWS and hybrid resources

Commvault Backup & Recovery is a software solution that makes it simple to unify and automate data protection across clouds, containers, SaaS, and traditional workloads. Using Commvault Backup you can configure region-aware plans that automatically discover and protect your AWS resources and edge-based workloads. Backups include AWS-native snapshots and service-independent deduplicated snapshot copies, which allow you to reduce backup storage costs as data ages. Backups are held as snapshots and snapshot copies using independent encryption, creating an additional layer of defense for backups. Commvault Backup automatically monitors and auto-tunes protection to reliably meet your business RPO and RTO targets, as your environment evolves. You can use the Command Center console, API, CLI, or SDK to configure, operate, and perform reporting and auditing on your backups.
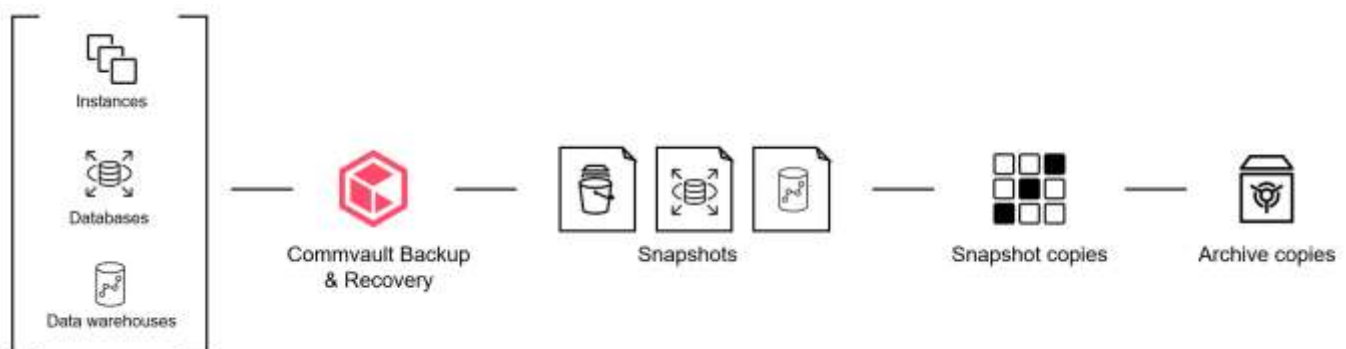
Commvault Backup understands your VMs, containers, applications, and databases and uses this information to enable migration between on-premises and AWS locations, including AWS Outposts. Backups may be replicated

between your cloud and edge locations to facilitate on-demand **Disaster Recovery** with RTOs ranging from minutes to hours.

Commvault has the broadest industry support for the backup and recovery of cloud and traditional workloads – see **www.commvault.com/supported-technologies** for protected workloads.

General information for all backups

Features available for all supported resources

Commvault Backup and Recovery provides the following key features for ALL supported AWS services and edge-based and SaaS-based protected workloads:

- **Automated backup schedules and region-aware protection**
  Commvault software provides *server plans* that automate protection by creating, replicating, and deleting cloud-native snapshots and service-independent backup copies across AWS regions and accounts.
- **Unified backup monitoring, alarming and insights**
  Perform centralized monitoring and alarming of your backup SLAs and end-to-end organizational *Recovery Readiness* across AWS regions and edge-locations, and optionally integrate with Amazon CloudWatch.
- **Full, Incremental, and Synthetic Full backups**
  Automatically create Full, Incremental, and zero-data transfer recovery points with Synthetic Full backups for your protected workloads.
- **Service-independent deduplicated backup copies**
  Create and replicate service-independent backup copies of AWS snapshots in workload-agnostic format for long-term and regulatory retention, and application mobility across AWS services and locations.
- **Commvault Combined Storage Tier backup and archive stores**
  Enable self-user user simplified and accelerated recall of protected workloads from Amazon S3 Glacier storage classes or long-term archives.
- **AWS KMS-integrated and independent encryption**
  Protect your KMS-encrypted workloads, encrypt your backups with KMS keys, and optionally encrypt your backups with your own keys via KMIP-compliant Key Management Servers.

- **Cross-account and cross-region data management**
  Protect your AWS workloads by copying backups across accounts and regions to protect backups from workload owner removal and or regional events.
- **Automated backup and recovery audit reporting**
  Generate audit logs on backup and recovery success/failure, user and administrator activity, and security hardening misconfigurations and remediations.
- **Immutable backup protection via S3 Object Lock**
  Create immutable data bunkers for very high-value workloads by writing backups with Amazon S3 Object Lock buckets.

Region and Availability Zone Support

Commvault Backup & Recovery may be deployed into and protect resources located in all 31 AWS Regions, Availability Zones, and edge locations. This includes GovCloud and China regions, and all associated Availability Zones, including AWS Local Zones, AWS Wavelength, and AWS Outposts.

Commvault Backup employs multiple protection methods including snapshots and network-streamed methods, if a particular API is unavailable in a location, Commvault can revert to network streaming protection.

> ⚠ **Note**
>
> Commvault does not support the use of FIPS service endpoints.

Commvault supports protection of supported workloads in the following 32 launched Regions (and their associated zones) with Platform Release 2023 and 2023e.

- us-east-1 (N. Virginia)
- us-east-2 (Ohio)
- us-west-1 (N. California)
- us-west-2 (Oregon)
- us-gov-east-1 (AWS GovCloud, US-East)
- us-gov-west-1 (AWS GovCloud, US-West)
- af-south-1 (Cape Town)
- ap-east-1 (Hong Kong)
- ap-northeast-1 (Tokyo)
- ap-northeast-2 (Seoul)
- ap-northeast-3 (Osaka)
- ap-south-1 (Mumbai)
- ap-south-2 (Hyderabad)
- ap-southeast-1 (Singapore)
- ap-southeast-2 (Sydney)
- ap-southeast-3 (Jakarta)

- ap-southeast-4 (Melbourne)
- ca-central-1 (Central)
- cn-north-1 (Beijing)
- cn-northwest-1 (Ningxia)
- eu-central-1 (Frankfurt)
- eu-central-2 (Zurich)
- eu-north-1 (Stockholm)
- eu-south-1 (Milan)
- eu-south-2 (Spain)
- eu-west-1 (Ireland)
- eu-west-2 (London)
- eu-west-3 (Paris)
- il-central-1 (Tel Aviv)[1]
- me-central-1 (UAE)
- me-south-1 (Bahrain)
- sa-east-1 (São Paulo

1: Support for Amazon S3 cloud storage libraries only, no native AWS workload backup & recovery.

See **Global Infrastructure Regions & AZs** for more information.

Features by protected resources

Commvault protects	Cross-region backup	Cross-account backup	Snapshot backup	Stream backup	Full Backup	Incr Backup	Item-level restore	Backup to any S3 class	Block-level deduplication
Aurora (via RDS protect)	✓	✓	✓	✓	✓	✓	✓	✓	✓
Cloud9 (via EC2 protect)	✓	✓	✓	✓	✓	✓	✓	✓	✓
CloudTrail logs (via S3)	✓	✓	✓	✓	✓	✓	✓	✓	✓
CloudWatch logs (via S3)	✓	✓	✓	✓	✓	✓	✓	✓	✓
DocumentDB	✓	✓	✓		✓				
DynamoDB	✓	✓		✓	✓	✓	✓	✓	✓
EC2	✓	✓	✓	✓	✓	✓	✓	✓	✓
EC2 on Outposts	✓	✓		✓	✓	✓	✓	✓	✓
EBS	✓	✓	✓	✓	✓	✓	✓	✓	✓
EFS	✓	✓		✓	✓	✓	✓	✓	✓
EKS, EKS-D	✓	✓	✓	✓	✓	✓	✓	✓	✓
EKS on Outposts	✓	✓	✓	✓	✓	✓	✓	✓	✓
FSx for Lustre	✓	✓		✓	✓	✓	✓	✓	✓
FSx for NetApp	✓	✓	✓	✓	✓	✓	✓	✓	✓
FSx for OpenZFS	✓	✓		✓	✓	✓	✓	✓	✓
FSx for Windows	✓	✓		✓	✓	✓	✓	✓	✓
RDS	✓	✓	✓	✓	✓	✓	✓	✓	✓
RDS on Outposts	✓	✓	✓	✓	✓	✓	✓	✓	✓
RDS Custom	✓	✓		✓	✓	✓	✓	✓	✓
Redshift	✓	✓	✓		✓			✓	✓
Red Hat OpenShift	✓	✓	✓	✓	✓	✓	✓	✓	✓
S3	✓	✓		✓	✓	✓	✓	✓	✓
S3 on Outposts	✓	✓		✓	✓	✓	✓	✓	✓
Storage Gateway	✓	✓		✓	✓	✓	✓	✓	✓
VMware on AWS	✓	✓	✓	✓	✓	✓	✓	✓	✓
VMC on Outpost	✓	✓	✓	✓	✓	✓	✓	✓	✓
Application -Aware	✓	✓	✓	✓	✓	✓	✓	✓	✓
WorkSpaces	✓	✓		✓	✓	✓	✓	✓	✓
Microsoft O365	n/a	n/a	n/a	✓	✓	✓	✓	✓	✓
Salesforce	n/a	n/a	n/a	✓	✓	✓	✓	✓	✓
GitHub, Azure DevOps	n/a	n/a	n/a	✓	✓	✓	✓	✓	✓

Server plans

In Commvault Backup & Recovery, a *server plan* represents your business policy for how often to perform backups, where to place the backups, and how many and how long to keep each copy. Backup plans are region-aware and will automatically direct backups to the appropriate region to enforce data residency needs. Commvault performs backup lifecycle management, which copies backups from snapshots to one or many service-independent copies stored in Amazon S3. You assign AWS and non-AWS workloads to plans to automate backups and backup copy lifecycle across all your hybrid workload locations.

Commvault Backup & Recovery stores your backups efficiently, using incremental snapshots (where supported by the AWS service) and then using Commvault block-level deduplication and compression to reduce storage and data transfer fees.

Commvault automatically manages your backup lifecycle and expires or *deletes* backups when they are no longer required, including snapshot removal.

Commvault uses A.I. and machine-learning (ML) to auto-tune the scheduling of protection and recovery operations to ensure your business RPOs and RTOs are met.

Commvault uses dynamic **rules** to discover and protect workloads, rules can be used to identify workloads by tags, region, availability zone, instance name, power-state, and many others. There are no limits on the number of rules and/or tags that can be used to select workloads for protection. Commvault will also recover tags that it protects during backup, ensuring dynamic discovery rules will auto-protect recovered and cloned workloads.

Ensuring all workloads are protected with default backups

Commvault Backup includes a default <u>safety net</u> called the *default subclient*. The *default subclient* is created for all workloads configured for protection and monitors for AWS resources that have not been included or excluded from backup coverage. If unprotected resources are identified, they are automatically protected with the *default server* plan, ensuring unintended data loss does not occur.

See **Adding Subclient Content** to discover your EC2 workloads for protection using tags, to ensure consistent protection based on your AWS Organizations **Tag Policies**.

Supported Amazon S3 storage classes

Commvault Backup can utilize all the available Amazon S3 storage classes to hold backups, long-term backups, and archives. You can tier your backups from EBS snapshots to S3 frequent access storage classes, to S3 infrequent access storage classes as data ages. Commvault provides a recall-optimized Commvault **Combined Storage Tier** which automates workload recall and recovery from archive storage classes.

Blackout windows

Blackout windows consist of a defined period and the data management operations that are not permitted to run during the blackout. If a data management operation is initiated or scheduled during the blackout window it will be queued and resumed when the window ends. By default, Commvault allows all data management operations to run 24x7.

See **Operations that support the blackout window** for more information.

AWS Regions and Zones

Commvault Backup protects and recovers your Commvault-supported resources deployed in any AWS Region (including GovCloud) and related Availability Zones. Additionally, edge-based locations that are serviced by AWS Local Zones. AWS Wavelength and AWS Outposts are also protected.

Federal Information Processing Standard (FIPS) 140-3 Support

Commvault Backup can utilize FIPS-140-3 validated ^{pending CMVP review} cryptographic modules when protecting workloads located in AWS US East/West, AWS GovCloud (US), or Canada (Central) regions that provide FIPs service endpoints.

See **FIPS Endpoints by Service** for more information.

Copying tags on backups

By default, Commvault Backup will copy tags from the resources it protects to copied *recovery points* and *restored resources*. For example:

- Tags that were originally associated with a protected Amazon EC2 instance are copied to its AMI and applied to any restored Amazon EC2 instances.
- Tags that were originally associated with a protected Amazon EBS volume attached to a protected Amazon EC2 instance are copied to its snapshots, and any volumes created from the snapshot.
- Tags applied to Amazon RDS, Amazon DocumentDB, Amazon DynamoDB, and Amazon Redshift instances are copied to their snapshots, and onto any instances created from their snapshots.
- Tags that were originally associated with a protected Amazon S3 object are applied to any restored object copies.

> ⚠ **Important Warning**
>
> Commvault <u>does not</u> protect tags associated with protected Amazon EFS, Amazon EKS, Amazon FSx*, Amazon Storage Gateway, Amazon FSx for NetApp online, or Red Hat on AWS (ROSA) resources.

Customizing tags during a restore

Commvault allows the customization of **AWS Resource Tags** during an out-of-place restore of full Amazon EC2 instances. Tags present during backup may be modified or removed. Additional tags may also be added.

Commvault will restore tags on Amazon EC2 instances, Amazon EBS volumes, Amazon VPCs, Subnets, and Security groups.

See **Set Tags on Restores VMs for AWS and VMware**.

AWS service	Tag backup	Tag restore	Tag customize (add, modify, delete) during a restore
Amazon Aurora	✓	✓	✗
Amazon DocumentDB	✓	✓	✗
Amazon DynamoDB	✓	✗	✗
Amazon EC2 (including Amazon EBS, Amazon VPC)	✓	✓	✓
Amazon EFS	✗	✗	✗
Amazon FSx	✗	✗	✗
Amazon RDS (including Amazon RDS Custom)	✓	✓	✗
Amazon Redshift	✓	✓	✗
Amazon S3	✓	✓	✗
Red Hat OpenShift on AWS (ROSA)	✗	✗	✗
VMware Cloud on AWS (VMC)	✓	✓	✓

Amazon EC2 support includes instances in AWS Local Zones, AWS Wavelength Zones, and AWS Outposts.

Amazon on RDS support includes instances on AWS Outposts, VMware Cloud, and VMware Cloud on Outposts.

Amazon S3 support includes buckets on AWS Outposts.

Creating snapshot backups across AWS accounts and regions

You can use the cross-account management feature to manage and monitor the creation of backups and recover across AWS accounts and regions. Your *server plans* can include one or more **regional mappings** that indicate the desired source and destination region for AWS-native snapshots, based on the workload region.

After the source snapshot has reached the `created` status, Commvault will automatically initiate a `CopySnapshot` activity to the destination Region and account. Commvault supports **multi-region KMS keys** for consistent encryption across regions, but may also re-encrypt snapshots with destination account KMS keys identified by the aliases `cvlt-master`, `cvlt-ec2`, or `cvlt-rds`.

Commvault performs cross-region and cross-account snapshot copying for **Amazon EBS**, **Amazon RDS**, and **Amazon Redshift** workloads.

Be aware that there are resource quotas on the maximum number of concurrent snapshots copies per service – see **AWS service quotas** for current limits.

Using cross-account backup shared resources

By default, Commvault will seek data management resources (Access Nodes, MediaAgents) within the AWS account under protection. In large multi-account multi-region deployments, Commvault infrastructure may be centralized in a **central backup account**. During backup, *recovery points* are created in the workload account and then shared or copied to the central backup account. This approach allows increased controls on the high-value backup account, while reducing the EC2 runtime costs to provide multi-account data management. During restores, this process is reversed, with *restored resources* being created within the central backup account, then shared with the workload account for workload deployment.

See **Using resources from a central backup account**.

Compute

Amazon EC2 protection

Commvault Backup & Recovery supports unified and automated data protection for your Amazon EC2 instances, Amazon EBS volumes, and associated network and security configuration. You can backup instances located in the AWS region and edge locations. Backups include the creation of Amazon Machine Images (AMIs) and associated EBS snapshots and service-independent snapshot copies. You can recover full EC2 instances, EBS volumes, and item-level files and folders across AWS accounts, regions, and zones.

Commvault Backup automatically discovers and protects new EC2 instances using AWS resource tags. *Server plans* dictate how many snapshots and service-independent snapshot copies are created. Snapshots may be shared and copied across AWS accounts and regions for increased protection from regional events.

There are no limitations on the protection and recovery of Amazon EC2 instances by instance type/family or size.

IAM permissions for creating and restoring backups

Commvault Backup requires an IAM identity (user or role) with the Commvault-supplied **Amazon EC2 IAM policy** attached to allow backup and recovery actions. You can remove actions for use-cases that you will not be using, per **Amazon Web Services Permission Usage**. You can also set a **PassKey** within Commvault as an additional control to ensure the identity performing a restore is authorized.

Commvault Backup supports and recommends using the Security Token Service (STS) `AssumeRole` action to obtain temporary security credentials from each protected workload account during data management activities. Use of STS:AssumeRole aligns with a well-architected secure-by-default approach by using temporary credentials that expire and require periodic renewal.

You should create **Permission boundaries for your IAM entities** by enhancing your Commvault IAM role with restricted Commvault access to AWS resources using **tags**. Your permission boundaries will be maintained for all resource copies as Commvault protects, recovers, and allows customization of resource tags during EC2 instance recovery.

See **Using Security Token Service (STS) AssumeRole for backup and recovery**.

Backup consistency

Commvault Backup, by default, captures **crash-consistent** backups using the `ec2:CreateImage` and `ebs:CreateSnapshot` actions. You can optionally select an **Application-Aware** backup type, which will auto-discover and push application agents to the EC2 instance to perform application-consistent backups (per **Application supportability**).

> ⓘ **Note**
>
> Commvault does not protect **user data** used during the launch of Amazon EC2 instances (Linux or Windows).

Reducing backup and restore time with Changed Block Tracking (CBT)

By default, Commvault Backup uses **Amazon EBS direct APIs** to determine the used blocks within Amazon EBS snapshots during full and incremental backups. Additionally, Commvault performs restores, replication, and disaster recovery by restoring EBS volumes using EBS direct APIs.

> ⓘ **Note**
>
> To use EBS direct APIs for accelerated incremental backup, a snapshot must remain in the workload account between backups to allow the EBS direct API to identify changed blocks.

Snapshots and Snapshot copies

Commvault Backup, by default, protects Amazon EC2 instances by creating a snapshot copy that consists of the used blocks for each protected EBS volume, to Commvault-optimized Amazon S3. Commvault identifies the used and changed blocks within each volume snapshot using the **Amazon EBS direct APIs**, which reduce backup time by up to 80% compared to traditional methods.

The Amazon EBS snapshots may be optionally retained after the backup by enabling **IntelliSnap®** snapshot management for the protected instances. EBS snapshots provide a rapid recovery point for EC2 instances and can be used in your backup lifecycle as the *primary recovery point* for EC2 instance recovery.

Use-case	EBS Direct API support
Streaming backup (with Amazon Access Node)	✓
Streaming backup (with on-premises Access Node)	✓
Snapshot backup (**IntelliSnap®**)	✓
Live Browse	✓
Cross-region backup copy from the snapshot	✓
Amazon EC2 Full Instance Restores	✓
Amazon EBS Volume Restores	✓
On-premises VM to Amazon EC2 **conversion**	✓
Disaster Recovery periodic VM to Amazon DC2 **replication**	✓

All use-cases function within multi-tenant environments

Encryption during backup and restores

By default, Commvault inherits the encryption settings of the protected instance during protection operations. Amazon EBS snapshots that are taken from encrypted volumes are automatically encrypted using the source volume data encryption key (per `ec2:CreateSnapshot`). Volumes that are created from encrypted snapshots are automatically encrypted using the source snapshot data encryption key (per `ec2:CreateVolume`). You can configure your account to enforce the encryption of new EBS volumes and snapshots by enabling automatic **account-level EBS encryption**.

See **Amazon EBS encryption** for more details.

Performance

In general, you can expect the following backup rates with Commvault Backup:

- 156 GB/hr. for snapshot copies taken via a c7g.xlarge instance using Amazon EBS direct APIs.

- 183 GB/hr. for restores via an Amazon c7g.xlarge using Amazon EBS direct APIs.
- There is no maximum duration to a backup activity.

You can vertically scale the Access Node instance size used to obtain additional network credits and bandwidth to meet your unique business RPOs and RTOs.

Alternatively, you can install operating system and application agents inside your EC2 instances to perform application-consistent snapshots, and snapshot copies over the production instance network.

On-demand backups

Using either the Command Center Console, API, CLI, or SDK, you can initiate an on-demand backup of a single resource (Instance) or a group of resources (VM Group). When you initiate an on-demand backup, Commvault will determine the appropriate storage location per your configured *server plan*.

Auto-scaling backup resources

By default, Commvault Backup will utilize your Commvault CommServe or MediaAgent as an Access Node to communicate with the AWS service endpoints or VPC endpoints for backup and recovery. As data volume grows, you can enable **automatic scaling of access nodes** to deploy required resources during backup, then terminate them when backup activities are complete.

> ⓘ **Note**
>
> Auto-scaled resources are used for backup and backup copy operations only. Restores will use the MediaAgent as the Access Node, if the Virtual Server Agent package is installed.

Concurrent backups and service limits

Commvault Backup limits backups to one concurrent backup per VM group. You can run multiple concurrent backups separated by AWS account, region, and selected resources. You should be aware that EC2 snapshot protection is dependent on the **Amazon EBS service quotas** for concurrent snapshot operations and EBS direct API operations.

Restoring a full EC2 instance

Using either the Command Center Console, API, CLI, or SDK, you can restore one or more Amazon EC2 instances to the original AWS account, region, and availability zone. Alternatively, you can restore to another AWS account, region, or zone including edge locations. To restore to a different EC2 instance type, EBS volume type, or re-encrypt with another KMS key, simply select the required parameters during restore.

Commvault will create an EC2 instance with provided Region, Availability Zone, Instance type, Instance size, and EBS volume type and encryption preferences. Nested or dependent resources will be recreated if not available in the destination account, including Elastic Network Interfaces (ENIs), Security Groups, and AWS KMS Keys.

Commvault supports supplying or customizing the following `ec2:RunInstance` parameters during instance restore:

- Instance name (Tag)
- Placement (Availability Zone)
- InstanceType
- VolumeType (all Volumes)

- KmsKeyId (all Volumes)
- SubnetId (Amazon VPC)
- NetworkInterface.N
- SecurityGroup.N

- SecurityGroupId.N
- TagSpecification.N (Instance, Volumes)

> ⓘ **Note**
>
> Commvault does not support supplying or customizing the following parameters to `ec2:RunInstances`. Commvault will create a new instance with these parameters matching the original instance. Use Commvault-created AMIs to manually deploy an instance using the EC2 console, or AWS CLI to modify these parameters:
>
> - AdditionalInfo
> - CapacityReservationSpecification
> - CpuOptions
> - CreditSpecification
> - DisableApiStop
> - DisableApiTermination
> - DryRun
> - EbsOptimized
> - ElasticGpuSpecification.N
> - ElasticInferenceAccelerator.N
> - EnclaveOptions
>
> - HibernationOptions
> - IamInstanceProfile
> - ImageId
> - InstanceInitiatedShutdownBehavior
> - InstanceMarketOptions
> - Ipv6Address.N
> - Ipv6AddressCount
> - KernelId
> - KeyName
> - LaunchTemplate
> - LicenseSpecification.N
>
> - MaintenanceOptions
> - MaxCount
> - MetadataOptions
> - MinCount
> - Monitoring
> - PrivateDnsNameOptions
> - PrivateIpAddress
> - RamdiskId
> - UserData

> ⓘ **Note**
>
> Commvault does not support supplying a private IPv4 address during instance restore. To assign a specific static IP to an instance, **Create a network interface** with a Private IPv4 address and select the network interface during restore.

