



GCP Cloud Architecture Guide

Commvault Platform Release 2023E

June 2023

Introduction

The public cloud megatrend is one of the most disruptive and challenging forces impacting customers' applications and infrastructure, requiring new business models and new architecture decisions. This impacts the decisions about solutions for the protection and management of data in public cloud.

Commvault utilizes attributes of public cloud to enable cost effective on-demand use cases for both data protection and data management both to and in public cloud platforms.

Cloud resources, bandwidth, and availability are often localized via massive regional presence to the proximity of on-premises corporate assets and human resources, allowing for an easy on-ramp to the public cloud. The cost model implications of pay-as-you-go do not just extend to only production workloads, but also to the ever-present challenge of providing a flexible, agile, yet capable, recovery solution for your applications and data. Today, many recovery environments have less computing and storage capacity than their production counterparts, resulting in an increased risk of an elongated business service outage.

With the public cloud model, the infrastructure availability and refresh aspect are disrupted by removing the need to maintain a hardware fleet that can meet both your recovery requirements and sustain your service level agreements. Public cloud instances can be rapidly provisioned to meet the needs tied to business requirements,

This dynamic shift allows you to begin costing per recovery event, instead of paying for availability, improving your level of disaster recovery preparedness through the application of flexible, unlimited resources to stage both recovery tests and execute actual recovery events – all without requiring pre-purchased hardware or disrupting production operations. While the recovery use case is the most common foray into a public cloud architecture, many other use cases such as application testing and development, business intelligence and analytics, and production bursting all benefit from the public cloud model.

Commvault® software is designed as an orchestrated, hardware and cloud agnostic, highly modular, distributed solution that conforms with cloud agility, allowing data protection and management solutions that remain flexible through a highly distributed infrastructure built on-top of cloud architecture – public, private or hybrid.

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, of which are 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 GCP	5
Terminology	7
Why Commvault?	8
Commvault Protection of GCP Products	10
Cloud Shared Responsibility	12
Security and Compliance	12
Cloud Sustainability.....	13
Google Cloud Marketplace	14
Getting started with Google Cloud Marketplace.....	14
Remote Access	17
Installation basics.....	18
Pre-packaging Commvault software within a VM template	18
Automating deployment with continuous delivery	18
Getting Started with Google Cloud Storage (GCS) for backup data.....	19
Zero trust model	21
Device Trust.....	22
Protection from threats	26
Hardening your CommServe	26
Design and Best Practices	29
Guides.....	30
Best Practices	58
Patterns.....	64
Anti-Patterns	68
Intelligent Data Use-Cases	69
Data Protection	69
Data Security.....	70
Protect Google Cloud Databases	71
Google Cloud specific workloads.....	73
Protecting and recovering active workloads in Google Cloud	73

Do you need help?	76
Revision History	77
Additional Resources	78
Community Forum.....	78
Documentation	78
Datasheets.....	78
Slack	78

Cloud Architecture Guide for GCP

This guide serves as an architecture guide for solutions architects and Commvault® customers who are building data protection and management solutions utilizing Google Cloud Platform (GCP) and Commvault® software.

It includes public cloud concepts, architectural considerations, and sizing recommendations to support Commvault® software in Google Cloud Platform (GCP). The approach defined in this guide applies to both running Commvault solely in GCP cloud and extending existing on-premises Commvault® functionality into hybrid cloud architectures. The guide covers several common use cases for protecting cloud-native infrastructure services, containerized applications, and scale-out database and storage services. This guide also addresses how to migrate your applications seamlessly and securely to GCP cloud and adopt an on-demand disaster recovery to GCP.

Currently this guide delivers architecture considerations and sizing recommendations for the Google Cloud Platform (GCP). Guides for other public cloud environments are available as well at docs.commvault.com.

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®.

1	Cloud Protection Coverage Get started quickly by leveraging existing GCP Marketplace solutions to protect your GCP services. Understand your data management responsibilities in Cloud.	Why Commvault Google Cloud Marketplace Commvault Protection of GCP Products Cloud Shared Responsibility
2	Cloud Architecture Guidance Review the Google Cloud Architecture Framework 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 Model Google Cloud Architecture Framework Ransomware Protection
3	Designing and optimizing for GCP 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. Adopt automation for hands-off/lights-off management at scale.	Data Management Use-Cases

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

Media Agent refers to the component that is responsible for data handling, both transfer and indexing to facilitate optimized backup and recovery operations. A Media Agent runs the Commvault Media Agent 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 'Media Agent' unless stated otherwise.

Why Commvault?

A common question asked when assessing **data management** solutions for the Enterprise is – what features, functions and capabilities are important? The following provide some high-level capabilities that should be considered when assessing your **data management** needs

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 isn't 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 GCP **projects, regions and zones**. Commvault automatically discovers your cloud workloads by labels, then manages the GCP snapshot lifecycle to meet your business rules and cost objectives.

Cloud-native protection – by default



Commvault orchestrates the creation of cloud-native snapshots (GCP Instances, Google Kubernetes Engine (GKE), Google Cloud VMware Engine (GCVE), Google Regional Persistent disks (standard and solid state drives), Encrypted and Non-Encrypted instance disks. Additionally, Commvault automates snapshot copies within and across regions and accounts using both regional and multi-regional methods.

Cloud mobility without compromise



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

Self-service backup and recovery



Enables **authorized** end users perform recovery using the **Commvault Command Center™** to self-service simple and complex recovery needs without specialist GCP 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 whether its data is protected or whether it represents risk (PII data, insecure data, orphaned data).

Information Management across your entire data estate

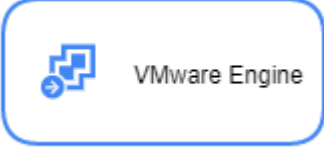


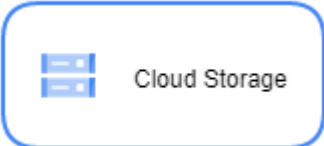





Commvault provides insight into your data through File Storage Optimization (FSO), Data Governance, and eDiscovery and Compliance capabilities. Visualize and identify data risks, inefficiencies across protected and live data sources.

Commvault Protection of GCP Products

Looking to get started quickly? Refer to [Cloud Feature Support for Google](#) for best practices and instructions.