AWS Resource tags attached to the original EC2 instance and EBS volumes will be restored allowing consistency of tag-based operations and automation for your restored instances. Additionally, when restoring an EC2 instance out-of-place, you may customize the tags by adding, modifying, or deleting tags that were collected during the backup of the instance.

> ⓘ **Note**
>
> Amazon EC2 Spot Instances will be restored as on-demand instances.

Auto-discovering, protecting, and recovering Amazon VPC resources

You can automatically discover and protect the following Amazon VPC resources as part of Amazon EC2 backup, and replication. Commvault will record the presence and configuration settings of these resources to allow forensic investigation of failed changes.

- Amazon VPCs
- Public subnets
- Private subnets
- Security groups
- DHCP Option sets
- Internet gateways
- VPN gateways

- Flow logs
- Endpoints
- Routers
- NAT gateways
- Elastic Network Interfaces (ENIs)

- Network Access Control Lists (NACLs)
- VPN connections
- Peering connections
- Carrier gateways
- Customer gateways

- Transit Gateway attachments

See **Amazon VPC Resources That Commvault Protects** for the most current information on what resources are protected.

> ⓘ **Note**
>
> Amazon VPC protection is enabled by default but dependent on the ability to create, read, update, and delete Amazon VPC resources. See the `amazon_restricted_role_permissions.json` for details. The VPC protection permissions are in the `VPCBackupPermissionsIfFeatureEnabled` and `VPCRestorePermissionsIfFeatureEnabled` blocks.

Commvault will check for the following Amazon VPC resources during the recovery of a full Amazon EC2 instance, and recreate the resource if not found in the destination account:

- Amazon VPCs
- Public subnets
- Private subnets
- Security groups
- Elastic Network Interfaces (ENIs)

See **Amazon VPC Resources That Are Restored by a Full In-Place Restore** for the most current information on which resources are recreated during a restore. It should be noted that a new Amazon VPC will receive a new VPC ID, Commvault records the historical VPC ID as an AWS Resource Tag on the restored restore.

> ⓘ **Note**
>
> As of Platform Release 2023e, restore of Amazon VPC resources is only supported to the original AWS account and Region. Out-of-place restores or 'cloning' of Amazon EC2 instances and associated Amazon VPCs is not supported.

You can also recover the network configuration for a protected Amazon EC2 instance or instances by selecting the instances, then selecting the **Instance configuration files** restore option. Commvault will restore a JSON file with the network configuration at the time of backup to the file-system of a Commvault client.

Instance configuration file restores are crucial to performing a forensic investigation of unsuccessful changes, failed automation, or malicious changes that impact your application availability. You can quickly restore your applications to a known-good running state, while investigating the changes made, and then investigate who made the changes using AWS CloudTrail.

Avoiding Amazon EBS volume initialization with HotAdd recovery

You can perform a full Amazon EC2 instance and Amazon EBS volume restores to newly created Amazon EBS volumes, which provides maximum performance without the delay of **initializing the EBS volume**. Select **Commvault HotAdd** transport type during recovery and new volumes will be mounted on an in-region, in-zone Access Node and populated with data from the backup. Once restoration completes, the volume(s) will be attached to the desired EC2 instance, already pre-warmed.

> ⓘ Note
>
> If the Access Node performing the restore resides in an alternate Availability Zone, the volume will require pre-warming initialization.

> **ⓘ Note**
>
> If cross-account shared backup resources are being used to perform recovery, Amazon EBS volumes will be created from a snapshot and require pre-warming. Deploy a Commvault Access Node into the destination account and availability zone if avoidance of volume initialization is needed.

Performing item-level restore for EC2 instances

Recovery time is crucial when business services are unavailable, and restoring item-level resources like files, folders, or volumes <u>without</u> the need to recreate a new instance *speeds recovery*. Commvault Backup will restore files and folders to the original instance, or another instance by securely staging files in Amazon S3, and then downloading them to the host using **Amazon Systems Manager Run Command**. Additionally, individual Amazon EBS volumes may be restored and attached to instances for application-owner-driven recovery.

This recovery approach is particularly useful in environments managed by Amazon Managed Services (AMS), which require that the Amazon EC2 instance ID be maintained during recovery events.

Database-consistent backup of Oracle databases on EC2 instances

Commvault provides snapshot-based incremental forever protection for Oracle and Oracle RAC instances running on Amazon EC2 instances. You can protect your large Oracle instances using Oracle Automatic Storage Management (ASM), Linux Logical Volume Manager (LVM), or raw devices using native Amazon EBS snapshots. Commvault will orchestrate database-consistent EBS snapshots and will only store incremental differences between each snapshot. This approach provides faster backup and recovery than traditional RMAN streamed database backups.

See **Forever incremental backups for Oracle databases hosted on Amazon EC2**.

VMware Virtual Machine protection

Commvault Backup and Recovery supports unified and automated data protection for VMware virtual machines (VMs) residing in VMware Cloud™ (VMC) on AWS and VMware Cloud™ (VMC) on AWS Outposts. You can utilize the VMware vSphere Storage APIs (formerly VADP APIs) to back up and restore VMs located in the AWS Region and edge locations, including on-premises clusters.

You can recover full VMware VMs, VMDK volumes, and item-level files and folders across AWS accounts, regions, and zones. You can perform Application-aware backups and use your backups to migrate workloads between locations or replicate them for on-demand Disaster Recovery to VMC.

See **VMware Cloud™ on AWS**.

Permission for creating and restoring backups

Commvault Backup requires a VMC vCenter user with **vSphere permissions** for the data management use-cases you would like to perform (i.e., Snapshot backup and restores, Streaming backup and restores, Replication).

Supported VM types and transport mode

Commvault Backup can backup and restore virtual machines on VMware ESXi 4.1 or later, 5.0.x, 5.1.x, 5.5, 5.5.1, 5.5.2, 5.5.3, 5.5.6, 6.0, 6.0.1, 6.0.2, 6.0.3, 6.5, 6.7, 6.7.1, 6.7.2, 6.7.3, 7.0, 7.0.1, 7.0.2, 7.0.3 running on NFS, VMFS, and VSAN datastores.

Commvault Backup supports SCSI Hot-Add, Network Block Device Secure Sockets Layer (NBDSSL), SAN, and NAS Transport modes to copy data from on-premises source VMs to backups.

Commvault Backup supports Raw Device Mapping (RDM) and independent disk protection using **IntelliSnap®** backup when protecting on-premises VMware clusters. During restoration, these disks will be converted to virtual RDMs, dependent disks, or VMDKs (see **Transformations for backup copies and restores**).

> ⓘ **Note**
>
> Only SCSI Hot-Add and NBDSSL transport modes are supported for use with VMware Cloud™ on AWS.

Auto-protecting workloads by tag

Commvault Backup will auto-discover and protect VMware VMs by **rules**. **Rules** automatically identify VMs by tag, attributes, VM name/pattern, and many other metadata fields exposed by vCenter at backup runtime. Tags are copied to restored VMs allowing consistency of tag-based management in VMC Cloud™.

Backup consistency

Commvault Backup, by default, captures **File-system and application-consistent** backups using vmtools within the guest to quiesce the file-system and any VSS-aware applications. You can optionally select an **Application-Aware** backup type, which will auto-discover and push application agents to the EC2 instance to perform application-consistent backups (per **Application supportability**). Alternatively, you may select **Crash-consistent** backups which take a VMware storage snapshot without coordinating with the VM.

Reducing backup and restore time with Change Block Tracking (CBT)

By default, Commvault Backup will enable and use **VMware Changed Block Tracking (CBT)** to read the allocated and modified portions of VMware virtual disks. Changed Block Tracking reduces backup time by up to 80% compared to traditional methods.

Restoring a full VM

Using either the Command Center Console, API, CLI, or SDK, you can restore one or more VMware VMs to the original AWS account, region, and availability zone. Alternatively, you can restore to another AWS account, region, or zone including edge locations. When performing a full VM restore, you may provide or modify the following parameters:

- VM name
- VMC cluster
- ESXi server or host
- Datastore
- Resource pool
- VM folder path
- Network
- IP address
- Volume encryption key

Commvault will restore one or more VMware VMs within provided Region, Availability Zone, VMC, or VMware cluster per the provided parameters and optionally power on the VM. Commvault will retain the source VM's instance UUID if it is not already present in the destination cluster.

> ⓘ **Note**
>
> Commvault restores disks with the original provisioning method. You may optionally select the preferred disk provisioning method during restore, including **Thick Lazy Zero, Thin**, or **Thick Eager Zero**.

vSphere tags and attributes and VM storage policies associated with the source VMs will be copied to the restored VMs allowing consistent tag-based operations as workloads migrate. Additionally, when restoring a VMware VM out-of-place (i.e., a copy or clone), you may add, delete, or modify the tags that will be applied to the new VM.

Performing item-level restore for VMware VMs

Recovery time is crucial when business services are unavailable, and restoring item-level resources like files, folders, or volumes underline the need to recreate a new VM *speeds recovery*. Commvault Backup will restore files and folders to the original VM, or another instance by using vmtools in the guest to perform an **agentless recovery**. Additionally, individual Virtual Machine Files (VMDKs) or volumes may be restored and attached to VMs for application-owner-driven recovery.

See **Restoring Guest Files and Folders** for details.

Additional resources

- **System Requirements for Amazon EC2 protection**.
- **System Requirements for VMware Virtual Machine protection**.
- **IAM permissions for creating and restoring Amazon EC2 backups**.
- **Auto-scaling Access Nodes for on-demand backup resources**.

Containers

Amazon EKS protection

Commvault Backup & Recovery provides unified and automated data protection for your Kubernetes applications running on Amazon EKS. You can backup containers located in the AWS region and edge locations running Amazon EKS Anywhere and Amazon EKS Distro (EKS-D). Backups include the protection of Kubernetes manifests, namespaced and non-namespaced resources, and persistent volumes. Protection automates the creation of storage snapshots using the CSI driver, then copies application data to a service-independent copy in Amazon S3. You can recover complete namespaces, applications, and persistent volumes, including item-level files and folders across your EKS clusters. Combined with AWS Cloud Databases and Amazon S3 protection, Commvault provides complete protection for your modern containerized applications.

Commvault Backup automatically discovers and protects new namespaces, applications, and persistent volumes using *label selectors*. *Server plans* dictate how many streaming backups and optional backup copies are created. Backup copies may be replicated across AWS accounts and regions for increased protection from regional events.

Any persistent storage volume presented via a **Container Storage Interface (CSI) driver** that supports dynamic provisioning and snapshots is supported.

Supported Amazon EKS releases

Commvault Platform Release 2022e supports Kubernetes `1.24`, `1.23`, and `1.22` only.

Commvault integrates directly with the EKS Kubernetes API server for backup and recovery and does not require any AWS IAM permissions or **EKS actions** to perform protection.

Ensure your **Amazon EKS Kubernetes version** is a supported Commvault release.

> ⓘ **Note**
>
> Protection of etcd is not supported on Amazon EKS Distro (EKS-D).

Full and Incremental Backups

Commvault Backup supports Full, Incremental, and Synthetic Full backups of Amazon EKS applications. Application manifests (metadata) are protected in their entirety with each backup, regardless of backup type. Incremental backups compare the last modified timestamp of files and protect any files modified since the last backup.

Backup consistency

Commvault Backup, by default, captures crash-consistent backups of application persistent storage using the underlying storage snapshot capability (i.e., `ec2:CreateSnapshot` action). You may optionally implement application consistency by implementing pre- and post-execution scripts to prepare your application for the storage snapshot. Commvault provides supported scripts for MySQL, PostgreSQL, MongoDB, and Cassandra.

See **Implementing Application-Consistent Backups for Kubernetes**.

Restoring namespaces and applications

Restoring namespaces and applications may occur as a *destructive restore* (in place) or a *non-destructive restore* (out of place). A *destructive restore* will remove the resource(s) to be restored if they exist in the destination EKS cluster, then restore from backup. A *non-destructive restore* will restore the resource(s) to an alternative cluster, namespace, or application name to avoid impacting the production application.

Commvault will use the registered CSI driver to provision volumes during restoration.

See **Kubernetes – Restores**.

Restoring item-level objects

Recovering item-level objects includes restoring application manifests, etcd snapshots, and application files and folders collected from persistent volumes (PVs) attached to the EKS application. When you perform application manifest or etcd snapshot restores, you select one or more files and specify the full restore path on your Kubernetes protection Access Node (i.e., `/user/ec2-user/my-eks-manifests`). When you perform application file recovery, you select files and folders to restore and then select a running EKS application, an attached persistent volume, and the full restore path in the volume (i.e., `/usr/local/my-containerized-app`). You may also restore application files and folders to the file-system of your Kubernetes Access Node.

Restores may be *destructive* or *non-destructive*. A *destructive* restore will delete the resource before attempting recovery. A *non-destructive* restore will restore to a new application name, or namespace, or fail if restore already exists.

Red Hat OpenShift on AWS (ROSA) protection

Commvault Backup & Recovery provides unified and automated data protection for your Kubernetes applications running on Red Hat OpenShift Service on AWS (ROSA). ROSA provides fully-managed production-ready OpenShift in AWS with self-service provisioning, automatic security enforcement, and streamlined deployment. You can backup containers located in the AWS region and edge locations running Red Hat OpenShift on-premises. Backups include the protection of Kubernetes manifests, namespaced and non-namespaced resources, and persistent volumes presented via production CSI drivers.

See **Amazon EKS protection** for details of Kubernetes application protection and recovery capability.

> ⓘ **Note**
>
> Red Hat OpenShift on AWS is not a supported or tested platform at the time of writing (**solution 5552181**).

Additional Resources

- **System Requirements for Amazon EKS protection** (including Amazon EKS Anywhere, Amazon EKS Distro).
- **Amazon EKS on Outposts protection**.
- **Creating a Kubernetes service account for Commvault Backup & Recovery**.

Database

Amazon Aurora protection

Commvault Backup & Recovery provides unified and automated data protection for your Amazon Aurora clusters, including serverless and provisioned resources. You can backup instances located in the AWS region and edge locations. Amazon Aurora is a fully-managed relational database management system (RDBMS) built for the cloud with full MySQL and PostgreSQL compatibility. Commvault protects multi-AZ database clusters and single-instance databases using snapshots and database-consistent dumps.

Commvault Backup automatically discovers and protects new Aurora clusters and/or instances using AWS resource tags. *Server plans* dictate how many snapshots, snapshot copies, or database dumps are created. Snapshots may be shared and copied across AWS accounts and regions for increased protection from regional events.

Commvault copies both Multi-AZ DB cluster snapshots and single-instance DB snapshots cross-account and cross-region.

See **Multi-AZ deployments** for details on creating high-availability DB deployments with automated failover.

IAM permissions for creating and restoring backups

Commvault Backup requires an IAM identity (user or role) with the Commvault-supplied **Amazon Aurora IAM policy** attached to allow backup and recovery actions. You can also set a **PassKey** within Commvault as an additional control to ensure the identity performing a restore is authorized.

Commvault Backup supports and recommends using the Security Token Service (STS) `AssumeRole` action to obtain temporary security credentials from each protected workload account during data management activities. Use of STS:AssumeRole aligns with a well-architected secure-by-default approach by using temporary credentials that expire and require periodic renewal.

You should create **Permission boundaries for your IAM entities** by enhancing your Commvault IAM role with restricted Commvault access to AWS resources using **tags**.

See **Using Security Token Service (STS) AssumeRole for backup and recovery**.

Auto-protecting instances by tag

By default, Commvault will configure a *default instance group* that protects all Amazon Aurora instances in all regions. You may optionally use *instance name wildcards* (i.e., my-finance-mysql-db*) or *instance tag values* to auto-discover and protect instances. Tags may include Key only or Key and Value matches. Wildcards are not permitted for instance tag discovery.

Instance and tag-based matching may only be configured after the *instance group* is created, by editing the content selection and adding rules.

> **ⓘ Note**
>
> In Multi-AZ DB clusters *instance name* and *instance tag* matching are performed on the Writer Instance only. Applying tags to the Regional Cluster or Reader Instance will not detect instances to protect.

Encryption during backup and restores

By default, Commvault inherits the encryption settings of the protected Amazon Aurora instance during protection operations. Amazon Aurora snapshots that are taken from encrypted Aurora instances are automatically encrypted using the source cluster data encryption key. Aurora instances that are created from encrypted snapshots are automatically encrypted using the source snapshot data encryption key.

See **Encrypting Amazon Aurora resources** for more details.

Copying snapshot backups cross-region and cross-account

Commvault can protect your Amazon RDS snapshot backups by taking a full copy out of the workload account into your central backup account. Additionally, Commvault can create copies of your RDS snapshots in one or many regions (fan-out) to provide disaster recovery copies to protect from regional events. Commvault can combine these controls and provide **cross-account and cross-region copies** that maintain encryption using **multi-region keys** or re-encrypted the instance in the destination region.

To replicate a copy of encrypted RDS snapshots, the destination AWS account must either have an AWS KMS key with an **alias** or **tag** name of `cvlt-rds` or `cvlt-master` in the destination region.

> **ⓘ Note**
>
> The IAM identity performing the backup operation must have access to perform backup and recovery of Amazon RDS instances in the source and destination regions and be listed as a KMS Key User for the destination region.

> **ⓘ Note**
>
> Tags attached to the source DB snapshot are copied to the destination AWS account or the regional DB snapshot copy.

Database-consistent dump

Commvault Backup can perform logical service-independent full backups of Amazon Aurora databases so that long-term retention copies or database migration may occur between cloud and on-premises. Database-consistent dumps use the database native dump or export utility to capture the entire instance. Dump-based backups may be performed with an ARM 64-bit or x86 64-bit Access Node (the access node does not need to match the architecture of the database being protected). Commvault recommends using **AWS Graviton base Access Nodes** (i.e., C7g, C6g) for the best price-performance and sustainability of long-term retention dump-based backups.

Dump-based backups may be used to restore to the original instance, or another pre-provisioned instance, or perform table-level restores. Backups are stored in a Commvault cloud storage location (i.e., typically an Amazon S3 bucket) with independent encryption to the source DB instance, providing another level of protection for your backups.

> **ⓘ Note**
>
> Amazon Aurora does not provide a method to perform transaction log backup and recovery.

Commvault Backup can use your existing Access Nodes to perform dump-based protection by installing the relevant Commvault database agents (MySQL, PostgreSQL) on your existing Access Node. Dump-based protection will utilize TLS 1.1/1.2 for all communication to and from the Access Node to the database endpoint (see **TLS 1.1/1.2 support for Amazon Aurora and Amazon RDS PostgreSQL**, **TLS 1.2/1.2 support for Amazon Aurora and Amazon RDS MySQL and MariaDB**).

Restoring an Amazon Aurora cluster from a snapshot

Using either the Command Center Console, API, CLI, or SDK, you can restore an Amazon Aurora DB cluster or DB instance to the original AWS account, region, and availability zone. Alternatively, you can restore to another AWS account, region, or zone including edge locations, if you have copied the snapshot to the destination region.

Commvault will create an Amazon Aurora instance within provided Region, Availability Zone, Instance type, Instance size, and EBS volume type and encryption preferences. Nested or dependent resources will be recreated if not available in the destination account, including Elastic Network Interfaces (ENIs), Security Groups, and AWS KMS Keys. Commvault restored Aurora instances clusters and/or instances will inherit the AWS tags from the original protected cluster as Commvault passes `CopyTagsToSnapshot = ` **`true`** when restoring clusters with **rds:RestoreDBClusterFromSnapshot** or **rds:RestoreDBInstanceFromDBSnapshot**.

> ⓘ **Note**
>
> Commvault does not permit changing the **DB instance class** during restoration. Restored instances will inherit the DB instance class (Serverless, Memory optimized provisioned, and Burstable provisioned) of the source instance.

Commvault supports supplying or customizing the following `rds:CreateDBcluster` or `rds:CreateDBInstance` parameters during restore:

- AvailabilityZones.AvailabilityZone.N
- DatabaseName / DBName

- DBClusterParameterGroupName
- DBSubnetGroupName

- MultiAZ
- PubliclyAccessible

> ⓘ **Note**
>
> Commvault does not support supplying or customizing the following parameters to `rds:CreateDBCluster` or `rds:CreateDBInstance` Commvault will create a new instance with these parameters matching the original instance. Use Commvault-created DB snapshots to manually deploy an instance using the RDS console, or AWS CLI to modify these parameters:

- AllocatedStorage
- AutoMinorVersionUpgrade
- BacktrackWindow
- BackupRetentionPeriod
- BackupTarget
- CharacterSetName
- CopyTagsToSnapshot
- CustomIamInstanceProfile
- DBClusterIdentifier
- DBClusterInstanceClass
- DBInstanceIdentifier
- DBParameterGroupName
- DBSecurityGroups.DBSecurityGroupName.N
- DeletionProtection
- Domain
- DomainIAMRoleName
- EnableCloudWatchLogsExports.member.N
- EnableCutomerOwnedIp
- EnableGlobalWriteForwarding
- EnableHttpEndpoint
- EnableIAMDatabaseAuthentication
- EnablePerformanceInsights
- Engine
- EngineMode
- EngineVersion
- GlobalClusterIdentifier
- Iops
- KmsKeyId
- LicenseModel
- MasterUsername
- MasterUserPassword
- MaxAllocatedStorage
- MonitoringInterval
- MonitoringRoleArn
- NcharCharacterSetName
- NetworkType
- OptionGroupName
- PerformanceInsightsKMSKeyId
- PerformanceInsightsRetentionPeriod
- Port
- PreferredBackupWindow
- PreferredMaintenanceWindow
- PreSignedUrl
- PromotionTier
- ReplicationSourceIdentifier
- ScalingConfiguration
- ServerlessV2ScalingConfiguration
- StorageEncrypted
- StorageType
- Tags.Tag.N
- TdeCredentialPassword
- Timezone
- VpcSecurityGroupIds.VpcSecurityGroupId.N

ⓘ **Note**

Recovery of Aurora clusters may only occur from an in-region snapshot. If a **cross-region** restore is required, ensure the *server plan* is configured with a region mapping and **snapshot replication** is enabled in the instance group.

Restoring an Amazon Aurora cluster from a database-consistent dump

Commvault Backup can use your service-independent database-consistent dump to restore to the original Amazon Aurora instance, to another pre-provisioned Amazon Aurora instance, or a compute instance running a compatible version of your database engine. You can restore the entire instance or individual databases. Restoring to the original instance is often faster than deleting and creating an entirely new instance. Restoring to the original instance maintains instance identifiers, critical in infrastructure as code (IaC) and managed services environments.

A compute instance running the same database engine version *or higher* is considered compatible.

You can restore user-defined databases. System databases are not restored.

⚠ **Note**

When you create an Aurora MySQL instance, the software automatically creates a database that is called innodb. This database does not include a format file, so the tables are not listed. Make sure that you do not restore the innodb database.

Amazon DocumentDB protection

Commvault Backup & Recovery provides unified and automated data protection for your Amazon DocumentDB clusters. You can back up the primary instance of your DocumentDB clusters located in the AWS Region. Amazon DocumentDB (with MongoDB compatibility) is a fast, scalable, highly available, and fully managed document database service that supports MongoDB workloads. Commvault protects DocumentDB clusters using regional snapshots to allow recovery to a new DocumentDB instance.

Commvault Backup automatically discovers and protects DocumentDB clusters to protect using AWS resource tags. *Server plans* dictate how many snapshots, snapshot copies, or database dumps are created.

> ⓘ **Note**
>
> Commvault does not perform *cross-account* or *cross-region* snapshot replication of Amazon DocumentDB clusters.

IAM permissions for creating and restoring backups

Commvault Backup requires an IAM identity (user or role) with the Commvault-supplied **Amazon DocumentDB IAM policy** attached to allow backup and recovery actions.

Commvault Backup supports and recommends using the Security Token Service (STS) `AssumeRole` action to obtain temporary security credentials from each protected workload account during data management activities. Use of STS:AssumeRole aligns with a well-architected secure-by-default approach by using temporary credentials that expire and require periodic renewal.

You should create **Permission boundaries for your IAM entities** by enhancing your Commvault IAM role with restricted Commvault access to AWS resources using **tags**.

See **Using Security Token Service (STS) AssumeRole for backup and recovery**.

Auto-protecting instances by tag

By default, Commvault will configure a *default instance group* that protects all Amazon DocumentDB clusters in all regions. The default instance group is a safety net that ensures that <u>every</u> DocumentDB cluster is always protected. You may optionally use *instance name wildcards* (i.e., docdb-production-*) or *instance tag values* to auto-discover and protect instances. Tags may include Key only or Key and Value matches. Wildcards are not permitted for instance tag discovery.

> ⓘ **Note**
>
> If your IAM user or role is restricted to a subset of AWS Regions and the **All clusters in all regions** scope is selected (default setting), the **PREVIEW** button will not function, and backups will succeed. Consider modifying your content selection to include **specific regions** (add clusters button) or **tags** (add cluster rule button).

Instance and tag-based matching may only be configured after the *instance group* is created, by editing the instance group content and adding rules.

Instances must be in `available` state when backup discovery occurs.

> ⓘ **Note**
>
> *Instance name* and *instance tag* matching are performed on the Primary Instance only. Applying tags to the Regional Cluster or Replica Instances will not detect instances to protect.

Encryption during backup and restores

Amazon DocumentDB snapshots are encrypted with the same KMS Key Id as the source database. Amazon DocumentDB clusters deployed from Commvault-created snapshots are encrypted with the KMS Key Id of the source snapshot.

Snapshot backups

Commvault performs Amazon DocumentDB **full backup** by automating the creation and deletion of **Manual Cluster Snapshots** and managing them per your assigned *server plan*. Backups utilize the `docdb:CreateDBClusterSnapshot` action which records the following information about your cluster and cluster snapshot:

- availabilityZones
- dBClusterSnapshotIdentifier
- dBClusterIdentifier
- snapshotCreateTime
- engine
- engineMode
- allocatedStorage
- status

- port
- vpcId
- clusterCreateTime
- masterUsername
- engineVersion
- licenseModel
- snapshotType
- percentProgress

- storageEncrypted
- kmsKeyId
- dbClsuterSnapshotArn
- IAMDatabaseAuthenticationEnabled
- tagList

> ⓘ **Note**
>
> Snapshots created by Commvault will have three tags applied with the following example values
> **Name** = SP_2_1015_2
> **Description** = Snapshot_created_by_Commvault_for_job_1015_Source_DocumentDB_docdb-2022-10-31-21-57-37
> **_GX_BACKUP_** = *null*

> ⓘ **Note**
>
> Commvault does not protect Amazon DocumentDB parameter groups.

Restoring an Amazon DocumentDB cluster from a snapshot

Using either the Command Center Console, API, CLI, or SDK, you can restore your Amazon DocumentDB backups from a regional snapshot to a new instance with a different Cluster identifier (name) and Instance class. When performing a restore from a snapshot using `rds:RestoreDBClusterFromSnapshot`, you can customize the following parameters with Commvault:

- DBClusterIdentifier
- DBInstanceClass
- ~~Tags.Tag.N~~
- DBSubnetGroupName
- VpcSecurityGroupIds.VpcSecurityGroupId.N

> ⚠ **Warning**
>
> Commvault will not restore the AWS Resource tags that were discovered on the original DocumentDB instance during its backup..

When performing a restore from a snapshot, the **DB subnet group** will be pre-populated from the source database, use the AWS CLI to list the available DB subnet groups in the target account and region:

```
aws rds describe-db-subnet-groups
{
    "DBSubnetGroups": [
        {
            "DBSubnetGroupName": "default-vpc-bd19ddd4",
            "DBSubnetGroupDescription": "Created from the RDS Management Console",
```

When performing a restore from a snapshot, the **number of instances** will default to one. Please ensure you enter the total number of instances you would like your DocumentDB cluster to contain.

When performing a restore from a snapshot, the **VPC security group IDs** field will be blank, it will not be populated from the source database. Use the AWS CLI to list the available security groups in the target AWS account. You may list multiple security groups separated by a comma.

```
aws ec2 describe-security-groups
{
    "SecurityGroups": [
        {
            "GroupId": "sg-0b45b05c1f70941f4",
```

> ⓘ **Note**
>
> Commvault does not support supplying or customizing the following parameters to `rds:RestoreDBClusterFromSnapshot` and subsequent `rds:RestoreDBInstance` calls. Commvault will create a new instance with these parameters matching the original instance. Use Commvault-created DB snapshots to manually deploy an instance using the RDS console, or AWS CLI to modify these parameters:
>
> - SnapshotIdentifier
> - DBClusterIdentifier
> - DBInstanceClass
> - Number of instances
> - Virtual Private Cloud (VPC)
> - DBSubnetGroupName
>
> - VPC security groups Ids
> - Port
> - KMS Key Id
> - Enable Performance Insights
> - Enable CloudWatch Logs Exports
>
> - Tags
> - Enable Deletion Protection

Amazon DynamoDB protection

Commvault Backup & Recovery provides unified and automated data protection for your Amazon DynamoDB tables. You can back up your DynamoDB tables located in the AWS Region. Amazon DynamoDB is a fully managed, serverless, key-value NoSQL database designed to run high-performance applications at any scale. DynamoDB offers built-in security, automated multi-Region replication, in-memory caching, and data import and export tools.

Commvault protects DynamoDB tables by downloading all records via the Amazon DynamoDB API and storing them in an independently encrypted Commvault Amazon S3 Cloud Library. Creating a service-independent backup allows Commvault backup copies to be retained over the 35-day retention limit provided by DynamoDB stand-alone backup.

Commvault Backup automatically discovers and protects DynamoDB tables to protect by region, Instance wildcards matching, and AWS resource tags. *Server plans* dictate how many backups and backup copies to create.

> ⓘ **Note**
>
> Commvault DynamoDB protection does not utilize stand-alone on-demand **DynamoDB backup and restore** capabilities. Commvault does not utilize DynamoDB **Exports to S3** to provide backup and recovery.

IAM permissions for creating and restoring backups

Commvault Backup requires an IAM identity (user or role) with the Commvault-supplied **Amazon DynamoDB IAM policy** attached to allow backup and recovery actions. You can also set a **PassKey** within Commvault as an additional control to ensure the identity performing a restore is authorized.

Commvault Backup supports and recommends using the Security Token Service (STS) `AssumeRole` action to obtain temporary security credentials from each protected workload account during data management activities. Use of STS:AssumeRole aligns with a well-architected secure-by-default approach by using temporary credentials that expire and require periodic renewal.

You should create **Permission boundaries for your IAM entities** by enhancing your Commvault IAM role with restricted Commvault access to AWS resources using **tags**.

See **Using Security Token Service (STS) AssumeRole for backup and recovery**.

Auto-protecting instances by tag

By default, Commvault will configure a *default instance group* that protects all Amazon DynamoDB tables in all regions. You may optionally use *instance name wildcards* (i.e., prod-analytics-dynamodb-*) or *instance tag values* to auto-discover and protect instances. Tags may include Key only or Key and Value matches. Wildcards are not permitted for instance tag discovery.

Instance and tag-based matching may only be configured after the *instance group* is created, by editing the content selection and adding rules.

Tables must be in an `active` state to be selected for backup.

Streaming backups

Commvault Backup performs Full and Incremental backups of Amazon DynamoDB tables. Backups use a combination of dynamodb:ListTables, dynamodb:Scan, dynamodb:GetShardIterator, and dynamodb:GetRecords to collect all data stored in the table.

DynamoDB supports the streaming of item-level change data capture records in near-real time, this capability is called **DynamoDB streams**. To perform incremental backups, Amazon DynamoDB streams must be enabled on the table, and the `StreamViewType` set to **`NEW_IMAGE`** or **`NEW_AND_OLD_IMAGES`**.

See **Enabling a stream** for more information.

> ⓘ **Note**
>
> If Amazon DynamoDB Streams is not enabled, or `StreamViewType` is set to `KEYS_ONLY` or `OLD_IMAGE`, all backups will be taken as a Full backup.

Increasing DynamoDB backup and restore performance

Amazon DynamoDB has two read/write capacity modes that dictate how you are charged for read and write throughput, and how to need to manage capacity over the table lifetime. You set the **read/write capacity mode** when creating a table, but it can be changed later.

If using **On-demand read/write mode**, there are no changes or optimizations required to accommodate Commvault Backup and Recovery, DynamoDB will instantly accommodate changes in read/write throughput.

If using **Provisioned mode** to control or match your read/write throughput rates to known performance baselines, you can enable the **Adjust read capacity** toggle in your Commvault DynamoDB *Table group*. When you supply a value to **Adjust read capacity**, Commvault updates the `ProvisionedThroughput – ReadCapacityUnits` during the backup job (see **ec2:UpdateTable**).

See **Read/write capacity mode**.

Encryption during backup and restore

Data stored in Amazon DynamoDB tables are encrypted at rest by default. When creating your DynamoDB tables you have the option to select Amazon DynamoDB managed, AWS managed (`aws/dynamodb`), or owned and managed AWS KMS keys.

> ⓘ **Note**
>
> Commvault will always restore your tables using the same encryption settings on the source table during backup. You cannot change the encryption settings during Commvault restores. You can make this change in the AWS DynamoDB console after the restore.

Commvault Backup downloads your DynamoDB tables over TLS 1.2 or greater connections and stores them in an independently encrypted Amazon S3 library. Commvault Cloud Libraries may be encrypted using Amazon S3 Server-Side Encryption (SSE) including SSE-S3, SSE-KMS, SSE-C, and DSSE-KMS.

See **Server-side encryption** for more details.

> ⚠ **Note**
>
> If you created your Commvault Cloud Library or Amazon S3 bucket using Commvault Backup & Recovery, the bucket encryption setting will be **disabled**. You should enable encryption on the bucket before writing any backup data to ensure all backups are encrypted following your internal policy.

Performance

Performance of DynamoDB backups and restores is dependent on three factors:

- The read/write throughput that is available from the DynamoDB table.
- The number of readers or streams configured in the Commvault *Table group* (backup) or *Restore job.*
- The instance type and size of the Amazon EC2 instance used as an Access Node.

Commvault has selected smart defaults based on the default recommended Access Node c7g.xlarge.

On-demand backups

Using either the Command Center Console, API, CLI, or SDK, you can initiate an on-demand backup of a single table or a group of tables. When you initiate an on-demand backup, Commvault will determine the appropriate storage location per your configured *server plan*.

Concurrent backups and service levels

Be aware of the Amazon DynamoDB service quotas that impact each of your AWS accounts and regions being protected.

Be aware that switching read/write capacity modes can only occur every 24 hours.

Be aware that there are **Throughput default quotas**, and limits to **Increasing provisioned throughput.**

See **Quotas and limits**.

Restoring a DynamoDB table

Using either the Command Center Console, API, CLI, or SDK, you can restore your Amazon DynamoDB table backups to the existing table or a new table. When performing a restore, Commvault will use the `dynamodb:CreateTable` action to create a new table and then populate it with data from the backup. You can supply the following parameters with Commvault:

- Number of Commvault streams
- Adjust write capacity
- Destination region
- TableName

> ⚠ **Note**
>
> Commvault does not restore or copy **Tags** from the source table to newly created DynamoDB tables. Manually re-apply AWS Resource Tags to any restored instances after the Commvault restore completes.

Amazon RDS protection

Commvault Backup & Recovery provides unified and automated data protection for your Amazon RDS clusters and databases, including serverless and provisioned resources. You can backup instances located in the AWS region and edge locations. Amazon Relational Database Service (Amazon RDS) is a collection of managed services that makes it simple to set up, operate, and scale databases in the cloud. Amazon RDS is available in seven popular engines — Amazon Aurora with MySQL compatibility, Amazon Aurora with PostgreSQL compatibility, MySQL, MariaDB, PostgreSQL, Oracle, and SQL Server. Commvault protects all engine types, single instance databases, and multi-AZ database clusters using snapshot and database-consistent dumps.

Commvault Backup automatically discovers and protects new RDS clusters using AWS resource tags. *Server plans* dictate how many snapshots, snapshot copies, or database dumps are created. Snapshots may be shared and copied across AWS accounts and regions for increased protection from regional events.

Commvault copies both Multi-AZ DB cluster snapshots and single-instance DB snapshots cross-account and cross-region.

See **Multi-AZ deployments** for details on creating high-availability DB deployments with automated failover.

> ⓘ **Note**
>
> Commvault also protects your **Amazon RDS Custom** Oracle database engines using database-native dumps, and application-integrated database, configuration, and transaction log backup.

> ⚠ **Important**
>
> Commvault does not protect Amazon RDS Custom instances using RDS snapshot technology, an agent must be installed on the Amazon RDS instance to collect database files and configuration.

IAM permissions for creating and restoring backups

Commvault Backup requires an IAM identity (user or role) with the Commvault-supplied **Amazon RDS IAM policy** attached to allow backup and recovery actions. You can also set a **PassKey** within Commvault as an additional control to ensure the identity performing a restore is authorized.

Commvault Backup supports and recommends using the Security Token Service (STS) `AssumeRole` action to obtain temporary security credentials from each protected workload account during data management activities. Use of STS:AssumeRole aligns with a well-architected secure-by-default approach by using temporary credentials that expire and require periodic renewal.

You should create **Permission boundaries for your IAM entities** by enhancing your Commvault IAM role with restricted Commvault access to AWS resources using **tags**.

See **Using Security Token Service (STS) AssumeRole for backup and recovery**.

> ⓘ **Note**
>
> When adding your first Amazon RDS instance or cluster for protection, you will need to create a **Cloud account credential** for use with Amazon RDS snapshot and dump-based protection. Previously configured AWS hypervisor accounts <u>cannot</u> be reused for RDS protection.

IAM and database permissions required for creating and restoring dump-based backups

Commvault Backup requires an IAM identity (user or role) with the Commvault-supplied **Amazon RDS discovery IAM policy** attached to allow backup and recovery actions.

Commvault Backup also requires that a database user account with the following permissions be configured to allow backup and recovery of MySQL and PostgreSQL instances using database-native dump utilities.

- **Permissions required for creating and restoring MySQL native backups**.
- **Permissions required for creating and restoring PostgreSQL native backups**.

> ⓘ **Note**
>
> When adding your first Amazon RDS instance or cluster for protection, you will need to create a **Cloud account credential** for use with Amazon RDS snapshot and dump-based protection. Previously configured AWS hypervisor accounts <u>cannot</u> be reused for RDS protection.

Auto-protecting instances by tag

By default, Commvault will configure a *default instance group* that protects all Amazon RDS clusters and instances in all regions, including Amazon Aurora. You may optionally use *specific regional discovery, instance name wildcards* (i.e., my-finance-mysql-db*), or *instance tag values* to auto-discover and protect instances. Tags may include Key only or Key and Value matches. Wildcards are not permitted for instance tag discovery. Resources located in edge locations like AWS Local Zones and AWS Outposts may be auto-discovered by selecting all resources in their parent region or using *instance name* or *tag-based* discovery*.

Instance and tag-based matching may only be configured after the *instance group* is created, by editing the content selection and adding rules.

> ⓘ **Note**
>
> In Multi-AZ DB clusters *instance name* and *instance tag* matching are performed on the Writer Instance only. Applying tags to the Regional Cluster or Reader Instance will not detect instances to protect.

> ⓘ **Note**
>
> When adding **instances** or **clusters** to a Commvault *instance group,* a default configuration of *All instances in all regions* is selected. The displayed list of regions does not reflect any region-based filters configured on your Commvault cloud credential or the IAM policy attached to the credential.

> ⓘ **Note**
>
> When adding **instances** or **clusters** to a Commvault *instance group,* instances in `Creating` state will be shown and may be selected. Please wait until instances reach `Available` status or backup will fail with `No database discovered in available state for backup`.

> ⓘ **Note**
>
> When adding **instances** or **clusters** to the backup scope, please ensure that the instance has had a database schema and initial data load performed. Commvault will fail to protect empty instances with

Encryption during backup and restores

By default, Commvault inherits the encryption settings of the protected Amazon RDS instance during protection operations. Amazon RDS snapshots that are taken from encrypted RDS instances are automatically encrypted using the source cluster data encryption key. RDS instances that are created from encrypted snapshots are automatically encrypted using the source snapshot data encryption key.

See **Encrypting Amazon RDS resources** for more details.

Copying snapshot backups cross-region and cross-account

Commvault can protect your Amazon RDS snapshot backups by taking a full copy out of the workload account into your central backup account. Additionally, Commvault can create copies of your RDS snapshots in one or many regions (fan-out) to provide disaster recovery copies to protect from regional events.

Commvault can combine these controls and provide **cross-account and cross-region copies** that maintain encryption using **multi-region keys** or re-encrypted the instance in the destination region. To re-encrypt an RDS snapshot, the destination AWS account and region must have an AWS KMS key with an **alias** or **tag** name of `cvlt-rds` or `cvlt-master`.

> ⓘ **Note**
>
> The IAM identity performing the backup operation must have access to perform backup and recovery of Amazon RDS instances in the source and destination regions and be listed as a **KMS Key User** for the destination region.

> ⓘ **Note**
>
> Tags attached to the source DB snapshot are copied to the destination AWS account or the regional DB snapshot copy.

> ⓘ **Note**
>
> When cross-region snapshot replication is configured with infinite retention, the **retain until** information displayed during an active snapshot backup may incorrectly reflect `Dec 31, 1969 11:59:58 PM`, this information can be safely ignored.

Snapshot backup

Commvault performs Amazon RDS and Amazon RDS Custom **full backup** by automating the creation and deletion of **DB snapshots** or **Multi-AZ DB cluster snapshots** and managing them per your assigned *server plan*. Backups utilize the `rds:CreateDBSnapshot` and `rds:CreateDBClusterSnapshot` actions which record the following information about your cluster/instance and cluster/instance snapshot:

- AvailablityZones
- DBClusterSnapshotIdentifier
- DBClusterIdentifier
- SnapshotCreateTime
- Engine
- EngineMode
- AllocatedStorage

- Status
- Port
- VpcId
- ClusterCreateTime
- MasterUsername
- EngineVersion
- LicenseModel

- SnapshotType
- PercentProgress
- StorageEncrypted
- KmsKeyId
- DBClusterSnapshotArn
- IAMDatabaseAuthenticationEnabled
- TagList

> ⚠ **Note**
>
> Commvault does not protect the **DB cluster parameter group** or **option group** that is used or associated with your protected Amazon RDS and Amazon Aurora resources. Commvault does protect the group names and attempts to re-associate your restored instances with their original parameter and options groups if they exist.

Database-consistent dump backup

Commvault Backup can perform logical service-independent full backups of Amazon RDS databases so that long-term retention copies or database migration may occur between database locations. You can use service-independent backups to migrate between **provisioned** and serverless RDS capacity types. Database-consistent dumps use the database native dump or export utility to capture the entire instance. Dump-based backups may be performed with an ARM 64-bit or x86 64-bit Access Node (the access node does not need to match the architecture of the database being protected). Commvault recommends using **AWS Graviton base Access Nodes** (i.e., C7g, C6g) for the best price-performance and sustainability of long-term retention dump-based backups.

Dump-based backups may be used to restore to the original instance, or another pre-provisioned instance, or perform table-level restores. Backups are stored in a Commvault cloud library with independent encryption to the source DB instance, providing another level of protection for your backups.

> ⓘ **Note**
>
> Amazon RDS does not provide a method to perform transaction log backup and recovery at this time, dump-based backups must be taken as **Full backups** only.

Commvault Backup can use your existing Access Nodes to perform dump-based protection by installing the relevant Commvault database agents (MySQL, PostgreSQL) on your existing Access Node. Dump-based protection will utilize TLS 1.1/1.2 for all communication to and from the Access Node to the database endpoint (see **TLS 1.1/1.2 support for Amazon Aurora and Amazon RDS PostgreSQL**, **TLS 1.2/1.2 support for Amazon Aurora and Amazon RDS MySQL and MariaDB**).

Commvault supports TLS 1.1/1.2 encrypted connectivity to the following Amazon RDS database engines:

- MySQL
- PostgreSQL
- MariaDB
- Oracle
- SQL Server

Commvault supports TLS 1.2/1.2 encrypted connectivity for the following Amazon RDS Custom database engines:

- Oracle
- SQL Server

Database-consistent backup of Amazon RDS Custom

Commvault Backup can perform the database-consistent backup of **Amazon RDS Custom** for Oracle and Amazon RDS Custom to SQL Server by installing Commvault database agents on your Amazon RDS compute instances.

Commvault can provide full, incremental, transaction log, and differential backup and recovery of the following database-integrated items when using an agent-in-RDS approach:

- Oracle database files, log files, and control files.
- SQL Server databases, log files.

See **Oracle**.

See **Microsoft SQL Server**.

Note: Block-Level Backup is not supported with Amazon RDS Custom.

> ⓘ **Note**
>
> Commvault does not support IntelliSnap® backup and recovery of Amazon RDS Custom instances to snapshot underlying Amazon EBS volumes, using the agent-in-RDS guest. Only streaming backup copies is supported.

Restoring an Amazon RDS cluster or instance from a snapshot

Using either the Command Center Console, API, CLI, or SDK, you can restore an Amazon RDS DB cluster or DB instance to the original AWS account, region, and availability zone. Alternatively, you can restore to another AWS account, region, or zone including edge locations, if you have copied the snapshot to the destination region.

Commvault will create an Amazon RDS instance within provided Region, Availability Zone, Instance type, Instance size, and EBS volume type and encryption preferences. Nested or dependent resources will be recreated if not available in the destination account, including Elastic Network Interfaces (ENIs), Security Groups, and AWS KMS Keys.

> ⓘ **Note**
>
> Snapshots created by Commvault will have three tags applied with the following example values
> **Name** = SP_2_1015_2
> **Description** = Snapshot_created_by_Commvault_for_job_1033._Source_RDSInstance_pg-aurora-1
> **_GX_BACKUP_** = *null*

Commvault will restore the cluster or instance with the same **Capacity type** (Provisioned or Serverless). Refer to *Restoring an Amazon RDS database from the dump* for details on migrating between capacity types with Commvault.