<p>Identity & Access Management</p> 	<ul style="list-style-type: none">Creating a Google Cloud Platform Service AccountService Account Permissions for Google Cloud PlatformCreating and managing service accounts in GCP
<p>Google Cloud Storage Library</p> 	<ul style="list-style-type: none">Configuring Cloud StorageGoogle Cloud StorageGoogle Cloud Storage – Access & Secret KeysGoogle Cloud Storage – OAuth 2.0 (Service Account)
<p>Google Compute</p>  	<ul style="list-style-type: none">Protecting Google Cloud Platform InstancesService Account Permissions for Google Cloud PlatformOperations for Google Cloud Platform InstancesOptions for Conversion to Google Cloud PlatformRestoring Full Instances for Google Cloud PlatformFull Restore Considerations: In Place and Out of Place Restores
<p>Google Kubernetes Engine</p> 	<ul style="list-style-type: none">Protecting Elastic Kubernetes ServiceAuto-protecting containers by label selectorApplication and data migration (cross-cluster, cross-region)

<p>Google Cloud VMware Engine</p> 	<p>Google Cloud VMware Engine Protection</p>
<p>Google Cloud Spanner (relational database)</p> 	<p>Google Cloud Spanner Protection</p> <p>Cloud Spanner</p>
<p>Google Cloud SQL</p> 	<p>Google Cloud SQL for MySQL Protection</p> <p>Google Cloud SQL for PostgreSQL Protection</p>
<p>Google Cloud Storage Protection</p> 	<p>Getting Started with Google Cloud Storage</p> <p>Backing up Google Cloud Storage</p> <p>Restoring Google Cloud Storage</p>
<p>Google Cloud Filestore</p> 	<p>Google Filestore Protection</p>
<p>Commvault Enhanced VM Conversion</p> 	<p>Cross-Hypervisor Restores (VM Conversion)</p> <p>Converting Virtual Machines to Google Cloud Platform</p>
<p>Disaster Recovery</p>	<p>Considerations for Google Cloud Platform Replication</p> <p>Preparing VMs for Replication to Google Cloud Platform</p>

	<p>Recovery Target Options for Google Cloud Platform</p> <p>Periodic Replication Group Options for Google Cloud Platform</p>
<p>App migration</p> 	<p>VM Conversion to Google Cloud</p> <p>Postgres database migration</p> <p>MySQL database migration</p> <p>Migrating Existing Data Using Google Transfer Appliance</p>

Cloud Shared Responsibility

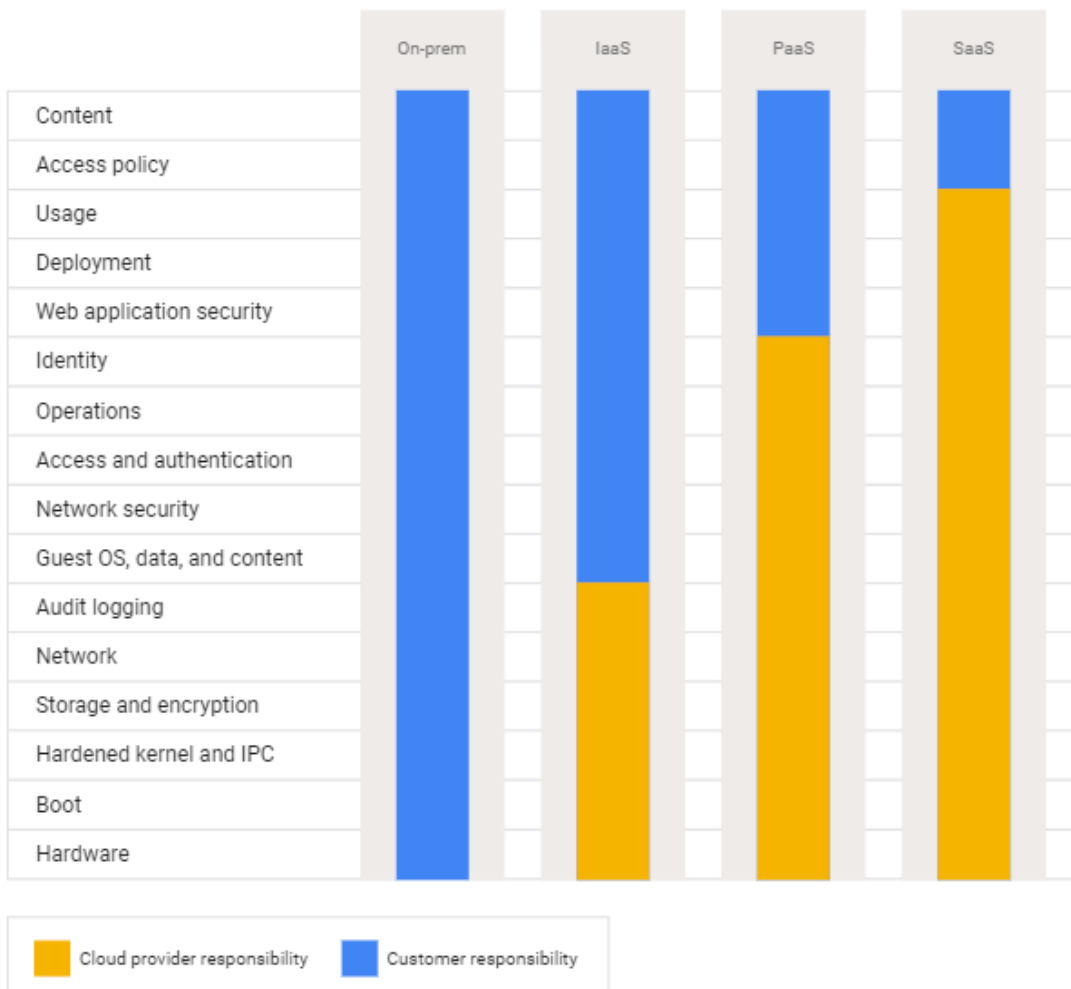
Security and Compliance

Security and Compliance of your data stored and handled in Google Cloud Platform (GCP) is a shared responsibility and shared fate. GCP is responsible for protecting the infrastructure that runs all GCP services. Infrastructure refers to hardware, software, and physical locations represented by region and zone locations. This is often referred to as security “of the cloud”.

GCP takes its security responsibility seriously with external validation to the leading industry regulations as found at [Compliance resource center](#).

Customers are responsible for protecting their data stored in GCP services by configuring and hardening the platform, application, and identity and access management controls that govern access to GCP 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 GCP shared responsibility model, [Shared responsibilities and shared fate on Google Cloud](#).



This document will help secure the Commvault and GCP infrastructure and provide **Recovery Readiness** for your GCP-hosted applications and **data**. Applying Commvault to the GCP data landscape lends recoverability for the following GCP services both within and across the projects and regions.

Cloud Sustainability

Google Cloud considers environmental sustainability another shared responsibility with its customers and partners. Google considers itself the cleanest cloud in the industry: **Cloud sustainability**.