Commvault supports supplying or customizing the following `rds:CreateDBcluster` or `rds:CreateDBInstance` parameters during restore:

- DatabaseName / DBName
- PubliclyAccessible
- MultiAZ
- AvailabilityZones.AvailabilityZone.N (instance restores only)
- DBInstanceClass
- StorageType
- DBSubnetGroupName
- DBClusterParameterGroupName / DBParameterGroupName
- OptionGroupName
- Port
- AutoMinorVersionUpgrade
- DBSecurityGroups

> ⓘ **Note**
>
> Commvault does not support supplying or customizing the following parameters to `rds:CreateDBCluster` or `rds:CreateDBInstance` Commvault will create a new instance with these parameters matching the original instance. Use Commvault-created DB snapshots to manually deploy an instance using the RDS console, or AWS CLI to modify these <u>additional</u> parameters:
>
> - DB engine
> - DB engine version
> - VpcId
> - Availability Zone(s)
> - EnableIAMDatabaseAuthentication
> - KMS Key Id
> - CopyTagsToSnapshot
> - EnableCloudWatchLogsExports.member.N
> - DomainIAMRoleName
> - MonitoringRoleArn
> - DeletionProtection

> ⓘ **Note**
>
> Commvault does not permit changing the **DB instance class** during restoration. Restored instances will inherit the DB instance class (Serverless, Memory optimized provisioned, and Burstable provisioned) of the source instance.

> ⓘ **Note**
>
> Recovery of RDS clusters or instances may only occur from an in-region snapshot. If a **cross-region** restore is required, ensure the *server plan* is configured with region mapping and **snapshot replication** is enabled in the instance group.

> ⓘ **Note**
>
> Commvault will attempt to restore with the **DB parameter group** and **Option group** captured during the backup, if they do not exist in the target AWS account and region the restore will fail with `OptionGroupNotFoundFault`. See **Errors**.

Commvault restored RDS clusters and/or instances will inherit the AWS tags from the original protected cluster as Commvault passes `CopyTagsToSnapshot = true` when restoring clusters or instances with **rds:RestoreDBClusterFromSnapshot** or **rds:RestoreDBInstanceFromDBSnapshot**.

Restoring an Amazon RDS database from a dump backup

Commvault Backup can use your service-independent database-consistent dump to restore back to the original Amazon RDS cluster or instance, to another pre-provisioned Amazon RDS cluster or instance, or a compute instance running a compatible version of your database engine.

You can restore the entire instance or individual databases. Restoring back to the original instance is often faster than deleting and creating an entirely new instance. Restoring back to the original instance maintains instance identifiers, critical in infrastructure as code (IaC) and managed services environments (i.e., Amazon Managed Services).

A compute instance running the same database engine version *or higher* is considered compatible.

You can restore user-defined databases. System databases are not restored.

> ⚠ **Note**
>
> When you create an Aurora MySQL instance, the software automatically creates a database that is called innodb. This database does not include a format file, so the tables are not listed. Make sure that you do not restore the innodb database.

> ⓘ **Note**
>
> Commvault will not recreate a new Amazon RDS instance when restoring from a dump-based backup, you must pre-provision the cluster or instance, then restore the dump-based backup into the instance.

Restoring an Amazon RDS Custom database

Using either the Command Center Console, API, CLI, or SDK, you can restore an Amazon RDS Custom database to the original AWS account, region, and availability zone. Alternatively, you can restore to another AWS account, region, or zone including edge locations. Additionally, Commvault RDS Custom backups are database-native backups that allow recovery to any compute instance running the same database engine version (or higher).

Recovery requires that an Amazon RDS instance or compute instance with database software has been pre-provisioned, and a Commvault database agent installed. Commvault supports item-level granularity including:

- Oracle
    - Restore the full database including log files and the control file.
    - Restore a partial database including one or more of the following items:
        - Archive logs.
        - Control file.
        - Individual data files and tablespaces.
        - Database archived redo logs.

Commvault granular item-level recovery allows complete control of your Amazon RDS custom recovery, including transaction log roll-forward, roll-back, and full or partial migration to new RDS instances or AWS services.

Commvault restored RDS custom clusters and/or instances (from RDS Custom snapshot) will inherit the AWS tags from the original protected cluster as Commvault passes `CopyTagsToSnapshot = true` when restoring clusters with **rds:RestoreDBClusterFromSnapshot** or **rds:RestoreDBInstanceFromDBSnapshot**.

> ⓘ **Note**
>
> Tags attached to the source DB snapshot are copied to the destination AWS account or the regional DB snapshot copy.

Amazon Redshift protection

Commvault Backup & Recovery provides unified and automated data protection for your Amazon Redshift clusters. You can backup instances located in the AWS Region. Amazon Redshift is a fast, fully managed, petabyte-scale data warehouse service that makes it simple and cost-effective to efficiently analyze all your data using your existing business intelligence tools. Commvault protects Amazon Redshift clusters using regional snapshots that allow full instance recovery to a new cluster.

Commvault Backup automatically discovers and protects new Redshift clusters using AWS resource tags. *Server plans* dictate how many snapshots and snapshot copies are created. Snapshots may be shared and copied across AWS accounts and regions for increased protection from regional events.

> ⓘ **Note**
>
> Commvault supports the protection of Amazon Redshift provisioned clusters <u>only</u>. **Redshift serverless** clusters cannot be protected or recovered.

IAM permissions for creating and restoring backups

Commvault Backup requires an IAM identity (user or role) with the Commvault-supplied **Amazon Redshift IAM policy** attached to allow backup and recovery actions.

Commvault Backup supports and recommends using the Security Token Service (STS) `AssumeRole` action to obtain temporary security credentials from each protected workload account during data management activities. Use of STS:AssumeRole aligns with a well-architected secure-by-default approach by using temporary credentials that expire and require periodic renewal.

You should create **Permission boundaries for your IAM entities** by enhancing your Commvault IAM role with restricted Commvault access to AWS resources using **tags**.

See **Using Security Token Service (STS) AssumeRole for backup and recovery**.

Auto-protecting instances by tag

By default, Commvault will configure a default instance group that protects all Amazon Redshift clusters in all regions. You may optionally use instance name wildcards (i.e., `my-redshift-wh-*`) or instance tag values to auto-discover and protect instances. Tags may include Key only or Key and Value matches. Wildcards are not permitted for instance tag discovery.

Instance and tag-based matching may only be configured after the *instance group* is created, by editing the content selection and adding rules.

> ⓘ **Note**
>
> In Multi-AZ DB clusters *instance name* and *instance tag* matching are performed on the Writer Instance only. Applying tags to the Regional Cluster or Reader Instance will not detect instances to protect.

Encryption during backup and restore

By default, Commvault inherits the encryption settings of the protected Redshift Cluster during protection operations. Amazon RedShift snapshots that are taken from encrypted Redshift clusters are automatically encrypted using the source cluster data encryption key. Redshift clusters that are created from encrypted snapshots are automatically encrypted using the source snapshot data encryption key.

See **Amazon Redshift clusters** for more details.

Snapshot backups

Commvault performs Amazon Redshift full backups by automating the creation and deletion of **Manual Snapshots** and managing them per your assigned *server plan*. Backups utilize the `redshift:CreateClusterSnapshot` action which records the following information about your cluster and cluster snapshot:

- AccountsWithRestoreAccess. AccountWithRestoreAccess. N
- ActualIncrementalBackupSizeInMegaBytes
- AvailabilityZone
- BackupProgressInMegaBytes
- ClusterCreateTime
- ClusterIdentifier
- ClusterVersion
- CurrentBackupRateInMegaBytesPerSecond
- DBName
- ElapsedTimeInSeconds

- Encrypted
- EncryptedWithHSM
- EngineFullVersion
- EnhancedVpcRouting
- EstimatedSecondsToCompletion
- KmsKeyId
- MaintenanceTrackName
- ManualSnapshotRemainingDays
- ManualSnapshotRetentionPeriod
- MasterUsername
- NodeType

- NumberOfNodes
- OwnerAccount
- Port
- RestorableNodeTypes.NodeType.N
- SnapshotCreateTime
- SnapshotIdentifier
- SnapshotRetentionStartTime
- SnapshotType
- SourceRegion
- Status
- Tags.Tag.N
- TotalBackupSizeInMegaBytes
- VpcId

> ⓘ **Note**
>
> Commvault-created Redshift snapshots will be created with the following example AWS tags for traceability:
>
> ```
> Name = SP_2_1025_3
> Description = Snapshot_created_by_Commvault_for_job_1025_Source_Redshift_redshift-cluster-001
> _GX_BACKUP_ = null
> ```

See **Amazon Redshift Snapshots** for more information on snapshots.

> ⚠ **Note**
>
> Commvault does not protect the Amazon Redshift Cluster **Parameter group**.

> ⓘ **Note**
>
> Commvault does not perform incremental or synthetic full backups for Amazon Redshift instances. Amazon Redshift snapshots are incremental by default – see **Amazon Redshift snapshots**.

Manual snapshots are retained indefinitely, even after you remove your Amazon Redshift cluster. Commvault Backup provides peace of mind as your Redshift Manual Snapshots will be retained and removed per your *server plan* or business policy.

Copying snapshot backups cross-region and cross-account

Commvault can protect your Amazon Redshift snapshot backups by taking a full copy out of the workload account into your central backup account. Additionally, Commvault can create copies of your Redshift snapshots in one or many regions (fan-out) to provide disaster recovery copies to protect from regional events. Commvault can combine these controls and provide **cross-account and cross-region copies** that maintain encryption using **multi-region keys**, or re-encrypting the snapshot in the destination region.

Commvault copies **tags** from the source snapshot to cross-region and cross-account copies, ensuring consistency in tag-based automated operations.

To replicate a copy of encrypted Redshift snapshots, the destination AWS account must either have an AWS KMS key with an **alias** or **tag** name of `cvlt-rds` or `cvlt-master` in the destination region.

> ⓘ **Note**
>
> The IAM identity performing the backup operation must have access to perform backup and recovery of Amazon Redshift instances in the source and destination regions and be listed as a **KMS Key User** for the destination region.

Restoring an Amazon Redshift cluster from a snapshot

Using either the Command Center Console, API, CLI, or SDK, you can restore your Amazon DocumentDB backups from a regional snapshot to a new instance with a different Cluster identifier (name) and Instance class. When performing a restore from a snapshot using `redshift:RestoreFromClusterSnapshot`, you can customize the following parameters with Commvault:

- ClusterIdentifier
- PubliclyAccessible
- AvailabilityZone

- NodeType
- ClusterSubnetGroup Name

- ClusterParameterGroupName
- AllowVersionUpgrade
- VpcSecurityGroupIds.VpcSecurityGroupId.N

Commvault copies **Tags** recorded with the cluster snapshot onto newly created Redshift clusters using the `redshift:CreateTags` action.

> ⓘ **Note**
>
> Commvault does not support supplying or customizing the following additional parameters to `redshift:RestoreFromClusterSnapshot` Commvault will create the new instance with these parameters matching the original instance. Use Commvault-created cluster snapshots to manually deploy an instance using the Redshift console, or AWS CLI to modify these parameters:
>
> - NumberofNodes
> - Port
> - IamRoles.IamRoleArn.N
> - VpcId
>
> - EnhancedVpcRouting
> - PreferredMaintenanceWindow
> - MaintenanceTrackName
>
> - AutomatedSnapshotRetentionPeriod
> - ManualSnapshotRetentionPeriod
> - AvailabilityZoneRelocation

> ⓘ **Note**
>
> Commvault will restore Amazon Redshift snapshots to the same cluster capacity type (Provisioned or Serverless) only. Use the Amazon Redshift, AWS CLI, or SDK to restore a Commvault-created snapshot to a Serverless namespace.

Additional Resources

- **System Requirements for Amazon Aurora PostgreSQL protection** (Snapshot).
- **System Requirements for Amazon Aurora MySQL protection** (Snapshot).
- **System Requirements for Amazon DocumentDB protection** (Snapshot).
- **System Requirements for Amazon DynamoDB protection** (Streaming).
- **System Requirements for Amazon RDS protection** (Snapshot).
- **System Requirements for Amazon RDS for PostgreSQL protection** (Streaming).
- **System Requirements for Amazon RDS for MySQL protection** (Streaming).
- **System Requirements for Amazon RDS for MariaDB protection** (Streaming).
- **System Requirements for Amazon RDS for Oracle protection** (Streaming).
- **System Requirements for Amazon RDS for SQL Server protection** (Streaming).
- **System Requirements for Amazon Redshift protection** (Snapshot).
- **IAM permissions for creating and restoring Amazon DocumentDB backups** (Snapshot).
- **IAM permissions for creating and restoring Amazon DynamoDB backups** (Streaming).
- **IAM permissions for creating and restoring Amazon RDS backups** (including Amazon Aurora).
- **IAM permissions for creating and restoring Amazon Redshift backups**.

Networking & Content Delivery

Amazon VPC protection

Commvault protects Amazon VPC resources that are associated with protected Amazon EC2 instances. Please see the **Auto-discovering, protecting, and recovering Amazon VPC resources** section for more details.

Storage

Amazon EBS protection

Commvault Backup & Recovery supports unified and automated data protection for your Amazon EBS volumes. You can backup volumes located in the AWS region and edge locations. Volume backups are initiated during the creation of Amazon Machine Images (AMIs) which create EBS snapshots for all attached volumes. Commvault can also take service-independent snapshot copies for long-term retention, reduced cost, and data mobility between AWS services.

By default, Commvault Backup & Recovery of Amazon EC2 instances creates crash-consistent backups of Amazon EBS volumes that are attached to protected Amazon EC2 instances. Crash consistency means inflight application and operating system file-system operations are not flushed to disk before taking the snapshot. Commvault creates EBS snapshots of all volumes via the `ec2:CreateImage` action. You do not need to power down your EC2 instance during the creation of snapshots, Commvault passes the `NoReboot` parameter to **ec2:CreateImage**.

> ⓘ **Note**
>
> Commvault does not use ec2:CreateSnapshots action to create crash-consistent snapshots of multiple EBS volumes in parallel. Snapshot start time may differ between volumes. If a volume is excluded from the backup scope, an EBS volume snapshot is created by ec2:CreateImage and then removed by Commvault.

There are no limitations on the types of volumes that may be protected (SSD-backed, HDD).

Commvault supports the use of **Amazon EBS direct APIs** for performing optimized EBS volume backup and recovery that reduces backup time by up to 80% from traditional methods. Additionally, Commvault supports Commvault

HotAdd backup and recovery, which enables the pre-warming of volumes as part of recovery, delivering production performance immediately after restoration.

Commvault recommends using Amazon EBS direct APIs for the best performance and reduced cost of providing service-independent backup and restore capability. EBS direct APIs are supported for all backup and recovery use-cases supported by Commvault Backup.

Protection Use Case	Supported
Streaming backup - *copying EBS contents to a Commvault-optimized copy*	✓
Snap backup & Live Browse - *performing single-file browse & recover from EBS snapshot*	✓
Snap backup copy - *creating Commvault optimized backup copy from EBS snap*	✓
Streaming, Snap/Snap copy in multi-tenant environments – *creating Commvault-optimized backup copies in multi-account environments*	✓
Live Browse for Windows guests using EBS direct – *browsing Windows-based volume snapshots and performing item-level recovery*	✓
Live Browse for Linux guests using EBS direct – *browsing Linux-based volume snapshots and performing item-level recovery*	✓

IAM Permissions for creating backups and restores

Commvault Backup requires an IAM identity (user or role) with the Commvault-supplied **Amazon EC2 IAM policy** attached to allow backup and recovery actions. The specific actions required for EBS direct API backup and recovery (recommended) are:

```
ebs:CompleteSnapshot
ebs:GetSnapshotBlock
ebs:PutSnapshotBlock
ebs:StartSnapshot
ebs:ListChangedBlocks
ebs:ListSnapshotBlocks
```

See **IAM permissions for EBS direct APIs**.

Commvault uses `iam:SimulatePrincipalPolicy` to validate these permissions are granted to the user or role performing the data management activity. Commvault will revert to using HotAdd backup if these permissions are not available.

> ⓘ **Note**
>
> Ensure that your IAM identities (users and roles) are granted access to execute the Commvault-required service actions. When using AWS Organizations and Service Control Policies (SCPs), ensure your SCPs grant access to the required service actions as SCPs will override permissions made at the account level.

Commvault checks whether you have configured a VPC endpoint for the Amazon EBS service when using EBS direct APIs using the `ec2:DescribeVpcEndpoints` action. Backups will not fail if a VPC endpoint is unavailable, but they will need to traverse your Internet Gateway (IGW) to transfer data to/from the EBS service endpoint.

Commvault **Amazon EC2 IAM Policy** also includes the permissions for creating, attaching, and detaching EBS volumes from Commvault Access Nodes to perform Commvault HotAdd backup and recovery. EBS direct APIs are

used by default for all backups and restores, HotAdd may be selected during recovery operations via the Command Center Console, API, command line, or SDK.

Enabling Change Block Tracking Incremental backup

Commvault Backup, by default, enables **Change Block Tracking (CBT)** using EBS direct APIs for all new VM groups. You can verify if you are using EBS direct API backups by checking the **Use changed block tracking** toggle on your Amazon EC2 VM groups in the Command Center console, or via API.

Commvault will always use EBS direct APIs to request changes between your **Incremental** backups (`ebs:ListSnapshotBlocks`, `ebs:ListChangedBlocks`), regardless of whether the EBS direct API actions for performing backups (`ebs:GetSnapshotBlock`) have been granted

See **Changed Block Tracking for AWS**.

Encryption during backup and restores

See **Amazon EC2 – Encryption during backup and restores** for details on how Commvault handles the protection and recovery of encrypted volumes.

Performance

See **Amazon EC2 – Performance** for details on the performance that has been measured for EBS direct APIs in Commvault lab testing.

Performing EBS volume recovery

See **Amazon EC2 – Performing item-level restores for EC2 instances** for details on restoring EBS volumes or files and folders into existing EBS volumes and EC2 instances. Commvault Live Browse provides the ability to browse your EBS volume snapshots and select files and folders for recovery back to your original EC2 instance, another EC2 instance, or a file-system. Item-level recovery allows rapid recovery of your business services without having to recreate Amazon EC2 instances and reconfigure your application(s).

Regardless of whether an Amazon EBS volume restore is occurring from an EBS snapshot or a service-independent backup copy, the AWS tags attached to the original volume will be restored or recreated.

Amazon EFS protection

Commvault Backup & Recovery provides unified and automated data protection for your Amazon EFS file-systems. Amazon Elastic File System (Amazon EFS) is a simple, serverless, set-and-forget elastic file system that makes it easy to set up, scale, and cost-optimize file storage in AWS. Commvault protects all EFS storage classes and can leverage provisioned throughput (MB/s) for additional backup and restore performance.

Commvault Backup automatically protects configured EFS file-systems using Network File System (NFS) protocol or backup, recover, and migrate your file-system. *Server plans* dictate how many backup copies are retained. Backup copies may be replicated across AWS accounts and regions for increased protection from regional events.

IAM permissions for creating and restoring backups

Commvault does not require any **Amazon EFS actions** to perform backup and recovery of Amazon EFS file-systems.

Commvault utilizes the Network File System (NFS) protocol to access, protect, and recover your Amazon EFS file-systems.

Permissions for creating and restoring EFS file-system backups

Commvault must be granted read and write permissions (as root) to your Amazon EFS file-system from the Access Node or Access Node(s) that will be used to perform backup and recovery services.

You will need to ensure that your **network access controls** include access from the VPC, subnets, and Access Nodes that will be performing backup and recovery.

You will need to ensure that **no root squashing** is performed on the file-system (this is the default behavior) to ensure Commvault can protect all files and restore file and folder permissions.

Auto-protecting file-systems

You can protect your entire Amazon EFS file-system or specific directories recursively by adding the top-level root folder or specific sub-directories to your Commvault **File System Subclient**. By default Commvault will protect all content in the configured **Network file-system** or locally mounted **EFS file-system**.

> ⓘ **Note**
>
> You cannot discover EFS file-systems or objects to protect using **AWS tags with Amazon EFS.** as Commvault uses the NFS protocol exclusively to access EFS.

Encryption during backup and restore

By default, when the Commvault Access Node accesses Amazon EFS files that are encrypted, it must have access to the KMS key used to encrypt the file-system. Data encryption and decryption at rest are handled transparently, however, the Amazon EFS KMS key policy must provide access to the Commvault Access Nodes that are mounting and accessing the file-system to perform `kms:Decrypt` actions.

See **AWS managed key for Amazon EFS KMS policy**.

Copying backups cross-region and cross-account

Commvault Backup stores your Amazon EFS backups in service-independent deduplicated and encrypted Amazon S3 storage. Commvault provides independent encryption for your Amazon EFS backups, providing another level of protection for your backup data.

Using Commvault **Storage copies** and **DASH Copies**, you can maintain multiple copies of your Amazon EFS backups across regions and AWS accounts at reduced transfer and storage costs.

Snapshot backups

Amazon EFS does not provide an API or action for performing snapshot-based backup and recovery. Commvault uses open-protocol NFS export access to backup and recover files and folders on Amazon EFS file-systems.

Streaming backups

Commvault Backup will perform **Full, Differential, Incremental,** and **Synthetic Full** backups for your Amazon EFS file-system providing the ability to efficiently protect very large EFS file-systems. Backups are performed over NFS v4.1 and v4.0 connections from an in-region Commvault Access Node or Nodes running the Commvault **Network Attached Storage (NAS)** agent and **Linux file-system** agent.

All Amazon EFS storage classes are supported for protection and recovery along with bursting and provisioned throughput modes.

Performance

Backup performance for Amazon EFS file-systems is dependent on the EFS **storage class**, file-system **throughput mode**, EFS **quotas**, and Amazon EC2 **network bandwidth** (burst, baseline) available on the Access Node.

Refer to **Amazon EFS quotas and limits**.

Restoring Amazon EFS files and folders from backup

Using either the Command Center Console, API, CLI, or SDK, you can restore Amazon EFS files to the original AWS account, region, and file-system. Alternatively, you can restore to another AWS account, region, or file-system. The Amazon EFS file-system must exist and be accessible (read/write) by Commvault before performing a restore, Commvault will not create new Amazon EFS file-servers.

You can use **Commvault NFS file-server migration** to automate the migration of an entire NFS export to Amazon EFS and between EFS file-systems.

You can restore your **Amazon EFS files and folders to Amazon S3** if you are looking to reduce the size of active data stored in Amazon EFS.

Optimizing Amazon EFS file-systems with Commvault Risk Analysis

Commvault Risk Analysis can scan your on-premises NFS and SMB/CIFS file-systems, and Amazon FSx file-systems for duplicate, unused, and insecure files. You can then use Risk Analysis (formerly Storage Optimization) to archive, delete, move, or tag files for further analysis. Using Storage Optimization, you can reduce the initial one-time migration cost and ensure storage optimization is part of your continual improvement processes for Amazon EFS file-systems.

See **Risk Analysis** (**Storage Optimization**) for additional details.

Amazon FSx protection

Commvault Backup & Recovery provides unified and automated data protection for your Amazon FSx file-systems. Amazon FSx lets you launch, run, and scale feature-rich and highly-performant file systems with just a few clicks, including FSx for Windows Server, FSx for OpenZFS, and FSx for Lustre. Commvault protects all FSx storage classes and can leverage provisioned throughput (MB/s) for additional backup and restore performance.

Commvault Backup automatically protects configured FSx file-systems using Network File System (NFS) and Server Message Block (SMB) protocols to backup, recover and migrate your file-systems. *Server plans* dictate how many backup copies are retained. Backup copies may be replicated across AWS accounts and regions for increased protection from regional events.

The following section details Commvault support for protecting Amazon FSx file-systems, excluding Amazon FSx for NetApp ONTAP (covered separately).

Commvault protects the following Amazon FSx file-systems using a common approach, which is detailed below:

- Amazon FSx for OpenZFS (NFS exports)
- Amazon FSx for Windows File Server (SMB file-shares)
- Amazon FSx for Lustre (Linux file-system)

IAM permissions for creating and restoring backups

Commvault does not require any **Amazon FSx actions** to perform backup and recovery of Amazon FSx file-systems.