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

- Minimizing hardware required to perform multi-project and multi-region data protection.
- Powering down Media Agents and Access Nodes running on GCP Compute instances when they are not actively required, with **MediaAgent Power Management**.

Google Cloud Marketplace

Getting started with Google Cloud Marketplace

GCP Marketplace modernizes your software acquisition, testing and governance across your organization. Commvault currently publishes two listings on the Google Cloud Platform

 COMMVULT	Commvault Backup & Recovery BYOL Commvault - Virtual machines <p>Commvault's industry-leading intelligent data management platform provides seamless backup, recovery, disaster recovery and data insight for cloud-based workloads. Applications can be recovered, migrated, and reported upon for recovery readiness, compliance and data insight initiatives.</p>
 COMMVULT	Commvault Backup & Recovery Commvault - SaaS & APIs <p>Commvault's industry-leading intelligent data management platform provides seamless backup, recovery, disaster recovery, and data insight for cloud-based workloads including VMs Databases, PaaS, Files, etc. Applications can be tested, recovered, and reported upon for recovery readiness, compliance, and data insight initiatives</p>

Marketplace, a **Software as a Service (SaaS) listing** and a **Bring Your Own License (BYOL) listing**. This gives a GCP Marketplace customer two methods for purchasing the Commvault software. The SaaS listing is setup for the customer to request a private offer for software and services from Commvault while the BYOL listing is a virtual machine that can be deployed as a CommServe® server and Access Node (MediaAgent + Virtual Server Agent) within their GCP environment and **GCP project** of their choice.

Search for “Commvault” on GCP Marketplace to obtain an image for your environment or purchase Commvault software through a private offer.

- **Commvault Backup & Recovery BYOL** deploys Commvault Backup & Recovery on a single GCP Compute Microsoft Windows Server 2019 instance within an existing GCP project with all dependencies. Solution provides a FREE 60-day license for testing, trials, and proof of concept (PoC) initiatives. You may purchase a license after 60-day trial expires.

- Commvault Backup & Recovery (SaaS)** can be used to purchase a subscription of Front End Terabyte (FET) or Virtual Machine (VM) licenses through the GCP Marketplace. Once a subscription is chosen, the customer will subscribe to the listing. Asking for a private offer allows Commvault to customize the SKUs needed for the customer so a purchase through the Marketplace is not limited to the options displayed.

	Data Protection Starter Bundle FET USD 7,483.00/mo Subscription Period* 1 year - USD 7,483.00/mo	Data Protection Starter Bundle VMs USD 735.00/mo Subscription Period* 1 year - USD 735.00/mo	Premium Data Protection USD 100,000.00/mo Subscription Period* 1 year - USD 100,000.00/mo
Complete Data Protection per Front End Terabyte Capacity	✓	no	
Complete Data Protection per Virtual Machines(10 pack)	no	✓	
Cloud Backup and Replication for Applications, VMs, DBs and Files	✓	✓	
Recovery Automation and Orchestration with Configurable RPO	✓	✓	
1-Click Failover, Failback and Recovery	✓	✓	
Disaster Recovery Readiness Validation and Reporting	✓	✓	
Automated VM and Data Migration with Native GCP Integration	✓	✓	
Ransomware Protection with Anomaly Detection and Reporting	✓	✓	
VM, Database and File Clone, Mount and Instant recovery	✓	✓	

Deploying Commvault Backup & Recovery BYOL

Commvault is available for download in the [Google Cloud Marketplace](#) via the [Commvault Backup & Recovery BYOL](#) listing. With Commvault Bring Your Own License (BYOL) you can take control of your cloud data, using a single instance for the CommServe®, Access Node, and Virtual Server Agent (VSA). Other instances can be deployed as needed.



Commvault Backup & Recovery BYOL

Commvault - Virtual machines

Commvault's industry-leading intelligent data management platform provides seamless backup, recovery, disaster recovery, and data insight for cloud-based workloads including VMs Databases, PaaS, Files, etc. Applications can be tested, recovered, and reported upon for recovery readiness, compliance, and data insight initiatives.

The Commvault Backup & Recovery BYOL listing allows the customer to choose and configure the instance by setting the Deployment name, the zone, the Machine Series, the Machine type, the boot disk type, the boot disk size, and the network interface.

Windows Server 2019 Datacenter Edition Usage Fee	USD 134.32/mo
SHOW MORE	
CommvaultVMDeploymentSolution usage fee (BYOL)	Varies
Google does not collect this license fee.	
Infrastructure fee	
VM instance: 4 vCPUs + 16 GB memory (n2-standard-4)	USD 159.69/mo
Standard Persistent Disk: 200GB	USD 8.80/mo
Sustained use discount	- USD 47.91/mo
Estimated monthly total	USD 254.90/mo
	+ BYOL license fee

Pricing for the Commvault Backup & Recovery BYOL listing is dependent on the series type, CPU, memory, disk type and size chosen. No cost is associated with the Commvault software 60 day trial.

Configure the options for your instance, accept the Terms of Service and click DEPLOY.

This will launch the Deployment Manager and create the instance within the current project. The template will show the estimated monthly cost based on the Machine Type and configuration

chosen. The template will default to the n2-standard-8 series machine type. The instance is sized for day zero work and can be utilized as an all-in-one instance. As the environment grows, the instance can be resized and access nodes added as needed.

Once deployed, the instance will be built as a Microsoft Windows 2019 Virtual Machine. The Commvault installer will begin at first login. This will install the Commvault CommServe® with VSA and Access Node configured.

New Commvault Backup & Recovery BYOL deployment

Deployment name *
commvaultpackage-1

Zone
us-east4-c

Machine type
Machine family
GENERAL-PURPOSE COMPUTE-OPTIMIZED MEMORY-OPTIMIZED

Machine types for common workloads, optimized for cost and flexibility

Series
N2
Powered by Intel Cascade Lake and Ice Lake CPU platforms

Machine type
n2-standard-4 (4 vCPU, 16 GB memory)

	vCPU	Memory
	4	16 GB

Boot Disk
Boot disk type *
Standard Persistent Disk

Boot disk size in GB *
200

Networking
Network interfaces
default default (10.150.0.0/20)

ADD NETWORK INTERFACE

I accept the [GCP Marketplace Terms of Service](#) and [Commvault Terms of Service](#).