Commvault utilizes the Network File System (NFS), Server Message Block (SMB) protocol, or direct host access to access, protect, and recover your Amazon FSx file-systems.

Permission for creating and restoring FSx file-system backups

Commvault must be granted read and write permissions (as root or admin) to your Amazon FSx file-system from the Access Node or Access Node(s) that will be used to perform backup and recovery services.

You will need to ensure that your **network access controls** include access from the VPC, subnets, and Access Nodes that will be performing backup and recovery.

You will need to ensure that **no root squashing** is performed on file-systems exported via NFS (this default behavior) to ensure Commvault can protect all files and restore file and folder permissions.

You will need to provide an Active Directory identity will admin credentials for CIFS/SMB exported filesystems.

See **Add a NAS File Server**.

Auto-protecting file-systems

By default, Commvault Backup will create a *default subclient* when adding a NAS File Server for protection. The *default subclient* will protect all files and folders within the file-system recursively. You may modify the **Backup Content** for the subclient to restrict which files and folders are protected.

> ⓘ **Note**
>
> You cannot discover FSx file-systems or objects to protect using **Tag your Amazon FSx resources.** as Commvault uses the NFS and SMB protocols exclusively to access your file-systems.

Encryption during backup and restore

By default, when the Commvault Access Node accesses Amazon FSx files that are encrypted, it must have access to the KMS key used to encrypt the file-system. Data encryption and decryption at rest are handled transparently, however, the Amazon FSx KMS key policy must provide access to the Commvault Access Nodes that are mounting and accessing the file system to perform `kms:Decrypt` actions.

See **Encryption at Rest** for more details.

Copying backups cross-region and cross-account

Commvault Backup stores your Amazon FSx backups in service independent deduplicated and encrypted Amazon S3 storage. Commvault provides independent encryption for your Amazon FSx backups, providing another level of protection for your backup data.

Using Commvault **Storage copies** and **DASH Copies**, you can maintain multiple copies of your Amazon FSx backups across regions and AWS accounts at reduced transfer and storage costs.

Snapshot backups

Amazon FSx does not provide an API or action for performing snapshot-based backup and recovery for all service variants. Commvault uses NFS and SMB protocols to access backup and recover FSx file-systems.

> ⓘ **Note**
>
> Commvault does not integrate or use the `fsx:CreateSnapshot` action to create snapshots of Amazon FSx for OpenZFS file-systems.

Streaming backups

Commvault Backup will perform **Full, Differential, Incremental,** and **Synthetic Full** backups for your Amazon FSx file-systems providing the ability to efficiently protect very large FSx file-systems. Backups are performed over CIFS/SMB and NFS v4.1 and v4.0 connections from an in-region Commvault Access Node or Nodes running the Commvault **Network Attached Storage (NAS)** agent, **Windows file-system** agent, or **Linux file-system** agent.

All Amazon FSx storage types are supported (SSD, HDD) for protection and recovery along with additional **throughput capacity** (if required) in single-AZ and multi-AZ deployments.

Performance

Backup performance for Amazon FSx file-systems is dependent on the FSx **storage type**, file-system **throughput**, FSx **quotas**, and Amazon EC2 **network bandwidth** (burst, baseline) available on the Access Node.

Refer to **FSx for Windows File Server performance** for details on tuning performed for FSx SMB-accessible file-systems.

Restoring Amazon FSx files and folders from backup

Using either the Command Center Console, API, CLI, or SDK, you can restore Amazon FSx files to the original AWS account, region, and file-system. Alternatively, you can restore to another AWS account, region, or file-system. The Amazon FSx file-server (file-system) must exist and be accessible (read/write) by Commvault before performing a restore, Commvault will not create new Amazon FSx file-servers.

You can use **Commvault CIFS/SMB file-server migration** to automate the migration of an entire SMB file share to an Amazon FSx file-system and between FSx file-systems.

You can restore your **Amazon FSx files and folders to Amazon S3** if you are looking to reduce the size of active data stored in Amazon FSx.

Optimizing Amazon FSx file-systems with Risk Analysis (Storage Optimization)

Commvault Risk Analysis can scan your on-premises NFS and SMB/CIFS file-systems, and Amazon FSx file-systems for duplicate, unused, and insecure files. You can then use Risk Analysis (Storage Optimization) to archive, delete, move, or tag files for further analysis. Using Storage Optimization, you can reduce the initial one-time migration cost and ensure storage optimization is part of your continual improvement processes for Amazon FSx file-systems.

See **Risk Analysis** (**Storage Optimization**) for additional details.

Amazon FSx for NetApp ONTAP protection

Commvault Backup & Recovery provides unified and automated data protection for your Amazon FSx for NetApp ONTAP (FSxN) file and block storage. Amazon FSx for NetApp ONTAP is fully-managed shared storage built on NetApp's popular ONTAP file-system. Amazon FSx for NetApp ONTAP provides the familiar features, performance, and APIs of on-premises NetApp file systems with the agility, scalability, and simplicity of a fully managed AWS service, Commvault protects all FSxN storage types (SSD, Standard) in Single-AZ and Multi-AZ deployments.

See **Amazon FSx for ONTAP documentation**.

> ⓘ **Note**
>
> FSx for NetApp ONTAP (FSxN) does not support Network Data Management Protocol (NDMP) at the time of writing. Commvault does not support the protection of FSxN via NDMP.

Commvault Backup orchestrates the creation of NetApp ONTAP snapshots, NetApp SnapVault® copies, and NetApp SnapMirror copies following your *server plan*. Commvault Backup will manage on-premises to FSxN snapshots, SnapVault®, and SnapMirror relationships for seamless backup, archive, and disaster recovery to AWS. Commvault Backup & Recovery can optionally take an FSxN independent snapshot copy by mounting the snapshot and copying it to a Commvault independently encrypted Amazon S3 bucket.

> ⓘ Note
>
> Commvault Backup and Recovery creates and manages NetApp ONTAP native snapshots, SnapVault®, and SnapMirror relationships for protection. Commvault does not utilize the **FSx APIs** to discover or manage FSxN-native snapshots.

Commvault Backup provides three core backup and recovery use-cases for FSxN data:

- **Automatic and on-demand protection of FSxN NFS and SMB file-systems**, providing backup, recovery, and migration of your network file-systems between on-premises and FSxN.
- **Automatic and on-demand protection of FSxN iSCSI block-mode LUNs** attached to your protected Amazon EC2 instances.
- **Automatic and on-demand protection for applications running on Amazon EC2 using FSxN**. Commvault Backup integrates application protection with NetApp snapshots for application-consistent snapshot-based protection of key enterprise applications and databases (see **supported agents**)*.*

Commvault Backup will take NetApp snapshot backups, SnapVault® copies, and SnapMirror copies following your *server plans*. *Server plans* dictate how many snapshots, SnapVault copies, SnapMirror replicas, and optional backup copies are retained. Commvault will replicate FSxN snapshots and Commvault FSxN independent backup copies across AWS accounts and regions for increased protection from regional events.

IAM permissions for creating and restoring backups

Commvault does not require any **Amazon FSx IAM permissions** to perform backup and recovery of Amazon FSxN file-systems and iSCSI LUNs.

Commvault integrates directly with the NetApp ONTAP API accessed via the FSxN management endpoint (or SVM management interface) to provide snapshot-based protection.

> ⚠ Important
>
> Amazon FSx periodically synchronizes with ONTAP to ensure consistency. If you create or modify volumes using NetApp applications, it may take up to several minutes for these changes to be reflected in the AWS Management Console, AWS CLI, API, and SDKs.

See **Managing FSx for ONTAP resources using NetApp applications**,

Permissions for creating and restoring FSxN file-system backups

Commvault Backup enables FSxN snapshot management by adding each Amazon FSxN file system as an Array within the Commvault software. For Commvault to perform backup, recovery, and replication of NetApp snapshots a Storage Virtual Machine (SVM) user account must be created with the following permissions is required on the FSxN instance:

- Create and delete privileges for:
  - Snapshots
  - Volume and LUN clones
  - Server Message Block (SMB) shares and Network File System (NFS) exports

o   Initiator groups

Consider creating a new SVM user account dedicated for Commvault Backup & Recovery with the pre-defined `vsadmin-backup` role attached. See **Predefined roles for SVM administrators**.

See **Configuring the NetApp Array Using Array Management**.

> ⚠ **Important**
>
> When accessing your Amazon FSx for NetApp ONTAP using the NetApp ONTAP REST API with the fsxadmin login, you will need to do one of the following:
>
> - Disable TLS validation.
> - Trust the AWS certificate authorities (CAs)
>
> See Using the **NetApp ONTAP REST API**.

> ⓘ **Note**
>
> Commvault does not support NetApp OnCommand Unified Manager (OCUM) or NetApp Blue/XP (formerly NetApp Cloud Manager) when adding Amazon FSxN file-servers.

Auto-protecting FSxN file-systems

You can protect your entire Amazon FSxN file-system or specific directories recursively by adding the top-level root folder or specific sub-directories to your Commvault **File System Subclient**. By default Commvault will protect <u>all</u> content in the configured **Network file-system** or locally mounted **FSxN file-system**.

> ⓘ **Note**
>
> You cannot discover FSxN file-systems (shares, exports) to protect using the process detailed in **Tagging your resources.** as Commvault uses the **ONTAP REST API** to orchestrate protection on statically configured file-systems.

Encryption during backup and restore

Commvault can leverage Kerberos-based encryption in transit provided by FSxN NFS and SMB exported file-systems. Commvault control plane traffic is encrypted by default, data plane traffic can be encrypted through the implementation of **network topologies** encrypted tunnels between the FSxN Access Node and MediaAgents.

See **Data encryption in FSx for ONTAP**.

Additionally, Commvault Access Nodes (Amazon EC2 machine identities) must be granted `kms:Decrypt` permission via an appropriate AWS KMS key policy that allows them to access FSx data encrypted with AWS-managed or Customer-managed KMS keys. File-system encryption and decryption will occur transparently, and data will be transferred to Commvault for independent encryption to a Commvault-controlled Amazon S3 bucket.

See **Amazon FSx key policies for AWS KMS** for more information.

Leveraging NetApp SnapVault and SnapMirror from on-premises

Commvault Backup will automatically create SnapVault® snapshot copies or SnapMirror snapshot mirrored copies following your *server plan*. You can configure your required data flows from **Source FSxN Storage Virtual Machine (SVM)** to **Target FSxN SVM** for SnapVault or SnapMirror relationships.

Commvault Backup supports Source or Target SVM's on-premises or on an Amazon FSx for NetApp ONTAP (FSxN) SVM. Using SnapVault replication, you can migrate, back up, or burst your file-based applications from on-premises to AWS without changing the application or operational process.

As you migrate your applications to Amazon EC2 and Amazon EKS containers, you can continue to use NetApp ONTAP snapshots, SnapVault, and/or SnapMirror replication to manage backup across availability zones or Regions.

Commvault can also establish NetApp SnapMirror relationships between your NetApp ONTAP file-systems and block LUNs to provide high-speed data replication as frequently as every 5 minutes. NetApp SnapMirror provides a *storage disaster recovery* solution between your NetApp instances, both on-premises to AWS, and between FSxN file-systems in AWS.

See **Adding a Snapshot Copy to a Plan**.

> ⓘ Note
>
> Commvault supports the following topologies for replication with Amazon FSxN
>
> Primary – SnapMirror
>
> Primary – SnapMirror – SnapVault
>
> Primary – SnapVault
>
> Primary – SnapMirror – SnapMirror
>
> <u>Note</u>: **SnapMirror to Cloud** is not supported.

Copying backups cross-region and cross-account

Commvault Backup uses NetApp SnapVault and SnapMirror relationships to maintain cross-region and cross-account NetApp-proprietary snapshot copies (see *Leveraging NetApp SnapVault and SnapMirror from on-premises).*

Commvault Backup can also create and store your Amazon FSxN backups in service independent deduplicated and encrypted Amazon S3 storage. Commvault provides independent encryption for your Amazon EFS backups, providing another level of protection for your backup data. Using Commvault **Storage copies** and **DASH Copies**, you can maintain multiple copies of your Amazon FSxN service-independent backups across regions and AWS accounts at reduced transfer and storage costs.

Snapshot backups of FSxN file-systems

Commvault Backup performs Network Attached Storage (NAS) backups of your Amazon FSxN file-systems by leveraging API integration with the NetApp ONTAP REST API. Commvault Backup utilizes Commvault **IntelliSnap®** to orchestrate NetApp snapshots of underlying NetApp SVM file-systems, then Snapvault's or SnapMirror's to remote FSxN or edge-based NetApp filters.

Based on the *server plan* associated with the file-system, the created NetApp snapshot may be Snapvault'ed or/or Snapmirror'ed following the supported topologies (shown above).

Snapshots are file-system consistent and may be optionally streamed to a service-independent copy (see *Streaming backups of FSxN file-systems*), providing additional protections and mobility for your file system-based applications.

Streaming backups of FSxN file-systems

Commvault Backup can also take NetApp-independent backup copies of FSxN snapshots into Commvault-controlled independently encrypted Amazon S3 storage. Commvault Backup leverages Data ONTAP API integration to create a snapshot, then scans the NAS file share or exports for changed data to protect.

Streamed backup copies of your FSxN file-systems allow for rehosting and refactoring your application architectures as part of application modernization efforts. You can use streamed backups to restore FSxN file-systems to Amazon EC2 compute, Amazon EKS containers, Amazon EFS, or Amazon S3 storage services.

Application-integrated snapshot and streaming backups of FSxN storage

Commvault Backup can protect your Amazon FSxN-enabled applications by coordinating snapshot creation with your operating system or applications. FSx for ONTAP provides high-performance solid-state drive (SSD) storage with sub-millisecond latencies often required with enterprise applications. **Commvault IntelliSnap® for NetApp** integrates with your Amazon EC2 Linux and Windows-based instances running the following applications:

- DB2
- DB2 MultiNode
- Exchange Database
- Microsoft SQL Server
- MongoDB
- MySQL
- Windows File System
- Linux File System
- Notes Database
- Oracle
- Oracle RAC
- PostgreSQL
- SAP for Oracle
- SAP HANA
- Sybase
- Network Share (SMB)
- Network Share (NFS)

See **Supported Agents for the NetApp Storage Array**.

Commvault will coordinate **Full**, **Incremental**, and **Differential** backups that ensure that application I/O is quiesced before backup copies are created. Backups of snapshots may occur directly on the production Amazon EC2 instance or optionally on shared data management infrastructure to reduce the impact on production applications.

Restoring Amazon FSxN data

Using either the Command Center Console, API, CLI, or SDK, you can restore Amazon FSxN files to the original AWS account, region, file-system, or EC2 compute file-system. Alternatively, you can restore to another AWS account, region, FSxN file-system, or EC2 compute file-system. The Amazon FSxN file-server must exist and be accessible (read/write) by Commvault before performing a restore, Commvault will not create new Amazon FSxN file-servers.

In recovery scenarios where very-rapid recovery without movement of data is desired for long recovery time objectives (RTOs), you can use the NetApp API, CLI, or SDK to **revert an SVM volume** directly.

> ⓘ **Note**
>
> Commvault does not support revert from FSxN primary snapshots, SnapVault copies, or SnapMirror copies. Use the NetApp command-line or REST API to perform these recovery scenarios.

You can use **Commvault NFS file-server migration** to automate migration of an entire NFS export to Amazon EFS and between EFS file-systems.

You can restore your **Amazon EFS files and folders to Amazon S3** if you are looking to reduce the size of active data stored in Amazon EFS.

Optimized Backups for Very Large Datasets

Commvault has helped customers with very-large NetApp datasets to achieve high-performance optimized backup using **extent-based backup technology.** Commvault Backup for FSxN can provide faster more efficient backups for datasets that have a large number of small files, where an FSxN-independent backup copy is desired.

See **Optimized Backups using Extend-Based Technology** for details on supported agents.

Enabling NetApp SnapLock Compliance and SnapLock Enterprise on Amazon FSx for NetApp ONTAP

Amazon recently announced support for write once, read many (WORM) protection on Amazon FSx for NetApp ONTAP using SnapLock (see **Amazon FSx for NetApp ONTAP supports write once, read many (WORM) protection with SnapLock**).

Commvault supports enabling NetApp® SnapLock Compliance and SnapLock Enterprise on Amazon FSx for NetApp ONTAP file-shares using the **Enable WORM Storage** workflow.

AWS Storage Gateway protection

Commvault Backup & Recovery supports unified and automated data protection for your Amazon Storage Gateway file-systems located within the AWS Region. Commvault protects and recovers your Amazon S3 objects, Amazon FSx file-system, and Amazon EBS volumes within and across AWS accounts and regions, including edge locations.

AWS Storage Gateway is implemented as an on-premises VMware ESXi, Microsoft Hyper-V, or Linux KVM Virtual Machine (VM) or Amazon hardware appliance. Commvault recommends protecting locally cached copies of your Storage Gateway data <u>before</u> it is uploaded to AWS using Hypervisor protection for local volumes and connected block-mode iSCSI volumes. Commvault provides crash-consistent protection for the following hypervisors:

- **VMware ESXi**
- **Microsoft Hyper-V**
- **Linux KVM**

Alternatively, you can install the Commvault **Linux file-system agent** in your Storage Gateway to collect critical configuration and cached files and folders.

Amazon S3 protection

IAM permissions for creating and restoring backups

Commvault Backup requires an IAM identity (user or role) with the Amazon S3 actions that support the backup and recovery of S3 objects (see below). You can also set a **PassKey** within Commvault as an additional control to ensure the identity performing a restore is authorized.

> ⓘ **Note**
>
> Commvault does not publish a recommended IAM policy for the protection of Amazon S3 objects, consider using the following definition:
>
> ```
> {
>   "Version": "2012-10-17",
>   "Statement": [
>     {
>       "Action": [
>         "GetAccelerateConfiguration",
>         "GetBucketAcl",
>         "GetBucketAnalyticsConfiguration",
>         "GetBucketCors",
>         "GetBucketEncryption",
>         "GetBucketInventoryConfiguration",
>         "GetBucketLifecycle",
>         "GetBucketLocation",
>         "GetBucketLogging",
>         "GetBucketMetricsConfiguration",
>         "GetBucketNotification",
>         "GetBucketObjectLockConfiguration",
>         "GetBucketOwnershipControls",
>         "GetBucketPolicy",
>         "GetBucketPublicAccessBlock",
>         "GetBucketReplication",
>         "GetBucketRequestPayment",
>         "GetBucketTagging",
>         "GetBucketVersioning",
>         "GetBucketWebsite",
>         "ListAccessPoints",
>         "ListBuckets"
>       ],
>       "Effect": "Allow",
>       "Resource": "*"
>     }
>   ]
> }
> ```

> ⓘ **Note**
>
> When performing cross-account protection, ensure that the account performing the backup has access to all KMS Keys used to encrypt objects under protection. If the user performing the backup does not have access to the required key, the backup will fail with `ReadFile() - failed to read the remote file - <FILENAME>, error: The ciphertext refers to a customer master key that does not exist, does not exist in this region, or you are not allowed to access`.

Commvault Backup supports and recommends using the Security Token Service (STS) `AssumeRole` action to obtain temporary security credentials from each protected workload account during data management activities. Use of STS:AssumeRole aligns with a well-architected secure-by-default approach by using temporary credentials that expire and require periodic renewal.

You should create **Permission boundaries for your IAM entities** by enhancing your Commvault IAM role with restricted Commvault access to AWS resources using **tags**.

See **Using Security Token Service (STS) AssumeRole for backup and recovery**.

Selecting content to protect

Commvault Backup supports the creation of one or more *Content groups* that identify the S3 buckets and objects to protect and the *server plan* that automates protection per business policy. Commvault Backup can protect *all buckets within the target AWS account* or *all objects within selected buckets* or *individual objects.*

> ⓘ **Note**
>
> Commvault cannot auto-discover or select buckets or objects to protect by **AWS resource tag**.

Streaming backups

Commvault performs a streaming backup of the selected *buckets* and *objects* using the Commvault Access Node configured on the object storage client.

Commvault performs **Full, Incremental**, and **Synthetic Full** backups and uses the last modified time of an object to determine if a subsequent incremental backup will re-protect the object. Object metadata is protected as part of a Full or Incremental backup, it is re-protected on new backups of the object. Commvault performs a single access node discovery of objects to protect, then distributes backup workload to **one or more access nodes** for fault-tolerant, load-balanced backup. Discovery occurs on each top-level prefix in the S3 bucket being protected, so distributing objects across multiple prefixes will increase backup throughput.

> ⓘ **Note**
>
> Commvault supports only a single Access Node per object storage client during **discovery**, to achieve greater parallel backup and restore throughput, configure multiple object storage clients (**content groups**) with different Access Nodes.
>
> ***Commvault recommends keeping the number of protected objects per content group and Access Node group at less than 100 million for optimal backup throughput.***

> ⚠ **Warning**
>
> After a full backup, if the tag or the ACL of an object is modified, but the data has not changed, then the next incremental backup does not include that object. Incremental backups include only objects whose data has changed, indicated by the **last modify time** changing between backups..

Restoring Amazon S3 objects

Using either the Command Center Console, API, CLI, or SDK, you can restore one or more Amazon S3 objects to the original AWS account, region, and bucket. Alternatively, you can restore to another AWS account, region, and bucket., Commvault will create new Amazon S3 buckets as part of a restore operation but will not restore settings and tags from the original bucket. Commvault will create new Amazon S3 buckets with the following default settings if created as part of restore:

- Bucket versioning = Disabled
- Multi-factor authentication (Mfa) delete = Disabled
- Tags = *null*
- Default encryption = Enabled
- Access = Bucket and objects not public
- Block *all* public access = On

> ⚠ **Note**
>
> Commvault does not protect or restore the **bucket** and **bucket properties**.

Commvault will restore to the same **Storage class** they resided on when originally protected.

Commvault will restore the following object properties when restoring objects.

- Object key (Object name).
- Storage class.
- Default encryption.
- Encryption key type.
- AWS KMS key ARN.
- Bucket Key (test it).
- Tags.
- System-defined metadata key-value pairs.
- User-defined metadata key-value pairs.
- Object Lock.
- Access control lists (ACLs).

> ⓘ **Note**
>
> Commvault will protect the latest version of an object only.

> ⓘ **Note**
>
> Commvault does not recover **Additional checksums** properties on restored objects.

Additional Resources

- **System Requirements for Amazon EBS protection**.
- **System Requirements for Amazon EFS protection**.
- **System Requirements for Amazon FSx protection** (including Amazon FSx for Lustre, Amazon FSx for Windows Server, and Amazon FSx for OpenZFS).
- **System Requirements for Amazon FSx for NetApp ONTAP protection**.

SaaS

Commvault can also protect supported Software as a Service (SaaS) workloads using AWS compute and writing to AWS storage.

Salesforce

You can use Commvault software to protect and restore Salesforce objects, documents, attachments, CRM content, files, and metadata. You can perform granular object-level records (i.e., all records of a selected object), record-level restores, and metadata restores. To aid in troubleshooting or development activities, you can restore an entire copy of your Salesforce instance to a separate sandbox. Commvault provides you and your developers flexibility by permitting restores to Salesforce cloud, a file-system, and a cloud-based or on-premises SQL server.

Commvault recommends that backup of Salesforce environments be performed with AWS Graviton Access Nodes for best price-performance and sustainability (see **Back Up and Recover Salesforce by Using AWS Graviton**).

Additional Resources

- **Salesforce protection**
- **Perform metadata and object comparison between Salesforce backups**
- **Perform data masking to protect sensitive data during restore**

Auditing backups and recovery readiness

You can use the Commvault Command Center™ console, API, and SDK to audit and report on your backup and recovery service levels and overall organizational **Recovery Readiness**. Commvault provides complete visibility of your Recovery Readiness across AWS regions and edge locations (i.e., your data centers).

Commvault Recovery Readiness and SLA reporting allows you to answer questions like:

- "Am I backing up all my resources or applications?"
- "Am I meeting my business recovery point and recovery time objectives?
- "Are all my backups encrypted?"
- "Are my backups being copied cross-region following business policy"

You use the **Health Report**, **SLA Report**, and **Recovery Readiness Report** to identify resources that are not compliant with key data management controls.

Creating an audit report of backup activity

Using either the Command Center Console, API, CLI, or SDK, you can run on-demand or scheduled creation of **Backup Job Summary Reports** that identify the detailed backup job history for your AWS and edge-based resources.

Consider scheduling these reports to export to a centralized file-system that is protected for audit and compliance (auditor) review.

See **Schedule a Report on the Command Center**., **Exporting and Saving Reports on Web Console**.

Creating an audit report on SLA compliance

Using either the Command Center Console, API, CLI, or SDK, you can run on-demand or scheduled creation of the **SLA Report**. The SLA Report allows you to track, report, and provide evidentiary proof of the success of resource backups over time. You can use this reporting to demonstrate internal compliance with the protection policy, or with external auditors as proof of protection.

A server meets SLA when all its subclients and databases are protected by at least one successful full, incremental, differential, or log backup job in each time range, such as 1, 3, 5, 7, 14, 21, or 30 days. Synthetic full backup jobs are not included in the SLA calculation.

Ensure that operational reviews and automated metrics and alarms track resources that appear in the **Missed SLA** report, and periodically review the **Excluded Servers in the SLA Report** for business resources that require SLA tracking.

Using Recovery Readiness to measure protection success

Commvault Backup and Recovery measures **Recovery Readiness** for your AWS and edge-based workloads for a true representation of your ability to recover your business services. *Recovery Readiness* takes backup success/failure reporting and trending one step further by modeling and estimating the **Actual Recovery Point Objective** and **Actual Recovery Time Objective** based on a supplied RPO in hh:mm:ss format.

Using either the Command Center Console, API, CLI, or SDK, you can run on-demand or scheduled creation of the **Recovery Readiness Report**.

Using the *Recovery Readiness Report* you can provide evidentiary proof and a 'what IF?' analysis of achievable protection SLAs, and actual protection SLAs based on observed recovery history.

Auditing data management controls with the Health Report

Maintaining a stable and secure *intelligent data management platform* across your AWS regions and edge locations requires continual monitoring and alarming of protection health. Commvault **Health Report** monitors several key controls that are crucial to monitor and resolve for continued recoverability.

Use the following health observations to tune your Commvault Data Management platform:

- **Anomaly jobs** that exceeded normal run time by at least 1 hour.
- **Disaster recovery** backups that have not run in alignment with best practices.
- **Backup copy jobs** that are lagging on their production snapshot copies, putting service-independent recovery or mobility at risk.
- **Secondary copies** that are lagging from their production region, putting cross-region DR recovery at risk.
- **Security assessment** which identifies insecure settings, required audit settings, and encryption of backup coverage.
- **Software currency** that tracks the software currency of Commvault software for validation against your software update policies.
- **Strike count** identifies resources experiencing repeated and systemic failures in protection operations.
- **Resources not protected in 60 days** highlights AWS resources and edge workloads that have not been protected for at least 60 days.

Protecting at the edge with AWS Outposts

In December 2019, Amazon announced the general availability of **AWS Outposts** (refer to **Announcing General Availability of AWS Outposts**). AWS Outposts provides local compute, storage, and networking resource from Amazon, within remotely owned/operated data centers.

Refer to the **Commvault + Amazon Reference Architecture** for a high-level view of the protection architecture.

AWS Outposts is designed for workloads that require low latency or must remain on-premises due to regulatory or data residency requirements. Commvault announced support for AWS Outposts in December 2019 (refer to **Commvault Data Protection Software Fully Tested and Validated to Support AWS Outposts**) and continues to extend this capability in line with Amazon advancements.
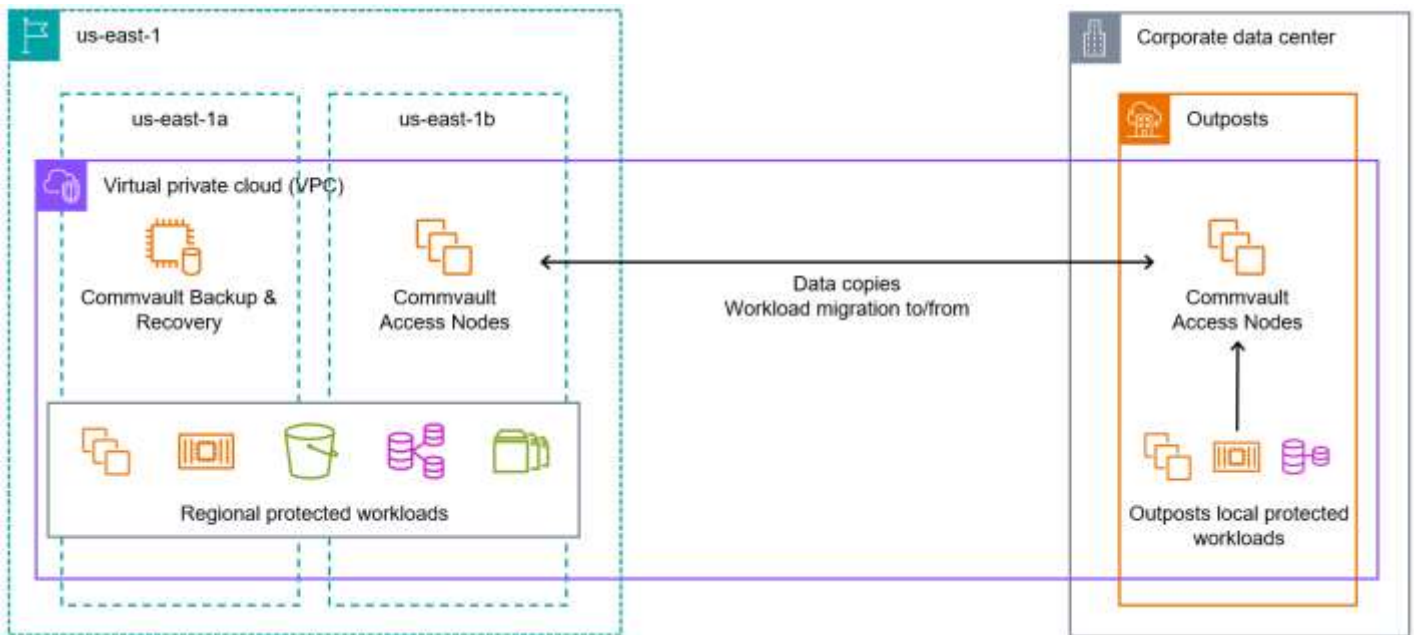
Commvault supports AWS Outposts for the following primary use cases:

- Protection of AWS Outposts workloads.
- Migration of Amazon workloads between the Outposts and the Amazon region.
- Replication between AWS Outposts, region, and on-premises

It should be noted that:

- Commvault Data Management components are fully supported when running **on AWS Outposts**.
- Commvault can **protect AWS Outposts** workloads, both within the Outposts and back to the Amazon region.
- Commvault components (CommServe, MediaAgent, Access Nodes) may all be deployed in a high availability architecture(s) **across the AWS Outposts and the Region**.

Extending protection into Outposts

When utilizing AWS Outposts to host Commvault workloads, a hybrid protection approach is recommended. Commvault Access Nodes may be deployed locally within the Availability Zone they protect (see below)

377
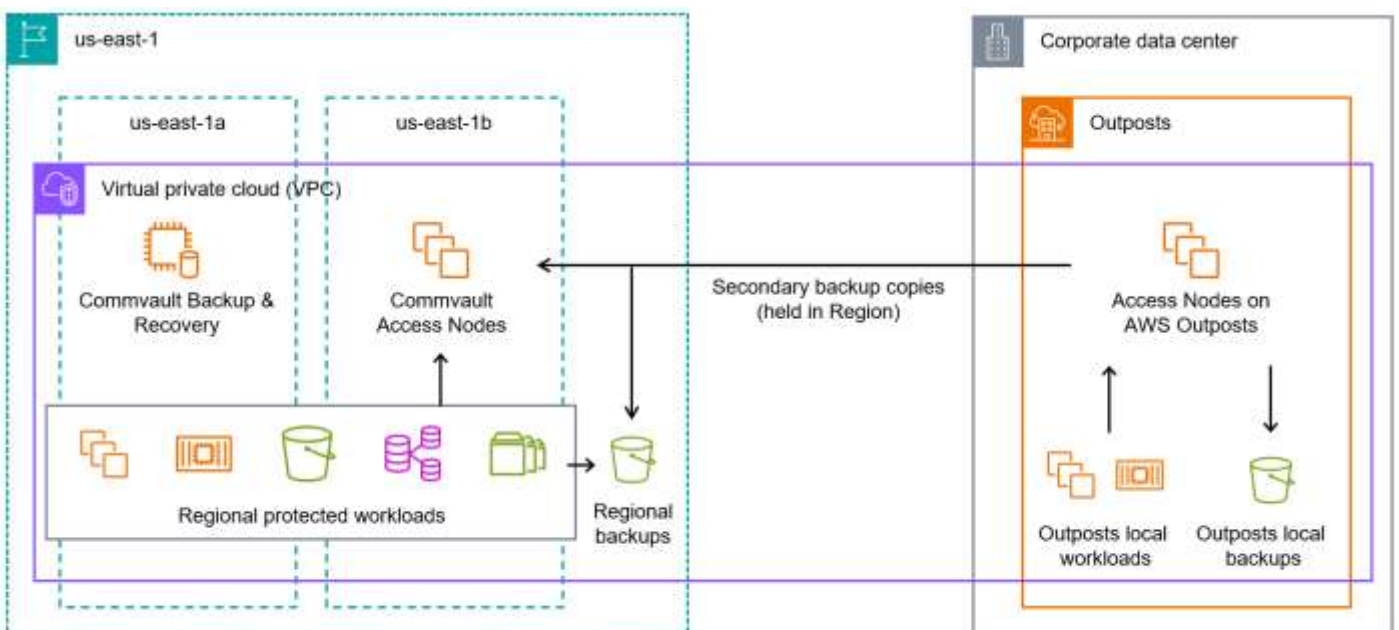
Commvault Access Nodes orchestrate the creation of Amazon snapshots and may perform application data transfers (Amazon RDS dump/export, Amazon EKS protection).

It is recommended that:

- Access node groups are used to provide failover for backup/recovery operations between the Outposts and the Amazon region.
- Access nodes are deployed locally within the AWS Outposts, dedicated to Outposts protection
- Access nodes within the region may be used as failover instances, if/when the Outposts Access Nodes are not available or fully consumed with other protection activities.

AWS Outposts Protection

Commvault protects the following workloads when running on AWS Outposts:

- Amazon EC2 instances and underlying Amazon EBS volumes
- Amazon EKS applications and underlying Amazon EBS, EFS storage.
- Amazon RDS instances
- Amazon S3 storage

- Refer to **Protecting AWS Outposts** for additional details.

AWS Outposts EC2 Compute

AWS Outposts supports General purpose (M5), Compute-optimized (C5), and Memory optimization (R5) instances which are all supported and recommended instance types for Commvault CommServe, MediaAgents, and Access Nodes (refer to **AWS Outposts – Features**).

Commvault **Automatic Scaling for Amazon Access Nodes** may be used to dynamically provision Access Nodes during protection operations, however, reserving a portion of your AWS Outposts for data management activity is considered the best practice to ensure compute resource is always available.

Amazon Elastic Kubernetes Service (EKS) Applications

Commvault protects Amazon EKS applications deployed on AWS Outposts. Commvault speaks natively with the Kubernetes kube-apiserver to automatically discover and protect applications. Commvault leverages the Container Storage Interface (CSI) to perform storage-level snapshots (where snap external snapshotter sidecar is supported). Commvault has been validated with the following Amazon CSI drivers:

- **ebs.csi.aws.com**

- **efs.csi.aws.com**

AWS Outposts S3 Storage

Commvault supports Amazon S3 (within AWS Outposts) for storing backup data and protecting it as a primary data source. As application modernization migrates data from traditional block-based and file-based storage solutions, Amazon S3 is becoming the primary storage location for applications. Commvault recommends protecting Amazon S3 data when it is the primary persistence layer of your application.

Commvault can direct backup data to a local Amazon S3 Cloud Library to ensure that data is kept local to protected workloads, and stored in an optimized format (deduplicated, compressed) to maximize AWS Outposts S3 investment.

> ⚠ **Important Warning**
>
> At the time of writing (October 2022), AWS Outposts does not support the EC2 CreateAMI action, and as such, Amazon EC2 backups are stored within the Amazon region. If backups must not exit the AWS Outposts location, Commvault recommends using an agent-in-guest protection method.

Fault domain separation for backup data

Commvault recommends that your backup data is stored in separate AWS Outposts to your primary workload to ensure that a failure event affecting your primary application workload does not also affect your backup data. If only a single Outposts rack exists, consider a **Backup Copy** back to the Amazon region to protect the backup data from an Outposts-wide failure event.

AWS Outposts Migration

Commvault backup data can be used to migrate applications into the AWS Outposts, or out of the AWS Outposts back to the region. Commvault supports the migration of the following workloads from the Amazon region into the AWS Outposts to accelerate the adoption of your Outposts investment:

- Amazon EC2 instances and underlying EBS volumes.
- Amazon RDS instances.
- Amazon EKS applications and underlying EBS or EFS storage.
- Amazon S3 buckets/objects.
- VMware Cloud on AWS (VMC) instances.

Recovery of Amazon EC2, RDS, EKS, and S3 data allows the seeding of new application environments within the Outposts or the region, depending on your development needs. Self-service migration may be performed directly by the application owner(s) within the Commvault Command Center™, removing the need for specialist knowledge on AWS Outposts.

See **Migrating Data Between AWS Outposts, AWS Regions, and On-Premises** for more information.

AWS Outposts Replication

In distributed application environments it is common to replicate data between one or many application locations. Commvault supports replication and synchronization of data between AWS Outposts, Amazon region, and on-premises deployments.

Commvault supports replication for the following data types:

- **Virtual Machine replication** (periodic backup/restore, direct snapshot replication)
- **File-system replication**
- **Database replication**
- **PostgreSQL**
- **Oracle / Oracle RAC replication**
- **Microsoft SQL Server**
- **SAP HANA**
- **DB2 replication**

Agent-in-guest (streaming)

Commvault provides the ability to install a small software agent within the operating system of infrastructure-based workloads (i.e., Applications residing on Amazon EC2 instances, or on-premises). An Agent-in-guest approach protects a wide variety of operating systems and applications.

See the full list of **Backup Agents** for supportability.

Agent-in-guest protection identifies data for protection, deduplicates, compresses, and optionally encrypts the data before transferring to a Commvault MediaAgent which writes the data out to an Amazon S3 (or equivalent) Cloud library.

Commvault can then replicate the data in a reduced format to one or many geographic regions, providing an alternate copy for Disaster Recovery.

When to Use Agent-in-Guest Approach

- When you require application-consistent backups – use an agent-in-guest either standalone or in conjunction with an Amazon Access Node backup. Deployment of agents can either be deployed via automation by Commvault® software, incorporated into AMI templates using de-coupled installation, or deployed as part of a continuous deployment method (i.e., Puppet, Chef, Ansible).
- When you require granular-level protection and restoration features for applications – the Commvault® iDataAgents can deliver granular-level protection for supported application workloads, such as SQL Server or Oracle Database, in comparison to a Full VM or File-level approach.

Architecture Requirements for Agent-In-Guest

- Minimum 1x iDataAgent per Amazon EC2 instance for the intended dataset (i.e., SQL database, File). Multiple iDataAgents can be deployed on the same Amazon EC2 instance.
- Minimum 1x MediaAgent per region. MediaAgents connect to the target object storage and can either be deployed on the same Amazon EC2 instance as the dataset iDataAgent or on a dedicated host for a fan-in configuration. The Amazon EC2 instance specifications of the MediaAgent should match the MediaAgent specifications within this Architecture Guide.
- Check the Systems Requirements section in the **documentation** to determine if the iDataAgent supports your application (see **Backup and Restore Agents**).

Architecture Recommendations

- The use of multiple readers to increase concurrency to the object storage target is recommended.
- Use of the **Amazon S3 VPC Endpoint** is highly recommended to improve throughput to/from Amazon S3 buckets.

Performing Disaster Recovery in the Cloud (Auto Recovery)

Commvault provides the ability to leverage the same intelligent data management activities used for backup and recovery to provide self-service **Auto Recovery** (formerly Disaster Recovery) to your business. Commvault supports disaster recovery by replicating Virtual Machine, Database, and File-system changes between your primary region and one or many secondary or 'DR' regions/sites.

Commvault provides **unified data** management for your backup & recovery and Disaster Recovery workloads. When using Commvault to perform Disaster Recovery replication, you can vary your replication schedules based on the individual application, ensuring a cost-optimized approach to disaster recovery.

Commvault **optimizes disaster recovery** by leveraging changed block tracking, deduplication, and compression to speed up the collection and replication of DR data. Additionally, the use of deduplication ensures that replication cross-region does not attract network egress fees from Amazon.

How it works
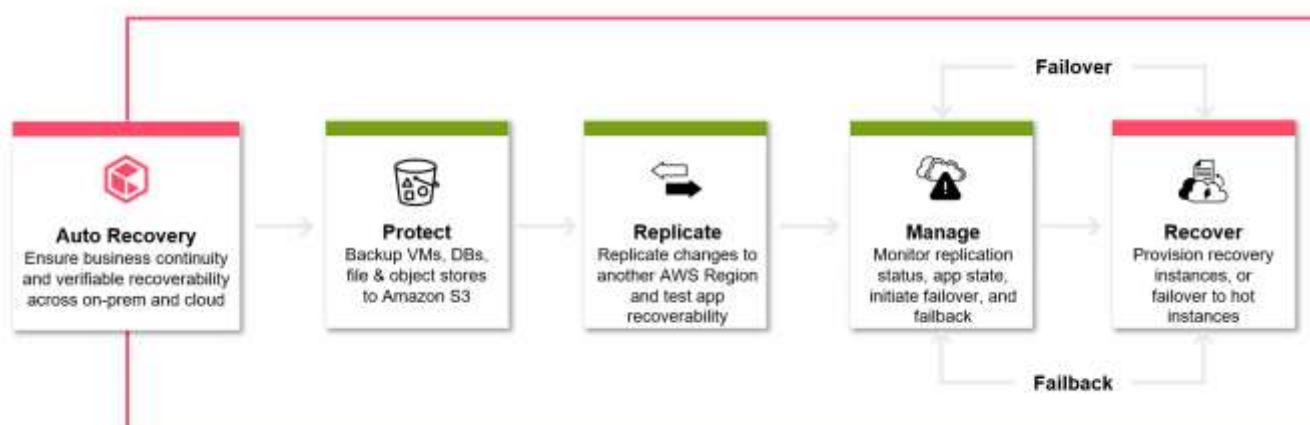
Use-cases

- **On-premises to AWS**
  Quickly recover mission-critical operations affected by datacenter-wide events, by restoring virtual machines, databases, and file systems to AWS with configurable RPOs and RTOs from hours to minutes.
- **Cloud to AWS**
  Use the elasticity and resilience of AWS to recover your other-cloud compute, database, and storage workloads into AWS native services to meet business dual-sourcing or compliance requirements.
- **AWS Region to AWS Region**
  Increase the availability of your AWS-based applications by adding cross-region replication and recovery to your most critical applications and AWS resources.

How to get started

You will require the following to get started with Commvault Auto Recovery:

- At least one **Replication group** for identifying the source workloads to protect and replicate to the DR location.
- At least one **Recovery Target** for identifying the DR location, frequency of replication (RPO), and method of replication.
- At least one MediaAgent to read/write to Amazon S3 storage in the DR region.

- At least one Access Node to write changes to the replicated workload in the DR region.

On the source site or location, it is expected the following resources already exist:

- At least one MediaAgent to receive backup data, write locally, and replicate to MediaAgent in the DR region.
- At least one Access Node to collect backup data from the workload and forward it to the local MediaAgent.
- When performing agent-in-guest protection of databases and/or file-systems, there is no Access Node required.

Supported AWS Regions

Commvault Auto Recovery is supported in all Regions (including GovCloud and China regions).

Commvault Auto Recovery is also supported to/from edge-based locations, including AWS Local Zones and AWS Outposts (subject to availability of the workload type in the edge location).

On-demand Cross-Vendor Recovery

Commvault Backup & Recovery provides a broad selection of Cross-Hypervisor restores or **VM conversions** from on-premises and Cloud VMs to Amazon EC2 instances.

See **Cross-Hypervisor Restores (VM Conversion)** for a full list.

Cross-vendor recovery can be used as an on-demand DR solution, only requiring the relevant VM backup to be replicated to the destination cloud in preparation for failover. During a DR event, an authorized backup administrator can use the Command Center console, API, command-line, or SDK to initiate the recovery of the VM.

The remainder of this section focuses on application types natively supported by the **Commvault Auto Recovery** *product.*

Supported workloads

Commvault Auto Recovery provides replication and disaster recovery for the following workloads:

Source workload	Destination	Replication	Failover/Failback
Amazon EC2	Amazon EC2	✓	Planned failover Unplanned failover Failback Test failover
VMware VM	Amazon EC2	✓	Planned failover Unplanned failover Failback
VMware VM	VMware VM	✓	Planned failover Unplanned failover Failback Test boot Test failover Automatic failover

Oracle database	Amazon EC2 instance running Oracle	✓	
SAP HANA database	Amazon EC2 instance running SAP HANA	✓	
SQL Server database	Amazon EC2 instance running SQL server	✓	
File-system data	Amazon EC2 instance running equivalent file-system	✓	
File-system data (Windows)	Amazon EC2 instance running Microsoft Windows	✓	
Amazon S3	Amazon S3	✓	
Amazon S3 Azure Blob Azure File	Amazon S3 Azure Blob Azure File	✓	
Hadoop	Amazon EC2 instance running Hadoop	✓	

References to Amazon EC2 include AWS Local Zones and AWS Outposts.

References to VMware VMs include VMware Cloud™ on AWS (VMC) and VMware Cloud™ on AWS Outposts.

File-system data includes GlusterFS, GPFS, Lustre, NAS, Nutanix, and Qumulo data.

Initial setup

Replication Groups

To start replicating workloads for disaster recovery, you will require a **Replication Group**.

The replication group is responsible for VMs, File-servers, SQL server databases, Oracle databases, SAP HANA databases, Hadoop instances, and Object storage.

During the setup of a Replication Group, you identify:

- Source **content** to protect and replicate to DR location.
- **Target** location to the replicated content (an AWS account, region, service, or instance).
- **Storage** location to write temporary cache information and/or indexing information on replication status.
- Specific workload **overrides** are required to failover the instance successfully (hostname, IP settings, etc.)

Recovery Targets

A **Recovery target** configures the location to recover replicated workloads or data. It includes:

- **Destination** service or instance and the credentials to perform provisioning of resources for replication and failover.

- **Availability zone** to provision recovered compute instances into.
- **Volume type** to set the type of Amazon EBS storage selected for all instances in a group, or all volumes on a specific instance.
- **Encryption key** to encrypt the restored Amazon EBS volumes.
- **Network** to provision elastic network interface(s) into.
- **Security group(s)** to attach to network interfaces
- **Instance type** and size of newly created compute instances will use CPU, memory, and operating system from the source instance to pick an appropriate instance type and size. May be manually selected.