DEPLOY

Commvault Backup & Recovery BYOL overview
Product provided by Commvault

Launching a BYOL product

Commvault Backup & Recovery BYOL is a BYOL (Bring Your Own License) product. Marketplace will deploy this product, but you are responsible for purchasing and managing the license directly from the provider

Windows Server 2019 Datacenter Edition Usage Fee	USD 134.32/mo
CommvaultVMDeploymentSolution usage fee (BYOL) Google does not collect this license fee.	Varies
Infrastructure fee	
VM instance: 4 vCPUs + 16 GB memory (n2-standard-4)	USD 159.69/mo
Standard Persistent Disk: 200GB	USD 8.80/mo
Sustained use discount	- USD 47.91/mo
Estimated monthly total	USD 254.90/mo + BYOL license fee

Price estimates based on 30-day, 24hrs per day usage of the listed resources in the selected region. The Estimated Monthly Infrastructure Fee calculation may not reflect all Google Cloud Platform IaaS resources actually created or consumed by this product (or the fees charged for such consumption). Commvault may be able to provide a more accurate estimate of monthly GCP IaaS consumption.

Software
Operating System Windows server(2019)

Terms of Service
By deploying the software or accessing the service you are agreeing to comply with the [Commvault terms of service](#), [GCP Marketplace terms of service](#) and the terms of applicable open source software licenses bundled with the software or service. Please review these terms and licenses carefully for details about any obligations you may have related to the software or service. To the limited extent an open source software license related to the software or service expressly supersedes the GCP Marketplace Terms of Service, that open source software license governs your use of that software or service.

By using this product, you understand that certain account and usage information may be shared with Commvault for the purposes of financial accounting, sales attribution, performance analysis, and support.

Google is providing this software or service "as-is" and any support for this software or service will be provided by Commvault under their terms of service.

Remote Access

As with all IaaS offerings, remote access to Virtual Machine instances can be achieved with RDP for Windows and SSH for Linux instances and Commvault® module deployment can be achieved with the current procedures listed in [Commvault Documentation](#)

Installation basics

The following links cover the steps when installing the CommServe® in the cloud. This is only needed when the primary CommServe® will be running on the hosted cloud VM or used for DR recovery. Multiple modules can be deployed in a single installation pass to streamline deployment.

Installation Overview

Installing the CommServe®

Installing the MediaAgent

Installing the Virtual Server Agent (GCP)

CommServe® Disaster Recovery solution comparison

Learn more about CommServe® DR Solution comparisons for building a standby DR CommServe® in the cloud, or simply restoring on-demand (DR backup restore), **Configuration of Disaster Recovery (DR) Backups**.

Pre-packaging Commvault software within a VM template

For environments where deployment time is reduced by preparing software and configuration within VM templates, the Commvault in-guest iDataAgents can also be deployed in Decoupled mode. This means that the in-guest iDataAgent is deployed within the Google Compute instance but will only be activated upon registration with the CommServe®.

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

Installing the Custom Package on Windows

Installing the Custom Package on Linux

Automating deployment with continuous delivery

For environments using Continuous Delivery toolsets such as **Terraform**, Puppet, Chef or Ansible, Commvault® supports deployment methods that allow administrators to control agent deployment and configuration to provide an automated deploy-and-protect outcome for applications and servers.

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

- **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 options available for each in-guest iDataAgent:

- [REST API – Overview](#)
- [Command Line – Overview](#)

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.

Getting Started with Google Cloud Storage (GCS) for backup data

Google Cloud Storage is a RESTful online file storage web service for storing and accessing data on Google Cloud Platform infrastructure. The service combines the performance and scalability of Google's cloud with advanced security and sharing capabilities.

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**). Google Cloud Storage can then be configured as a secondary or secondary copy location, providing offsite protection while benefiting from scalability and durability. The secondary copy is replicated periodically as an encrypted network-optimized **Deduplicated Accelerated Streaming Hash Copy**.

The link below lists all the supported direct cloud storage targets.

- [Supported Cloud Storage](#)

The link below covers cloud storage target setup and management.

- [Cloud Storage - Overview](#)

① Note

Use the following GCP permissions to create a GCS bucket as a Commvault library.

Bucket Target permissions:

- storage.buckets.create
- storage.buckets.delete
- storage.buckets.get
- storage.buckets.list
- storage.buckets.update
- storage.objects.create
- storage.objects.delete

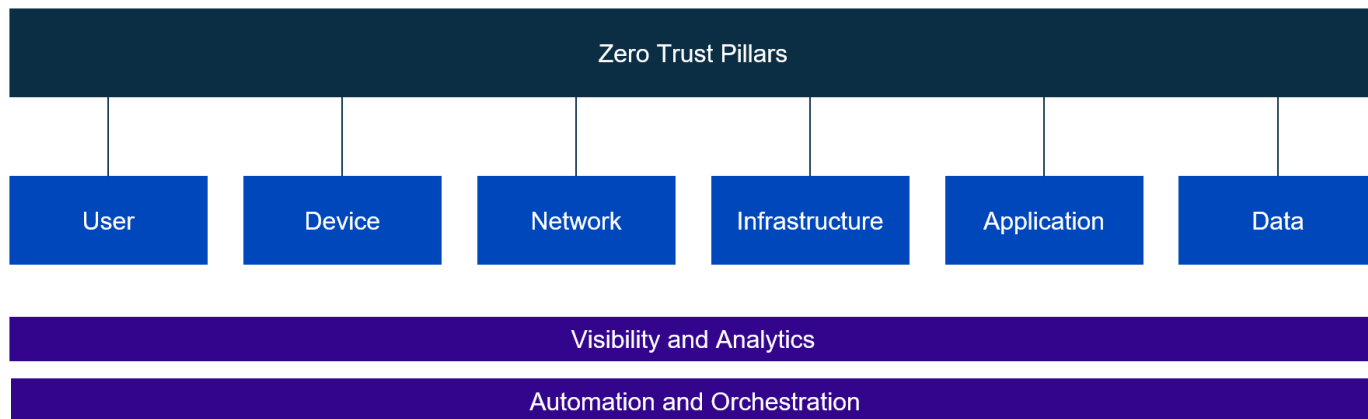
Zero trust model

Zero trust (ZT) cybersecurity models are built on the following base assumption:

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

NIST Special Publication 800-207 – Zero Trust Architecture

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). Identity and Access Management (IAM) lets administrators authorize who can take action on specific resources, giving you full control and visibility to manage Google Cloud resources centrally.

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 devices.

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, etc..) that the

user is permitted to act upon. All sessions are authenticated with a configurable timeout which defaults to 30 minutes.

Commvault has multiple anomaly detection and machine-learning algorithms that track and automatically respond to events that are indicative of a threat (i.e., modification of honey pots).

Commvault software can be implemented as a secure multi-tenancy model where tenants are created as securely separated 'companies'. Commvault considers tenant isolation and granular role-based access control a specialized form of macro-segmentation between companies. Individual tenant user rights and access controls provides micro-segmentation within the tenant.