Warm site vs. Hot site replication

For virtual machine migration, there are two methods of performing replication:

- Hot site replicates the changes, provisions a new instance with the changes, deletes the previous instance for every replication event
- Warm site replicates the changes to Amazon S3 storage in the destination location but defers the creation of Amazon EBS volumes, and Amazon EC2 instances until the DR events

Hot site replication

1. Create an Amazon EBS snapshot in the source location.
2. Replicate changes by creating a new Amazon EBS snapshot using Amazon EBS direct APIs in the destination location.
3. Populate the snapshot with blocks from the latest backup using a regional access node.
4. Provision a new Amazon EC2 instance matching the parameters in the *Recovery Target*.
5. Delete the previous Amazon EC2 instance and Amazon EBS volumes.

> ⓘ **Note**
>
> A snapshot of the EBS volumes is taken and retained as an integrity snapshot for reference during the next incremental replication. Snapshots will have a `CV_Integrity_Snap` tag attached to identify them.

Warm site replication

This process only occurs once a planned or unplanned failover is initiated using the Commvault Command Center, API, command-line, or SDK.

1. Create an Amazon EBS snapshot in the source location.
2. Replicate changes by creating a new Amazon EBS snapshot using Amazon EBS direct APIs in the destination location.
3. Populate the snapshot with blocks from the latest backup using a regional access node.
4. Delete the previous Amazon EC2 instance and Amazon EBS volumes.

The Amazon EC2 instance will only be provisioned during the DR event.

> ⓘ **Note**
>
> A snapshot of the EBS volumes is taken and retained as an integrity snapshot for reference during the next incremental replication. Snapshots will have a `CV_Integrity_Snap` tag attached to identify them.

See the Replication Process using **Amazon EBS Direct APIs**.

Preparation of on-premises VMs for Conversion

You will need to prepare your source-site VMs by pre-installing the Amazon operating system (OS) drivers and preparing the OS configuration to boot within Amazon EC2.

See Requirements for **VM Replication** for details.

Replication of databases, file-servers, and object stores

Replication of databases, file-servers, and object-stores occurs automatically, while recovery is *manually initiated* by the administrator breaking or *pausing* a replication group during a failover event.

The replication process is identical to virtual machines. except that the destination workload must already be running and accessible by the destination site access node.

Replication process

1. Perform backup of the database, file-server, object store at the source location.
2. Replicate backup data to the destination location, typically stored in Amazon S3.
3. Once replication is complete, apply changes to a running database, file-system, or object store.

Database changes are made by the

Monitoring and initiating failover and failback

You can use the **Disaster Recovery dashboard** for an *at-a-glance summary* of your replication health.

You can monitor the synchronization status of all configured workloads using the **Replication Monitor.**

For virtual machine or compute instance workloads, you can perform:

- Planned failover
- Unplanned failover
- Test failover
- Failback

> ⓘ **Note**
>
> There are no predefined alerts for failed or delayed synchronization. You may create **custom alerts** to publish events to administrators or Amazon CloudWatch Log metrics when static thresholds are breached.

Additional Resources

- **Disaster Recovery (DR) Architecture on AWS, Part III: Pilot Light and Warm Standby**.
- **Disaster Recovery of Workloads on AWS: Recovery in the Cloud**.
- **AWS Well-Architected Labs – Disaster Recovery**.
- **AWS Pattern – Implement cross-Region DR**.
- **AWS Strategy – Disaster recovery strategy for databases on AWS**.
- **AWS Guide – Disaster recovery options for VMware Cloud on AWS**.

Cloud migration made simple

As your organization looks to accelerate cloud adoption and modernize your application landscape with AWS, you will be faced with the challenge of *Where do I start*? Fortunately, AWS guides how to start assessing and migrating your applications from small and simple applications to large complex, and mission-critical applications.

See **Migration to AWS: Best Practices and Strategies**.

Commvault complements these strategies by accelerating **Portfolio discovery and Planning**, and **Designing, Migrating, and Validating Applications.** Commvault is a member of the **AWS ISV Workload Migration Program (WMP)**, which provides AWS Promotion Credits to offset migration costs for customers – including labor, training, and running parallel environments during migration.

Commvault Backup & Recovery is your single unified data management platform for AWS resources, Cloud, Containers, SaaS, and traditional workloads and has the complete view of your organization's application landscape. Amazon recommends **automating migration discovery and repetitive migration tasks**, Commvault provides intelligent automation that accelerates the identification of workloads and migrates them using one of the **6 Rs** strategies (detailed below).

Re-using your Commvault Backup & Recovery system and existing data management practices accelerate your migration efforts, meaning you can spend more time on your application modernization.

Migration strategy

Amazon recommends six different strategies be employed to assess, migrate, and modernize workloads as they migrate to AWS Cloud, made popular in the **Six Strategies for Migrating Applications to the Cloud** blog post. Migration assessment may decide to **Retire** an application, providing the seventh R.
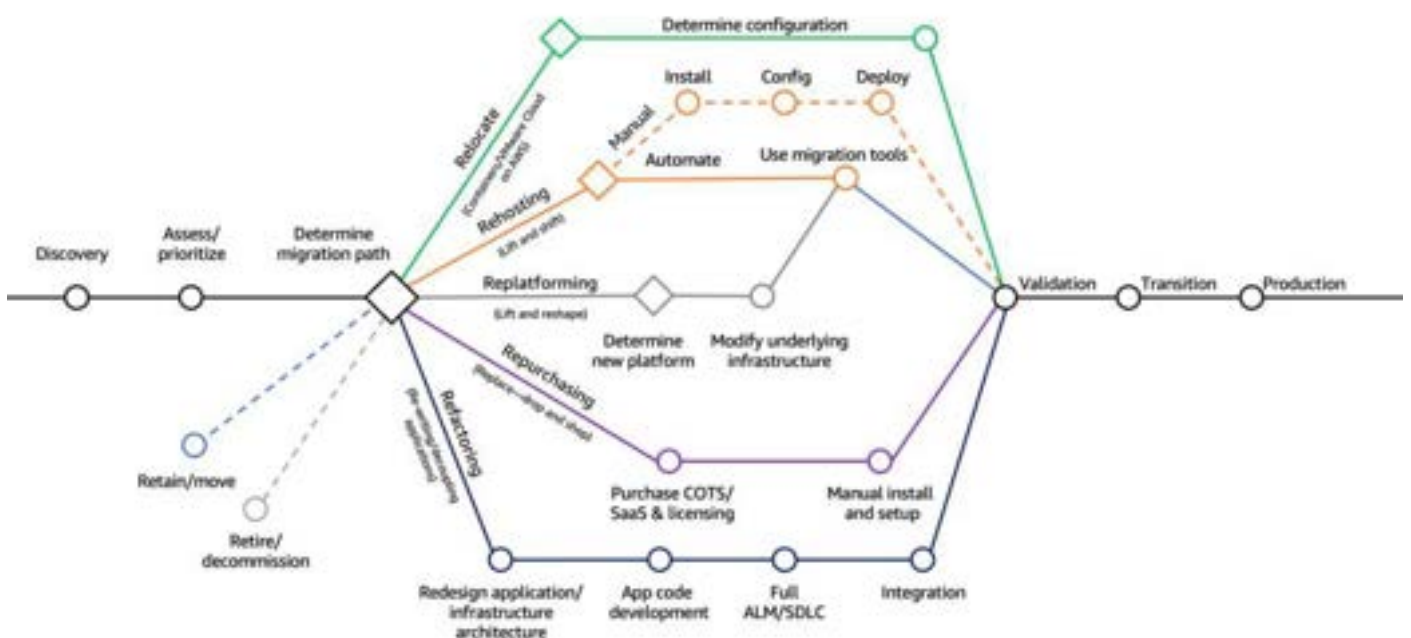
Source: **7 Strategies for Migrating Applications to the Cloud, introducing AWS Mainframe Modernization and AWS Migration Hub Refactor Spaces**, Jonathan Allen, 30 NOV 2021 AWS Cloud Enterprise Strategy Blog.

Commvault Backup & Recovery is the one *data management platform* in your organization capable of providing data-driven guidance and automation for your migration efforts. Your Commvault data management platform contains

details on all your applications and workloads being used by the business and an indication of the business value or **per-application data classification**.

Amazon classifies the migration strategies by the level of effort required to complete a migration from easiest (Retire) to hardest (Refactor).

See **Migration terms – 7 Rs**.

Retire

Retiring an application is the simplest migration – simply archive the application, then decommission its infrastructure from your IT portfolio. Commvault provides **Backup Job Summaries** and Client Monthly Growth reports that can show you the amount of data change occurring in the application backups over time. **File Risk Analysis** can show you a breakdown of the age of files to better understand how *active* an application is.

You may choose to write an application-consistent backup, virtual machine backup, or **archive the files** to your **Commvault Combined Storage Tier** storage, which stores your archive on cost-effective Amazon S3 Glacier Flexible Retrieval, and Amazon S3 Glacier Deep Archive. You can still extract value from retire data leveraging Amazon S3 Batch Operations to increase restore throughput (see **faster restores from Amazon S3 Glacier**).

Retain (revisit)

Retaining an application refers to *doing nothing* or leaving an application in-place within your owned or operated data center and self-managed infrastructure. You will likely revisit this application in future evaluation phases, but for now, continue to **protect your application** with Commvault Backup & Recovery so you have the *data-driven insight* to inform future migration decisions.

Relocate

Relocating refers to a targeted *hypervisor-level lift and shift*, specifically for VMware workloads which can migrate to **VMware Cloud™ on AWS** and **VMware Cloud™ on AWS Outposts**. Commvault provides VMware-integrated snapshot backup & recovery for your on-premises VMware VMs today. Using Commvault **out-of-place recovery** and **replication technology**, you can relocate your VMware VMs and application workloads to AWS Cloud.

See **VMware Cloud on AWS**.

Rehost

Rehosting refers to an application or infrastructure *lift and shift*, without making any modifications to the application architecture to leverage cloud-native capabilities. Commvault provides the **broadest industry coverage** for lift and shift migration across **Cloud compute**, **Virtual machines**, **Databases**, **Containers**, and traditional **applications**.

Commvault can natively migrate **on-premises virtual machines directly into Amazon EC2** instances, automatically provision and migrate **Oracle databases onto Amazon EC2**, and **SQL Server databases onto Amazon EC2**.

A migration for an application-integrated backup is as simple as an *out-of-place* restore for Commvault Backup and Recovery. If you have protected an application on-premises and have an Amazon EC2 instance with Commvault application agents installed – you can rehost.

See **Cloud Migration**.

Repurchase

Repurchasing is typically more concerned with changing your application purchasing strategy and moving from on-premises owned and operated software to adopting Software as a Service (SaaS) solutions. Commvault continues to protect your SaaS applications after migration, including:

- Microsoft Office 365 (O365)
- Microsoft Dynamic 365
- Salesforce
- GitHub
- Azure DevOps
- MongoDB Cloud Atlas

You may also choose to repurchase your Commvault Backup & Recovery platform by leveraging **Metallic Data Management as a Service (DMaaS)**, which offers Commvault-powered Backup & Recovery in a SaaS acquisition model.

Replatform

Replatforming is often referred to as *lift-tinker-and-shift* as it allows for minor changes as applications and infrastructure are lift-and-shifted from your on-premises location to AWS Cloud. Commvault has several solutions to assist your application optimization as part of migration:

- Replatform compute instances, containerized applications, databases, and file and object storage onto **AWS Outposts** for consistent hybrid cloud operations (including Commvault Backup & Recovery).
- Replatform **Oracle databases**, and **SQL Server databases** into fully-managed Amazon RDS databases.
- Replatform Commvault data management infrastructure onto **AWS Graviton instances** for improved price, performance, and overall sustainability.
- Replatform existing file storage (**file-servers**, NAS devices) into elastic Amazon S3 cloud storage.

Refactor or re-architect

Refactoring is the most complex and time-consuming migration as it requires modifying the application architecture to take advantage of cloud-native features and services. This can involve changing database technologies, and operating systems, or complete architectural rewrites from monolith-based applications to containerized and micro-services architectures.

Commvault has several solutions to help in re-factoring existing protected applications into AWS Cloud, including:

- **Migrate AIX, Solaris, and HP-UX Oracle databases to Linux**
- Migrate on-premises **MySQL** and **PostgreSQL** databases to Amazon Aurora serverless instances.

Once an application is re-factored into a containerized application complete with cloud-native storage and databases, Commvault will continue to protect your application with protection for:

- Amazon EKS
- Amazon Aurora
- Amazon RDS
- Amazon RedShift
- Amazon DocumentDB
- Amazon DynamoDB
- Amazon S3
- *many others…*

See the **Re-Architect Pattern List** for common re-architecture approaches.

Discovery

Any Cloud migration effort starts with a ***Discovery Phase***.

You can use your existing Commvault Backup & Recovery system to identify what virtual machines, containers, databases, file storage, object storage, and traditional applications you are using. Ideally, you will be using **Server plans** or **entity tags** to identify the underlying business value, data sensitivity, or ***data classification*** of your

protected workloads. Start with a <u>top-down approach</u> that identifies applications by environment production, pre-production, dev-test, and sandbox.

Start with the simplest workloads or environments first, then progress to the next environment until Production workloads are reached.

Start with **Backup Health Reports** which identify Virtual machine, File storage, Database, and Client group protection statistics (total protected instances).

Migration

The following section provides guidance on the migration approach to use for each workload type.

Virtual machines

Commvault Backup and Recovery takes virtual machine backups from Azure, Azure Stack HCI, Hyper-V, and VMware (including VMware Cloud™ on AWS) and converts them to Amazon EC2 instances (**see Cross-Hypervisor Restore Support**). You can migrate to Amazon EC2 running in the AWS Region, AWS Local Zones, or AWS Outposts.

Conversion requires preparing the source virtual machine by pre-installing the PV drivers, EC2 drivers, and NVMe and ENA drivers for nitro-based instances.

Commvault allows prescriptive selection of AMIs to be used for migration and supports migration of Universal Extensible Firmware Interface (UEFI) and Basic Input/Output System (BIOS) boot modes. Commvault supports migration of Operating System (OS) and data volumes, both GUID Partition Table (GPT) and Master Boot Record (MBR) volumes (see **Convert and Replicate VMs with Unified Extensible Firmware Interface (UEFI) Boot to AWS**).

Migrate uses one of three methods (in order of migration speed):

- **Amazon EBS direct API** migration creates Amazon EBS snapshots, populates them with backup data, and then provisions an Amazon EC2 instance. Migration uses **Amazon EBS direct APIs** from a regional, zonal, or on-premises Access Node.
- **Commvault HotAdd** migration creates Amazon EBS volumes, attaches them to a regional or zone Access Node, populates with backup data, and then provisions an Amazon EC2 instance. This approach delivers production-grade EBS volume performance by avoiding **EBS volume initialization** (if using an Access Node in the target Zone).
- **Amazon VM Import/Export** migration leverages the **Amazon VM Import/Export** service and uploads VHD formatted volumes to Amazon S3 from backup data, then requests conversion and provisioning of an Amazon EC2 from the Import/Export service. *This method is the slowest migration method.*

Commvault recommends always staging a copy of the source VM backup data in Amazon S3 and performing the migration with an Amazon EC2 instance using VPC Endpoints to reach the EBS and S3 service for optimal migration performance.

See below for a high-level summary of the differences between migration approaches.

Criteria	Amazon EBS direct APIs	Commvault HotAdd	Amazon VM Import/Export
Cost	Requires a backup	Requires a backup	Free

Recovery speed	Fastest	Fast	Slower
Requires Commvault Access Node running in AWS	Recommended	Mandatory	
Requires preinstallation of AWS drivers on VM	Yes	Yes	
Supports Windows Server 2019, 1803, 1709, 2016, 2012 R2, 2008 R2	✓	✓	✓
Supports Windows Server 2003 R2, 2003			✓
Supports Windows 10	✓	✓	✓
Supports Windows 8.1, 8, 7			✓
Supports CentOS 8-7.x, RHEL 8-6.x, RHEL, Ubuntu 22-14.x, Oracle Linux 7.x-6.x	✓	✓	✓
Supports Oracle Linux 8.x	✓	✓	
Supports Amazon Linux 2, CentOS 5.1-5.11, Debian 6.0.0-6.0.8 / 7.0.0-7.8.0 / 8.0.0, Fedora Server 18-21, Oracle Linux 5.10-5.11, RHEL 5.1-5.11, SUSE Linux 11 SP1-4 / 12 SP1-3 / 15, Ubuntu 12.04, 12.10, 13.04, 13.10			✓
Limited to 21 volumes (max.) per guest		✓	✓
Supports selecting of licensing model for converted VM (AWS supplied, BYOL)			✓
Requires antivirus or intrusion detection software disabled on the source			✓

Databases

Commvault Backup and Recovery takes your database backups and restores them into Amazon EC2 compute instances, Amazon RDS provisioned instances, and Amazon Aurora serverless instances. You can migrate to Amazon RDS databases running in the AWS Region, AWS Local Zones, or AWS Outposts.

> ⓘ **Note**
>
> SQL Server database conversion requires that the on-premises database backup used for migration be taken as a single stream backup (see **Restoring an On-Premises SQL Instance to an Amazon RDS for SQL Server Instance**).

Commvault provides three migration methods for your on-premises databases:

- **RDS Migration Workflow** is a self-service migration workflow that uses your **Oracle** full backups and a pre-created Amazon RDS instance with a database administrator account, to seamlessly restore your entire database, schemas, or tablespaces.
- **Migrate to Amazon EC2** is a self-service migration process that uses your **Oracle** and **SQL Server** full backups, and an authorized AWS IAM User or Role to optionally provision a new Amazon EC2 compute instance, Amazon EBS volumes, and then restore your database binaries and data into the EC2 compute instance.

- **Database restore** is the simplest form of migration that uses your database's full backups or dump-based backups to restore into a pre-existing Amazon RDS, Amazon Aurora, or Amazon EC2 instance running your database engine. *This method allows migration to serverless database instances.*

Each of the available migration methods available per-database technology is shown below

Criteria	Amazon RDS workflow	Amazon EC2 migration	Database restore
Amazon RDS capacity type	Provisioned	Provisioned	Provisioned Serverless
Oracle	✓	✓	✓
SQL Server		✓	✓
MySQL			✓
PostgreSQL			✓
MariaDB			✓
Use-case	*Use to migrate to fully-managed Amazon RDS instances in the AWS Region or on AWS Outposts.*	*Use to migrate to Amazon EC2 compute instances that comply with the organizational hardening policy.*	*Use to migrate to fully managed Amazon RDS, or Amazon EC2 compute.* *Optionally, migrate to fully-managed Amazon Aurora serverless instances.*

File storage

Commvault Backup and Recovery takes your file-system and network-attached storage (NAS) backups and restores them to appropriate cloud storage solutions. You can migrate per-host file storage and shared storage from your existing file servers and NAS arrays.

Migration occurs using one of three methods:

- **Migration to Amazon S3** takes file-system backup data (**Windows**, **Linux**, **Macintosh**) and restores it into an Amazon S3 bucket and storage class for migration or long-term archival.
- **Migration of NAS data** takes your CIFS/SMB and NFS backups and restores them using open-protocol SMB 2.0 and NFS v4/v4.1 protocols to cloud-native equivalent cloud file storage services. Ensure share or export migration can be **automated** for cut-over from on-premises to cloud-based shared storage services.
- **Migration of OneDrive for Business data** takes your Microsoft OneDrive for Business backup data and selectively restores it into a **file-storage destination** accessible via SMB/CIFS or NFS.

See below for the compatibility of migrating your file storage data to AWS.

Source data	Amazon FSx for Windows	Amazon FSx for ONTAP	Amazon EFS	Amazon S3
Linux file-systems		✓	✓	✓
Windows file-systems	✓	✓		✓
CIFS/SMB shares	✓	✓		✓

		✓	✓	✓
NFS exports				
Microsoft OneDrive	✓	✓	✓	

Object storage

As your business adopts more cloud-native technology options, you will find that traditional block and file storage will be replaced with infinitely scalable and high-performance object storage.

Commvault can protect the following Cloud storage locations and restore them to Amazon S3 or optionally a file storage location accessed via a regional or zonal Access Node.

- Alibaba Cloud Object Storage Service (OSS)
- Amazon S3
- Azure BLOB
- Azure Data Lake Storage
- Azure File Storage
- Google Cloud Storage
- IBM Cloud Object Storage

Commvault will migrate object metadata when restoring between Azure Blob and **Amazon S3**.

See **Object Storage Protection**.

> ⓘ **Note**
>
> Commvault supports the object storage systems listed (above) for backup and restoration. If you have another object storage system you would like to migrate, please contact your Commvault Sales Representative.

Application workloads

Commvault can perform application-integrated application migration between any compute instances running the same application software. Commvault Backup & Recovery will protect your traditional application on-premises, replicate your backup to Amazon S3 (optional), and may then be restored to an Amazon EC2 instance that has the Commvault file-system and application agent(s) installed.

Commvault has the broadest industry coverage for traditional application workloads.

Check **www.commvault.com/supported-technologies** for your workload, and **documentation.commvault.com/2022e/essential** for specific application compatibility.

Post-migration

The last step in any migration is to ensure that the new workload is both protected and can be safely recovered. Ensure before cut-over to your new AWS-powered compute or container instance, cloud database, or cloud storage that you have performed a successful backup and restore in a pre-production environment.

Additional Resources

- **6 Strategies for Migrating Applications to the Cloud**.
- **AWS Cloud Enterprise Strategy Blog – Migration**.
- **Migrating to AWS: Best Practices and Strategies**.
- **Best practices for assessing applications to be retired during migration to the AWS Cloud**.
- **Migrating to AWS: Best Practices and Strategies**.
- **AWS large-migration strategy and best practices**,
- **Lift and shift: Rehost your workload on AWS to accelerate your cloud journey**.

- **Application portfolio assessment strategy for AWS Cloud migration**.
- **Migration strategy for relational databases**.

Adopting APIs for Automation

Your Commvault data management platform comes with a rich set of developer tools aimed any automating your cloud operations at scale.

How it works
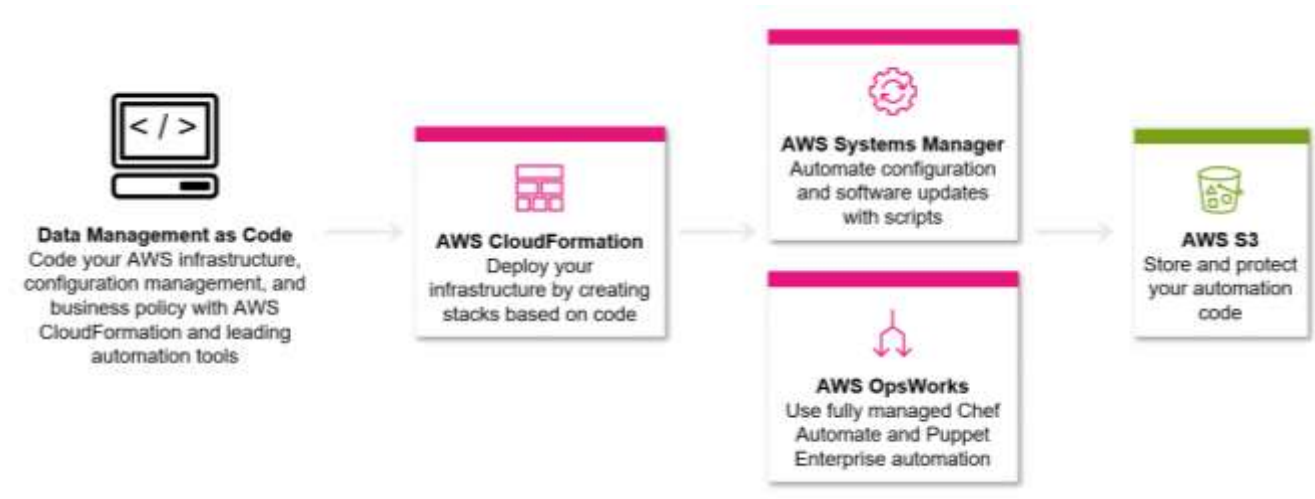
Use cases

- **Provision and manage Infrastructure as Code**
  Automate, test, and deploy infrastructure templates for Commvault compute, storage, networking, and security resources, integrated with your continuous integration and delivery (CI/CD) automation.