Device Trust

Commvault performs active vulnerability management and reporting, for issues that impact Commvault products at **Security Vulnerability and Reporting**. Vulnerabilities may require an update to Commvault software and/or Operating system and third-party software on client 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 AES-256 cipher.

Commvault software assists in device management by providing centralized device configuration, reporting and insights. Additionally, Commvault software automatically downloads and deploys Commvault software patches and updates per the defined policy on all client and core data management devices/appliances. Commvault can also manage the deployment of Microsoft 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 user-configurable 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-140-2 cryptographic module validation program compliant**.

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 can write to encrypted-only object storage buckets, with cloud-provider or customer-managed keys (CMKs) specified.

Commvault in essence 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 HTTP and HTTPS 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 (WAF) like **GCP Cloud Armor** 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 employs a least privilege approach to cloud data management requesting only the minimum permissions required to protect a GCP service. Commvault role-based access controls (RBAC) are then overlaid to further restrict the individual user and/or user group.

Commvault software is accessible via Command Center 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-140-2 compliant cryptographic library 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 **Google Cloud Key Management**. Customers may stay in control of their keys with a cloud-based or on-premises KMIP-compliant hardware security module (HSM) like **GCP Cloud HSM**.

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 and managed encryption keys.

Certifications and Compliance

Commvault is responsible for the integrity of data persistent within Commvault-controlled cloud, HyperScale™, disk, and tape storage locations. Commvault can perform 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.

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 webhook, Syslog, SDK, or custom action.

Commvault provides continuous diagnostics and mitigation capability via centralized reporting and alerts. This helps with infrastructure that are behind in software updates or configuration upgrades.

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 Job Manager also performs policy enforcement from the centralized CommServe® instance which monitors, initiates, and terminates communication as required to complete data management activities.

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%2020210610%20\(2\).pdf](https://www.gsa.gov/cdnstatic/Zero%20Trust%20Architecture%20Buyers%20Guide%20v11%2020210610%20(2).pdf)
- 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>

Protection from threats

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 data protection plan.

Hardening your CommServe

Best practice is to harden your CommServe® instance by restricted access to authorized users, authorized hosts, and authorized ports. See [Securing the CommServe®](#)

Access to the centralized configuration MS SQL Server database should also be limited, see [Securing the CommServe Database](#).

Account segregation

Utilize multiple GCP accounts to segregate users, departments, and backup copies. Utilizing GCP [Separation of duties](#) ensures that one individual does not have all necessary permissions to complete a malicious action. Create [separate IAM roles](#) to establish Separation of duties. Commvault supports the protection and storage of data across multiple accounts, projects, regions, and 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 one copy 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**.
 - Commvault **Cloud MediaAgent Power Management** can be used to automate the power-down/power-up of the MediaAgent.
 - (optional) Consider utilizing **GCP Cloud Scheduler** to schedule the powerup/power-down of air gapped MediaAgents.
 - A **powered-down MediaAgent** is considered ‘air gapped’ from your primary operational recovery infrastructure.
- Always utilize **Google Cloud Storage** with dedicated credentials and authorized MediaAgents

- The use of *object-based storage* without end-user access is considered another **air gap** for your data.
- Traditional storage mediums (disk-based storage) can be directly accessed by malware and infected if security credentials are discovered/breached.

Detecting Ransomware activity

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**

See **Ransomware protection** and **Offline Backup Copies** for additional information.

Partial Recovery

Be sure to consider the impact of a ransomware event on your organization. Ensure a copy of your data is readily accessible, and within a storage tier aligned to your recovery time objectives. Commvault optimizes recovery events by allowing multiple MediaAgent / Access Nodes to read your backup data in parallel. This infrastructure can be **temporary**, existing only for the period of the recovery event.

Additional resources

- **GCP Best practices to protect against ransomware**
- **GCP security transformation**
- **NIST Cybersecurity Framework & Google Cloud**

Google Cloud Immutable Storage with Bucket Lock

Commvault can create immutable data by activating Google Cloud Storage Bucket 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).

Immutable storage, also known as WORM (Write Once, Read Many) storage, enables users to

① Note

WORM should ideally be enabled on the backup library before writing any backup data, to ensure all data is immutable from initial creation.

store business-critical data with the requirement that it cannot be modified or deleted based on a user defined retention period. Designing a solution with offsite copies to protect against ransomware and cyber threats is imperative. Commvault provides the ability to utilize cloud immutable storage with Google Cloud storage for enhanced data security.

Important considerations for *immutable cloud storage* include:

- Data within the Cloud storage location is set to write-once-read-many (WORM).
- Data retention for the copy should be set to half of the total desired retention age
- Deduplication is supported and recommended.
- For deduplication enabled storage pools, the seal frequency of the deduplication database (DDB) is set to the same number of days as retention set in the storage policy copies of selected storage pool, with a maximum of 365 days.

See the following information for more details:

- [Configuring WORM Storage Mode on Cloud Storage](#)
- Setting up [Google Cloud Bucket Lock](#)

Effects of DDB sealing

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**, or double the space, of the backup content being re-sent to the Cloud Library. This is intentional and provides multiple, segregated data copies to protect from corruption or other unforeseen data access issues.

See [Sealing the Deduplication Database](#).

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.

As you design and build your 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 GCP and edge-based workload protection.

Right-sizing Over Forecasting

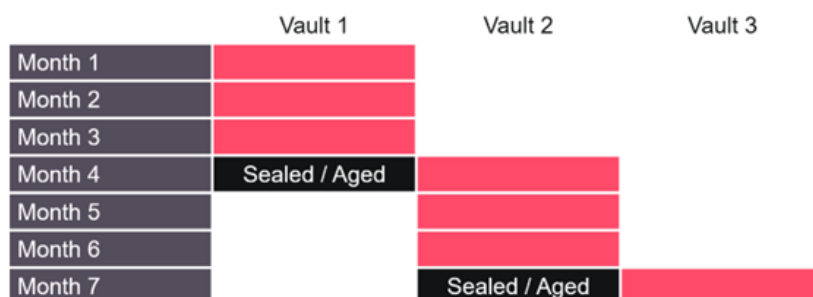
Commvault recommends that traditional infrastructure t-shirt sizing and forward-looking capacity provision practices are discarded when designing for Commvault cloud workloads. Always select the smallest recommended compute instance and scale to meet your business service-levels.

Apply **machine type recommendations** when resources are being exhausted, and review your achieved objective levels before increasing resource sizing.

Match protection to business impact

Use a *business-value data lifecycle* approach for backups that stores backups using GCP snapshots, then copies in Google Cloud Storage Standard class, and finally in Google Cloud Storage Nearline/Coldline/Archive classes. Applications rarely require a single storage technology across their lifetime.

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 low cost of the GCS Archive tier means the storage cost is negligible.



Centrally managed GCP recovery points

Commvault provides a **single unified data management platform** to perform GCP 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 GCP resources.

Optimize At Rest

Ensure that all copies of data are stored in a least-cost optimized format for the organization's SLAs to reduce storage fees and the overall sustainability of your data management landscape. As your recovery needs change and no longer require rapid-recovery GCP snapshots, take service-independent backup copies to Google Cloud Storage. Use Commvault deduplication, compression to optimize your Google Cloud Storage 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 GCP-native snapshots (where supported) and Commvault deduplication enhanced replication (DASH Copy) 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 GCP-native snapshots and Commvault independently encrypted Google Cloud Storage 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 Google Cloud Operations to automate operations where required.

Guides

The following section guides planning and implementing Commvault data management in GCP with a focus on best practice guidance and technical limitations to consider during design.

Backup and recovery approaches on GCP and beyond

This guide describes a high-level process for assessing and implementing a consistent data protection and data management approach for your GCP and hybrid workloads.

Hybrid protection

The availability of compute instances, network, and storage resources has changed how business services are delivered, and how data protection services can be designed and implemented. GCP and Commvault provide several services and technologies to protect your workloads running in the GCP Region, **Google Anthos**, 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 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 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.

When determining an appropriate PTO, consider the costs and complexity of implementing multi-region failover and fallback are often higher than in-region 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

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. If the organization permits, application owners can restore directly back to their existing GCP and edge-based resources or opt to recreate and replace malfunctioning resources.

Optimize for scalable cloud resources

Leverage the scalable nature of GCP compute instances 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:

- **Zero cloud infrastructure** by performing the backup activity and cloud orchestration from on-premises (Commvault does not require infrastructure in the GCP region to write backups to Google Cloud Storage) or reading data directly from cloud snapshot copies or streamed backups (see [Restores for Google Cloud Platform](#)) for procedures for both in-place and out-of-place restores.
- **Power Managed infrastructure** that is powered on only when a data management activity is performed (see [Cloud MediaAgent Power Management](#)).

The following table summarizes Commvault backup infrastructure options when deployed in the Google Cloud (multi-region, region, and zone).

Commvault Component	Long-Running	Power-managed
CommServe® instance	√	
MediaAgents		√
Access Nodes for Backup		√
Access Nodes for Restore		√

Protection approaches

The introduction of Google Cloud to your application landscape has also introduced enterprise-grade protection technologies often available only to the largest organization. Protection in GCP 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.

GCP service/product	API	Snapshot	Streaming	Comments
Cloud Spanner			√	Export backup is used for streaming copy
Persistent Disk	√	√	√	Snapshots can be region or multi-region
Compute Instance	√	√	√	Snapshots can be region or multi-region
Filestore			√	Protection via Linux File System Agent
Anthos	√		√	Cloud-native protection via Kubernetes API server (kube-apiserver).
Big Query				Coming in future release
Cloud PostgreSQL, MySQL			√	Export Backup is used for streaming copy
Google Cloud Storage	√		√	Direct API request for objects via CloudApps agent

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 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 Google Persistent disks, google Cloud Storage (GCS), Google Filestore, and GCP 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 GCP Region out to on-premises or another cloud, consider the impact of network transfer times on your recovery times. When utilizing edge-based optimized infrastructure like Google Anthos, 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 be 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

Google Cloud Storage provides 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 can store your data in Google Cloud Storage in an optimized, deduplicated, and compressed format which reduces backup storage and replication costs. This approach creates multiple logical containers within your Google Cloud Storage and prevents the use of Google Cloud 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.

Programmatic Data Management

Commvault® software provides a robust **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. Since format conversions are handled by the GCP 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. 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 **Google Cloud Marketplace**.

Selecting compatible Google Cloud Compute Engine instance types

Availability of Google Cloud Compute Engine instance types varies by region.

Commvault does not limit deployment to a specific instance type.

Commvault publishes GCP 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 GCP Marketplace to only instances that meet Commvault [minimum requirements](#).

Commvault recommends the use of current generation, Compute optimized (C Class family), Memory optimized (M Class family), and General Purpose (burstable) (E2 and N Class families) for Commvault components.

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.

Use of compute-optimized burstable instances

Commvault **does not recommend** the use of **burstable** instance types (E2 class 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 burstable families 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 any 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.

Usage of GCP VM instances with networking bandwidth exceeding 12.5Gbps will result in significant underutilization of provisioned resources leading to waste, higher Compute instance runtime costs, and reduced sustainability of the solution.

Use of Ephemeral disks (local SSD)

Avoid using ephemeral **local SSDs** for Commvault GCP instances as these local volumes will not retain their data if the GCP instance is powered off or terminated, which makes them unsuitable choices for the Index Cache or DDB.

Where to place your instances in a multi-region deployment

CommServe placement

Place your primary CommServe instance in the GCP Region where the majority 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.

Access Node placement

Access Nodes perform agent-less protection for GCP Compute instances, Or PaaS databases such as Spanner, Big Query, PostgreSQL, MySQL, and Google Filestore file-level data. No Commvault software agents are required inside your GCP Compute 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 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.

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 Google Cloud Storage, 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. Note that the index data and DDB data will still go to the respective Media Agent.

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-zone **data transfer fees** are incurred in normal operations as Commvault distributes I/O for availability in multi-zone deployments.

Multi-region placement

Commvault does not support high-availability *MediaAgent Grids* or Access Node groups that span GCP 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.

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 (Compute Instance) to perform rapid backup and recovery operations.

- Use Commvault HotAdd to create and restore service-independent backup copies of GCP persistent disks, when the per region, per-account GCP 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 Access Nodes to support more parallel read streams from Google Cloud Storage.
- Increase the number of **data readers** and **device streams** to improve concurrency to Google Cloud Storage.

Right-sizing your data management resources

Commvault recommends using automated **GCP Machine type recommendations** to review when resources require right-sizing to larger or smaller instance sizes.

us-east4-	Mar 2, 2022,	e2-standard-8	
c	1:08:43 PM UTC-		Save \$10 / mo
	05:00		

CommServe scaling

Using the GCP Machine type recommendations, 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 **Designing resilient systems**).
- 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).

Access node scaling

For non-GCP Compute instance protection and recovery, such as on-prem or GCVE virtual machines, 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 [Increase the size of a persistent disk](#) to resize a volume).

Note: This is the Commvault recommendation when building a CommServe® on a GCP Compute instance.