- **Automate configuration management**
  Automate day-one configuration and day-two operations by centrally managing configuration, software updates, and system changes.

- **Protected automated operations**
  Store, version, and protect your automated cloud operations by storing your Infrastructure as Code (IaC) templates and automated change workloads in Commvault-protected Amazon S3, GitHub, and Azure DevOps.

How to get started

Checkout the developer tools available at **docs.commvault.com**, these include:

- **REST API** actions and service endpoints @ **github.com/Commvault/Rest-API-Postman-Collection**.
- **Commvault Terraform module** @ **github.com/Commvault/terraform-provider-commvault**.
- Commvault Ansible modules @ **github.com/Commvault/ansiblev2**
- Python SDK @ **github.com/Commvault/cvpysdk**.
- PowerShell Module @ **https://github.com/Commvault/CVPowershellSDKV2**.

You can learn more and begin testing with the 800+ restful APIs available for Commvault data services with Postman at **api.commvault.com**.

Commvault software can be automated with your favorite automation tools, with customers leveraging Commvault command-line and other developer interfaces with:- AWS CloudFormation, Ansible, Azure ARM, Cfengine, Chef, Crossplane, GitOps, Google Deployment Manager, Pulumi, Puppet, Saltstack, Terraform, and Vagrant.

Multi-cloud mobility

While this guide (Public Cloud Architecture Guide for AWS) is focused primarily on **Amazon Web Services** protection, many organizations employ multiple cloud solutions to meet their business needs. Some industry regulators will require that *disaster recovery and backup be implemented outside the primary cloud in which workloads are deployed*, when this occurs – **Commvault hybrid multi-cloud mobility** is crucial.

Only Commvault provides the **broadest industry support** for your hybrid multi-cloud, containers, SaaS, and edge-based applications

---

How it works

Commvault Backup & Recovery collects your virtual machine/instance, containers, databases, and storage data and transfers it to Commvault in a service-independent format. This service-independent format can then be used to restore your data to an alternative cloud or edge-based location. Commvault records common metadata like resource tags and restores them to the new location, allowing consistency of cloud operations as your workloads migrate location.

Use Cases

Multi-cloud mobility simplifies hybrid multi-cloud data management by removing the requirement to maintain per-cloud and per-application runbooks to recover and migrate your applications.

- **Migrate applications between cloud providers and locations**
  Move virtual machines/instances, containers, databases, applications, and storage data between clouds and edge locations.

- **Perform disaster recovery to the cloud**
  Recover mission-critical workloads between clouds and edge locations – either one-time or replicated periodically and powered on during disaster events.

- **Seed cloud environments for development**
  Restore copies of your compute, container, database, and storage workloads between clouds and edge locations to seed sandbox, dev/test, or production environments.

How to get started

Consult the **compute**, **container**, **database**, and **storage** mobility migration destination (below) to validate that your workload migration scenario is supported.

Configure your **destination cloud** to enable migration.

Run an **out-of-replace recovery** for your workload to your destination cloud.

Consider using **offline migration** using AWS Snow family, Azure Data Box, Google Transfer Appliance, or Oracle ZFS Appliance if you have limited network bandwidth in your destination cloud.

Compute services

Source cloud	Destination cloud											
	Alibaba	AWS[1]	Azure	Azure Stack Hub	Azure Stack HCI	Google Cloud Platform	Hyper-V	Nutanix AHV	OpenStack	Oracle VM	Oracle Cloud Infrastructure	VMware[2]
AWS	--	Yes	Yes	--	--	Yes	--	--	--	--	--	Yes
Azure	--	Yes	Yes	Yes	Yes	Yes	Yes	--	--	--	--	Yes
Azure Classic	--	--	Yes	--	--	--	--	--	--	--	--	--
Azure Stack Hub	--	--	Yes	--	--	--	--	--	--	--	--	--
Azure Stack HCI	--	Yes	Yes	Yes	Yes	--	Yes	--	--	--	--	Yes
Google Cloud Platform	--	--	--	--	--	--	--	--	--	--	--	Yes
Hyper-V	--	Yes	Yes	Yes	Yes	--	Yes	--	--	--	--	Yes
Nutanix AHV	--	--	Yes	--	--	--	--	--	--	--	--	Yes
Oracle VM	--	--	Yes	--	--	--	--	--	--	--	--	--
VMware[2]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

1 Includes AWS Local Zones, Amazon Wavelength, and AWS Outposts.
2 Anywhere VMware: including VMware Cloud on AWS, Azure VMware Solution, Google Cloud VMware Engine, and Oracle Cloud VMware Solution.

Commvault will restore common configuration metadata when restoring across vendors (i.e., restoring tags on virtual machines and **object storage** objects).

Container services

Commvault supports the protection and migration of stateless and stateful Kubernetes container-based **applications** on any **Cloud Native Computing Foundation (CNCF) certified distributions**.

Source cloud	Alibaba ACK	Amazon EKS	Amazon EKS-D	Amazon EKS on Outposts	Azure AKS	Azure Stack Hub	Azure Stack HCI	Google GKE	Google Anthos	HPE Ezmeral	Nutanix Karbon	Oracle OKE	Red Hat OpenShift	VMware Tanzu
Alibaba ACK	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Amazon EKS, EKS-D, EKS on Outposts, ROSA	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Azure AKS, AKS on Azure Stack	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Source cloud	Alibaba ACK	Amazon EKS	Amazon EKS-D	Amazon EKS on Outposts	Azure AKS	Azure Stack Hub	Azure Stack HCI	Google GKE	Google Anthos	HPE Ezmeral	Nutanix Karbon	Oracle OKE	Red Hat OpenShift	VMware Tanzu
Google GKE, Google Anthos	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
HPE Ezmeral	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Nutanix Karbon	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Oracle OKE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Red Hat OpenShift	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
VMware Tanzu	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Database services

Commvault protects fully-managed cloud databases and traditional database deployments running on virtual compute instances. Commvault leverages database-native dump and export utilities to take a full copy of your database and restore it to another cloud or compute instance with compatible database software pre-installed.

When the cloud database service allows installation of database agent-in-guest protection (i.e., **Amazon RDS Custom**), a more granular database and logs approach may be used.

Source cloud	Alibaba[1]	AWS[2]	Azure[3]	Google Cloud Platform[4]	Oracle Cloud infrastructure[5]	Traditional databases[6]
Alibaba AsparaDB (dump)	Yes	Yes	Yes	Yes	Yes	Yes
Amazon RDS (dump)	Yes	Yes	Yes	Yes	Yes	Yes
Azure Database (dump)	Yes	Yes	Yes	Yes	Yes	Yes
Google Cloud (dump)	Yes	Yes	Yes	Yes	Yes	Yes
Oracle Cloud (dump)	Yes	Yes	Yes	Yes	Yes	Yes
Traditional Databases	Yes	Yes	Yes	Yes	Yes	Yes

[1] Restore to pre-provisioned Alibaba ECS compute instances or Alibaba AsparaDB fully-managed database services, for supported source cloud databases.
[2] Restore to pre-provisioned Amazon EC2 compute instances or Amazon RDS fully-managed database services in Region or on AWS Outposts, for supported source cloud databases.
[3] Restore to pre-provisioned Azure Virtual Machines or Azure fully-managed database services in Region or on Azure Stack Hub/HCI, for supported source cloud databases.
[4] Restore to pre-provisioned Google Cloud Virtual Machines or Google Cloud fully-managed database services, for supported source cloud databases.
[5] Restore to pre-provisioned Oracle Cloud instances or Oracle Cloud fully-managed database services, for supported source cloud databases.
[6] Restore to any supported pre-provisioned database running in on-premises hypervisors, cloud edge, or cloud regions, via file-system and database agents-in-guest destination instance.

Commvault is continually working with our cloud partners to protect traditional and modern cloud databases, check the following resources to validate supportability for your source and destination cloud databases:

- **Alibaba AsparaDB**.
- **Amazon Relational Database Services** (Amazon RDS), including Amazon Aurora.
- **Google Cloud**, including Cloud Spanner.
- **Microsoft Azure**, including Cosmos DB.
- **Oracle Cloud Infrastructure**, including Database on OCI, Database on Exadata.

Commvault also supports a broad array of traditional databases that may be migrated between on-premises and cloud locations using a Commvault agent-in-guest backup and recovery. Consult **docs.commvault.com** for the latest list, a subset of supported databases is provided below:'

- DB2, Documentum, Greenplum, IBM Notes, Informix, MongoDB Atlas, MariaDB, MySQL, Oracle, Oracle RAC, PostgreSQL, SAP, SQL Server, Sybase.

Storage services

Commvault supports the migration of file-storage that is protected using open protocols (NFS, SMB) or agent-in-guest file-system protection (iSCSI volumes).

Source cloud service	Alibaba	AWS	Azure	Google Cloud Platform	Oracle Cloud Infrastructure	NetApp ONTAP	Traditional locations
File storage [1, 2, 3, 4, 5, 6, 7]							
Alibaba File Storage NAS (NFS, SMB)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Amazon EFS (NFS)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Amazon FSx (NFS, SMB)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Azure Files (NFS)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Google Filestore (NFS)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Oracle Cloud File Storage (NFS)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
NetApp ONTAP (NFS, SMB)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
File-servers (inc. Microsoft OneDrive for Business)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Block storage [8]							
Alibaba EBS	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Amazon EBS	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Amazon FSx (iSCSI)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Azure Disks	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Google Persistent Disk	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Oracle Cloud Block Volumes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
NetApp ONTAP (iSCSI)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Object storage [9]							
Alibaba Object Storage Service (OSS)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Amazon S3 (inc. Outposts)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Azure Blobs	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Google Cloud Storage	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Oracle Cloud Object Storage	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Any supported Object storage	Yes	Yes	Yes	Yes	Yes	Yes	Yes

[1] Restore to pre-provisioned Aspara File Storage NAS or file-servers running on Elastic Compute Service (ECS).

[2] Restore to pre-provisioned Amazon EFS, Amazon FSx for Windows, Amazon FSx for NetApp ONTAP, or file-servers running on Amazon EC2.

[3] Restore to pre-provisioned Azure Files or file-servers running on Azure Virtual Machines.

[4] Restore to pre-provisioned Google Filestore or file-servers running on Google Cloud Virtual Machines.

[5] Restore to pre-provisioned Oracle Cloud Infrastructure File Storage or file-servers running on OCI Virtual Machines.

[6] Restore to on-premises and managed NetApp ONTAP instances including – Amazon FSx NetApp, ONTAP Select, Cloud Volumes Service, Cloud Volumes ONTAP, and Azure NetApp Files,

[7] Restore to pre-provisioned on-premises Network Attached Storage (NAS) or file-servers running on supported hypervisors or operating systems.

[8] Restore to a supported cloud and edge locations via agent-in-guest compute or destination-supported file-storage or object-storage services.

[8] Restore to a supported cloud and edge location supported cloud storage (object-storage services and products).

Do you need help?

Do you still need help architecting, designing, and deploying your cloud-native AWS data management solution?

Consider engaging in Commvault **community** discussions and **events** to gain insight into the latest trends and advancements in cloud-native protection.

Commvault provides a wide array of services to help your organization succeed in deployment and maintaining your intelligent data management services, see the AWS Marketplace for details on:

- **Commvault Technology Consulting**

  Commvault technical consultants ensure that your data management environment is designed for optimal results, configured quickly, and easy to maintain.

- **Commvault Training**

  Learn skills to effectively manage your Commvault environment and give your career a boost. We offer content for learners at all levels. Our On-Demand Learning Library is free for customers and partners. Looking for a more structured learning environment? Register for self-paced eLearning or instructor-led courses.

- **Commvault Managed Services**

  Remote Managed Services compliments the Commvault software platform and provide results-oriented data protection to customers worldwide. Expert Commvault engineers deliver secure, reliable, cost-effective, remote monitoring and management of your Commvault software environment. You retain full ownership of your data management infrastructure while we provide secure service delivery.

- **Commvault Enterprise Support**

  Enterprise Support Program is Commvault's most comprehensive support offering and is designed to provide strategic, world-class technical management for all aspects of our customers' enterprise data management solution. We partner fully with our customers to enable their success, and to provide business stakeholders with the highest level of customer satisfaction, all while safeguarding technology investments and intellectual property.

Reach out to Commvault at **aws@commvault.com** and we can help connect you with Commvault and **Commvault partners** that can help you architect, design, deploy, and maintain your Commvault Intelligent Data Management Platform.

Commvault can integrate and operate with your cloud operations delivered by **AWS Managed Services**, reach out to us to learn more.

Additional Resources

Community Forum

Commvault Community Forum (**community.commvault.com**) is your one-stop location to discuss technical questions, connect with experts, and share knowledge and ideas. Commvault product management, customer support, and engineering all monitor and participate in discussions.

Some examples of the content you can access within the forums include:

- Technical blogs and articles
- Onboarding guides and Q&A
- Ransomware protection best practices
- Commvault Education updates
- Feature release updates

Documentation

Cloud Storage

The **Cloud Storage** documentation covers technical procedures and information on Supported Cloud Targets, Advanced procedures, Troubleshooting, and FAQ sections for Commvault customers.

AWS IAM Permissions

All required Amazon user permissions can be accessed from the documentation here:

- **JSON Templates for IAM Role Definition and User Permissions**
  - o **Performing Backups with Restricted Access** (includes Amazon EC2 and EBS protection, VM conversion, and performing backups to an S3 Library)
  - o **Performing Backups to an S3 Library**
  - o **Amazon Web Services User Permissions for Backups and Restores**
  - o **Amazon Web Services User Permissions for RDS Backup and Restores** (snap method)
  - o **Amazon Web Services User Permissions for VM Conversion**
  - o **Amazon Web Services User Permissions for DocumentDB Backup and Restores**
  - o **Amazon Web Services User Permissions for Redshift Backup and Restores**
  - o **Amazon Web Services User Permissions for RDS Snapshot-based Backup and Restores**
  - o **Amazon Web Services VM Import Role**
- **AWS Permissions – EBS Direct API / Change Block Tracking** (included in Backup/Restricted role above)
- **AWS Permission – EBS Direct API Restores** (permissions required for API-based recovery)

A full breakdown of all permissions required by Commvault for each task may be found at **Amazon Web Services Permission Usage**.

For quick and easy **IAM Role** and **Policy** definition, you may download these permission definitions from the Commvault GitHub repo – here **github.com/Commvault/aws-permissions**

Slack

Commvault operates a slack workspace **commvaultsystems.slack.com**, reach out to your Commvault sales representation for an invite to dedicated channels for Commvault masters, and API-based automation with Commvault

Solutions, References, and Videos

Datasheets and Press Releases

Commvault Named a Leader for 12th Consecutive Time in the 2023 Gartner® Magic Quadrant™ for Enterprise Backup and Recovery Software Solutions (**link**) **new**

Commvault and AWS Expand Partnership through AWS ISV Workload Migration and Service Ready Programs (**link**) **new**

Commvault Redefines Data Protection With New Security Capabilities And Ecosystem Integrations To Combat Increasingly Sophisticated Cyber Threats (**link**) **new**

Commvault Leads the Industry in Kubernetes Data Protection for Third Consecutive Year (**link**) **new**

Global Research and Advisory Firm Names Commvault a Leader in Data Resilience Solutions Suites (**link**) **new**

Commvault achieves AWS Outposts Ready designation (**link**)

Whitepaper: Commvault for Amazon Web Services (AWS) (**link**)

Datasheet: Commvault for VMware® Cloud on Amazon Web Services – Datasheet (**link**)

Commvault.com/aws – Amazon Web Services (AWS) Solutions (**link**)

Solution Briefs

Backup and recovery of Amazon EC2 Instances (**link**)

Long-term data retention with AWS (**link**)

Backup Done Differently: One Solution to Solve & Simplify All Your Backup Needs (**link**)

Disaster Recovery On Demand: Keep Your Enterprise Up and Running (**link, link**)

Commvault Data Protection Software Fully Tested and Validated to Support AWS Outposts. (**link**)

Government cloud solutions from Commvault and Amazon Web Services (**link**)

Backup to AWS (**link**)

References

Reduce recovery time and optimize storage costs with faster restores from Amazon S3 Glacier storage classes and Commvault (**AWS Storage Blog**) **new**

Webinar: Why Parsons Considers Its Data Protection Strategy a Business Advantage (**Awscloud**)

Webinar: Optimizing cloud data management with Commvault integrations for AWS (**webinar**)

University of Canberra integrates Commvault Complete Backup & Recovery with Amazon S3 and Glacier. (**YouTube**, **case study**)

Dow Jones' move to AWS with Commvault (**case study**)

College of the Holy Cross: long-term retention with Commvault and AWS Cloud (**YouTube**)

Moving Forward Faster: How Monash University Automated Data on AWS with Commvault (**YouTube**)

See all **AWS case studies**

Videos

AWS re:Invent 2022 - Build and automate for resilience with modern data protection (PRT040) (**YouTube**) **new**

AWS re:Invent 2020: Optimizing protection for AWS service workloads at petabyte scale & beyond (**YouTube**)

AWS on Air 2020: Howdy Partner Featuring Commvault (**YouTube**)

AWS re:Invent 2019: Ensuring data protection readiness across hybrid environments (**ENT323-S**)

Getting Started with Commvault in Amazon Marketplace (**Vidyard**)

Amazon Best Practices

Architecture

AWS Well-Architected
https://aws.amazon.com/architecture/well-architected/?wa-lens-whitepapers.sort-by=item.additionalFields.sortDate&wa-lens-whitepapers.sort-order=desc

Compute

Best practices for Amazon EC2
https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/ec2-best-practices.html

Security

Security best practices in IAM
https://docs.aws.amazon.com/IAM/latest/UserGuide/best-practices.html

Security of your AWS Cloud environment from ransomware
https://d1.awsstatic.com/WWPS/pdf/AWSPS_ransomware_ebook_Apr-2020.pdf

Storage

Best practices design patterns: optimizing Amazon S3 performance
https://docs.aws.amazon.com/AmazonS3/latest/userguide/optimizing-performance.html

Security Best Practices for Amazon S3
https://docs.aws.amazon.com/AmazonS3/latest/userguide/security-best-practices.html

Amazon EFS performance tips
https://docs.aws.amazon.com/efs/latest/ug/performance-tips.html

Security Best Practices for Storage Gateway
https://docs.aws.amazon.com/storagegateway/latest/userguide/security-best-practice.html

Database

Best practices with Amazon Aurora
https://docs.aws.amazon.com/AmazonRDS/latest/AuroraUserGuide/Aurora.BestPractices.html

Best Practices for Amazon DocumentDB
https://docs.aws.amazon.com/documentdb/latest/developerguide/best_practices.html

Best Practices for Designing and Architecting with DynamoDB
https://docs.aws.amazon.com/amazondynamodb/latest/developerguide/best-practices.html

Best practices for Amazon RDS

https://docs.aws.amazon.com/AmazonRDS/latest/UserGuide/CHAP_BestPractices.html

Amazon Redshift best practices

https://docs.aws.amazon.com/redshift/latest/dg/best-practices.html

Revision History

Version	Data	Changes
5.2	August 2023	**Commvault Platform Release 2023E** (**Newsletter**) • Extended Amazon VPC protection coverage. • Added new region support (UAE, Spain, Switzerland, Hyderabad, Melbourne). • Added new AWS Graviton 3 instance support (M7g, R7g). • Added AWS Graviton export-based protection of Amazon Aurora and Amazon RDS databases. • Added AWS Graviton-based protection of Salesforce CRM instances. • Added Application-Integrated protection for Amazon RDS Custom Databases (Oracle, SQL Server). • Add incremental forever backup for customer-managed Oracle databases on Amazon EC2. • Added multi-node fault-tolerant, load-balanced Amazon S3 backup & recovery. • Added File Indexing V2 for Amazon EC2 backup & recovery.
5.1	February 2023	• Added *faster restore for S3 Glacier* support. • Added Amazon S3 Dual-Layer Server-Side Encryption with Keys Stored in AWS Key Management Service (DSSE-KMS) support. **Commvault Platform Release 2023** (**Newsletter**) • Added Amazon VPC backup and recovery in-place support. • Adding VM conversion and replication for UEFI boot-enabled VMs. • Added Transport Layer Security (TLS) 1.1 and 1.2 for Amazon Aurora, and Amazon RDS MySQL. • Added Transport Layer Security (TLS) 1.1 and 1.2 for Amazon Aurora, and Amazon RDS PostgreSQL. • Added ability to add, remove, and customize AWS Resource Tags during Amazon EC2 restore. • Added IntelliSnap® support for Amazon FSx for NetApp ONTAP.
5.0	November 2022	**Commvault Platform Release 2022E** (**Newsletter**) • Added Disaster Recovery for VMware VMs Using EBS Direct APIs (**more**). • Added Warm Site Recovery for Disaster Recovery, *including DR to AWS* (**more**). • Added Amazon S3 Object-lock support for immutable Write Once Read Many (WORM) storage libraries (**more**). • Added Back Up and Restore for Entire Kubernetes Clusters and Namespaces, *including Amazon EKS and EKS-Distro* (**more**). • Enabled Cross-Region Copy of Amazon Redshift Snapshots (**more**). • Added support to restore AWS RDS snapshots to a different AWS account (**more**). • Added ability to convert Hyper-V Virtual Machines to Amazon EC2 Instances (**more**). • Expanded guest OS supportability for VMware, Azure, and Hyper-V conversion to Amazon EC2 instances (**more**).

		- Added support for Linux CommServe Server (**more**). - Added support for strong authentication using AWS Key Management Service (AWS KMS) via a designated access node (**more**). - Added Application Validation of VMware IntelliSnap Backups, *including VMware Cloud on AWS* (**more**). - Added Analyze Sensitive Data For Cloud-Based Object Storage, *including Amazon S3* (**more**). - Added ability to Use Your Own Key for Commvault-managed data encryption (**more**). - VM, disk, and network interface tags are maintained during AWS to AWS and Azure to AWS replication or Disaster Recovery (**more**). - Two Factor Authentication (2FA) to Commvault now supports security keys such as Yubikey from Yubico (**more**) - Added Space Reclamation for Cloud Storage, *including Amazon S3* (**more**). - Added support for Linux-based Access Node to File Index Windows Virtual Machines (**more**). - Containerized Commvault Web Server support (**more**). - Containerized Commvault Command Center support (**more**). *You can download the AWS CAG 2022e Edition for additional information on updates.*

Commvault remains committed to ensuring the **Cloud Architecture Guide** remains current and relevant to currently available public cloud capabilities.

The latest copy of this document is available at **Virtualization White Papers**.

Commvault publishes updates to the Cloud Architecture Guide with each Long-Term Support (LTS) release.

Index

COMMVAULT
Be ready