GCP Commvault Backup & Recovery - Drive Layout				
Drive Letter [Label]	Initial Capacity (GiB)	File-system Block size	Type / Performance	Purpose
C:\ [WINOS]	35 (default)	4096 (default)	Balanced PD / 3000 IOPS at 4KB / 125 Mbps	Microsoft Operating System.
E:\ [CVLT] E:\SoftwareCache	60	4096 (default)	Balanced PD / 3000 IOPS at 4KB / 125 Mbps	Commvault binaries, MS SQL Server binaries, Commvault log files, & Commvault software cache
F:\ [MSSQL]	40	65536 (64K)	Balanced PD / 3000 IOPS at 64KB / 125 Mbps	MS SQL database files + tempdb
G:\ [TLOGS]	10	65536 (64K)	Balanced PD / 3000 IOPS at 64KB / 125 Mbps	MS SQL transaction logs
H:\ [DDB1]	50	32768 (32K)	Balanced PD / 3000 IOPS at 32KB / 125 Mbps	Commvault Deduplication Database (DDB)
I:\ [INDEXC] I:\IndexCache	50	32768 (32K) (default)	Balanced PD / 3000 IOPS at 32KB / 125 Mbps	Commvault Index Cache

J:\ [JOBS] J:\JobResults J:\DR	50	32768 (32K)	Balanced PD / 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 be formatted as **NTFS** with the specified block size.

NOTE: Stated IOPS are a *day one* configuration and can be increased as demand requires by reviewing **GCP disk performance**.

Reference – Cloud Access Node Storage Layout

The following is the default storage layout for an arm64 or x86_64 Cloud Access Node that can be manually deployed referencing [Deploying a Linux Machine as an Access Node](#). These nodes may be used as MediaAgents, Access Nodes, or perform both roles.

GCP Commvault Cloud Access Node - Drive Layout				
Drive Letter [Label]	Initial Capacity (GiB)	File-system Block size	Type / Performance	Purpose
nvme0n1	10	4096	Persistent Disk / 3000 IOPS at 4KB / 125 Mbps	Linux Operating System.
nvme1n1 [vg_commvault]				
vg_commvault-lv1 /opt/commvault	10	4096	Persistent Disk / 3000 IOPS at 4KB / 125 Mbps	Commvault binaries, Commvault software cache
vg_commvault-lv2 /var/log/commvault	4.9	4096	Persistent Disk / 3000 IOPS at 64KB / 125 Mbps	Commvault log files
vg_commvault-lv3 /mnt/commvault_jobresults	40	4096	Persistent Disk / 3000 IOPS at 32KB / 125 Mbps	Commvault Job Results, 3DFS Cache, temporary upgrade location.
vg_commvault-lv4 /opt/commvault_indexcache	25	4096	Persistent Disk / 3000 IOPS at 32KB / 125 Mbps	Commvault Index Cache
nvme2n1 [vg_commvault2]				
vg_commvault2-lv_ddb /mnt/commvault_ddb	20	4096	Persistent Disk / 3000 IOPS at 32KB / 125 Mbps	Commvault Deduplication Database (DDB)

Understanding Google Cloud Storage (GCS) classes

Google Cloud Storage provides a broad selection of storage classes with differing performance and cost characteristics.

Google Cloud Storage (GCS) can be configured in three location types:

- **Multi-region:** Highest availability across largest area.
- **Dual-region:** High availability and low latency across 2 regions.
- **Region:** Lowest Latency with a single region.

Google Cloud Storage provides four basic classes with an option for Autoclass:

- **Autoclass:** Automatically transitions each object to hotter or colder storage based on object-level activity, to optimize for cost and latency. Recommended if usage frequency may be unpredictable. Can be changed to a default class at any time.
- **Standard:** Best for short-term storage and frequently accessed data
- **Nearline:** Best for backups and data accessed less than once a month
- **Coldline:** Best for disaster recovery and data accessed less than once a quarter
- **Archive:** Best for long-term digital preservation of data accessed less than once a year

Choosing the correct Google Cloud storage tier

Choosing the correct Google Cloud Storage tier for backup data is crucial when considering performance, cost, and accessibility. If, after reading through the information below, you are still unsure where to store backup data, we recommend **Google Cloud Nearline storage** for best cost versus performance and should be used for most primary copies. Keep in mind this is just a recommendation and there are always exceptions.

Each tier of storage has specific characteristics which need to be fully understood. However, before choosing the storage tier, you first need to determine what the business requirements demand for accessibility and retention.

For example:

Data Storage Requirements	Hot (Standard)	Cool (Nearline/Coldline)	Cold (Archive)
Data retention (Days, Months, Years)	Days	Months	Years

Accessibility (Frequent, Occasional, Rare)	Frequent	Occasional	Rare
Time to recover (Minutes, Hours, Days)	Minutes & Hours	Minutes & Hours	Hours & Days

Additional factors:

- Location of data to be protected (edge-based or in the cloud)
- Location of MediaAgent (edge-based or in the cloud)
- Location of backup storage (edge-based or in the cloud)
- Internet bandwidth (Public internet or GCP Cloud interconnect)

Consideration of the above factors can help determine if the same storage tier for all data being protected is sufficient or may be different depending on the type of data (databases, healthcare, financial, filesystems, etc). Understanding these factors will determine the storage tier best suited for the data type.

- **Google Cloud Storage (Standard)** is designed for fast and easy access to data, but at the highest cost per GB. Data stored in Standard storage class is recommended for short term retention (primary backup copy).
- **Google Cloud Storage (Nearline/Coldline)** is similar in performance as the Standard tier storage and is designed for fast and easy access to data. However, it's cheaper per GB. The reason for the price break is because Google Cloud is expecting the data to remain in Nearline/Coldline storage for a minimum of 30 days. If the data is removed prior to 30 days, the storage cost is prorated. Data stored in the Nearline/Coldline tier is recommended for short and long term retention that requires access for restores on frequent or occasional use (primary backup copy and long term retention).
- **Google Cloud Storage (Archive)** is the lowest price storage tier per GB. The purpose for this tier is for a very specific use case. Data placed in Archive storage is expected to remain for a minimum of 365 days. Plan tiering the data in advance depending on the retrieval requirements. Retrieval of data from archive storage is more expensive than the other tiers.

See [Google Cloud Storage pricing model](#) for the latest cost of each tier for each region.

① Note

Commvault Combined Storage Tiering is not used for Google Cloud Storage. Combined Storage Tiering is for optimally utilizing archive tiers where data is not accessible unless an explicit recall is needed. This is not required for GCS.

Google Cloud storage redundancy

Google Cloud Platform provides redundancy of data stored in a storage account to ensure durability and high availability. Data that is geo-redundant is stored in at least two separate geographic places. Objects stored in **multi-regions and dual-regions** are geo-redundant, regardless of their storage class. These redundancy levels are transparent to Commvault and can be used if required.

Auto-discovery and protection at scale

Commvault recommends using your **GCP Labels** strategy to inform Commvault of the workloads that need to be protected and what data classification or business value applies to each workload. Commvault uses auto-discovery rules to dynamically discover new workloads at backup runtime.

Additionally, Commvault can **auto-detect** applications running inside your GCP Compute Instance 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.

Using Commvault deduplication for reduced storage costs

Commvault recommends using deduplication and compression for all backup and archive data stored in Google Cloud Storage.

Deduplication is about reducing the amount of data stored by removing duplicate data items from the data store. Commvault uses MediaAgents and optionally 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 Google Cloud Storage costs and data transfer costs when replicating backups between Regions.

Commvault recommends utilizing deduplication to reduce the cost of data stored and transferred as a general rule. While there are specific data types that do not deduplicate, such as Oracle and SQL transaction logs or encrypted data, these are often a small percentage of the overall data footprint within the business.

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.

Replicating Cloud Storage

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.

Sizing Guidelines

Commvault software provides the ability to build a distributed data management platform servicing a small single region single-account environment, to a very large multi-project, 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 Google Compute 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.

The following section provides the Commvault-recommended GCP Compute Instances for both deployment models. As your data management needs expand, you should refer back 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

On day 1, you should deploy a **seed architecture** that deploys all components on a single Google Compute instance, located within a single region and 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 zone 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 Google Cloud Storage 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 Google Cloud Storage Archive tier.
- **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.

📌 Important

The sizing recommendations in this document have been tailored specifically for deployment within Google Cloud Platform (GCP). 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 **GCP Machine Type Recommendations** to receive recommendations on the optimal GCP

- **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 located in an edge location outside the Region.

Commvault recommends the following Google Cloud 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 [minimum requirements](#).

💡 Pro-Tip

Commvault recommends selecting the instance with the **least cost** to get started and using the **GCP Machine Type Recommendations** to guide when to right-size your instance type.

Instances are listed in priority order. Options for x86_64 (Intel, AMD) and arm64 architecture are provided, where supported.

	Commvault Recommended Google Compute Instance Types	
	Least Cost	Most Performant
Seed all-in-one CommServe® instance	n2-standard-8 (8/32/intel)	n2d-standard-8 (8/32/amd)
Seed MediaAgent (snapshot protection)	n2d-standard-2	n2-standard-2
Seed MediaAgent (snapshot + streaming protection)	c2-standard-4	c2d-standard-4
Scale-out all-in-one CommServe® instance	n2-standard-8 – 16 (snapshot) n2-highcpu-16 – 32 – (streamed) n2-standard-8 – 32 (mixed workload)	n2d-standard-8 – 16 (snapshot) n2d-highcpu-16 – 32 – (streamed) n2d-standard-8 – 32 (mixed workload)
Scale-out MediaAgent Grids (snapshot and streamed protection) ⁴	n2-highcpu-4 – 32 (snapshot and stream)	N2d-highcpu-4 – 32 (snapshot and stream)

Seed – Commvault CommServe® Instance

The following is the recommended day-one minimum configuration for getting started with Commvault software.

Pro-Tip

Get started in GCP Marketplace with **Commvault Backup & Recovery BYOL** product.

GCP Quick Start Specifications – Seed Commvault CommServe® Instance

Configuration	An all-in-one deployment with Commvault CommServe®, MediaAgent, Access Nodes, and Cloud Controller in one Google Compute instance. Hosts deduplication cloud libraries for backups	
Instance type	Least cost n2-standard-8	Best price/performance n2d-standard-8
Operating systems	Red Hat Enterprise Linux 8 x86_64 (recommended) Microsoft Windows 2019 x86_64	
In GCP Marketplace	Yes (Microsoft Windows 2019) x86_64	
Required Persistent Disk Storage	All volumes are zonal or regional persistent disks storage with initial baseline IOPS configuration. <ul style="list-style-type: none">- 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 GCP services in the region, zones, and edge-based hybrid locations. Protects Google Compute Instances, Google Cloud Storage (GCS), Google Kubernetes Engine (GKE), Google Spanner, PostgresSQL, MySQL, Google Filestore,	

Protects	Avg. throughput 504GB/hr. (Backup)/ 1424GB/hr.(Restore) observed on an n2d-standard-8 equivalent instance*** Plan for up to 100TiB of written streamed backup data, 5% streamed per day.
----------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

*** 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 minimum 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).

GCP 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	Least cost n2d-standard-2	Best price/performance n2d-standard-2
Operating systems	Red Hat Enterprise Linux 8 x86_64 (recommended) Microsoft Windows 2019 x86_64	
In GCP Marketplace	No	

<p>Required Persistent Disk Storage</p>	<p>All volumes are zonal or regional persistent disks storage with initial baseline IOPS configuration.</p> <ul style="list-style-type: none"> - 80GiB Commvault binaries / Logs / Job Results / Index Cache, 3000 IOPS @ 4K - 25GiB Commvault deduplication database, 3000 IOPS @ 4K <p>Your backup metadata is stored in scalable, durable, and secure Google Cloud storage.</p>
<p>Use for</p>	<p>Performing GCP snapshot creation, sharing, and copying between regions and accounts for Google Compute Instances</p> <p>Writing backup activity for snapshot-only jobs to a hosted non-deduplicated cloud library.</p> <p>(Optional) Sending backup activity for snapshot-only jobs to a remote cloud library.</p>
<p>Protects</p>	<p>**Protect up to 8000 snapshots per project or 375 requests/minute</p>

** Commvault software creates 5 snapshots at a time by default, to avoid exceeding **GCP Resource Usage**.

Seed – Snapshot and Streaming MediaAgent Grids

The following is the recommended day-one minimum 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).

GCP 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	Least cost c2-standard-4	Best price/performance c2d-standard-4
Operating systems	Red Hat Enterprise Linux 8 x86_64 (recommended) Microsoft Windows 2019 x86_64	
In GCP Marketplace	No	
Required Persistent Disk Storage	All volumes are zonal or regional persistent disks storage with initial baseline IOPS configuration. - 80GiB Commvault binaries / Logs / Job Results / Index Cache, 3000 IOPS @ 4K - 25GiB Commvault deduplication database, 3000 IOPS @ 4K	
Use for	Performing GCP snapshot creation, sharing, and copying between regions and accounts for Google Compute Instances. Creating service-independent streamed backup copies of GCP Workloads Google Cloud Storage (GCS), Google Kubernetes Engine (GKE), Google Spanner, PostGRES SQL, MySQL, Google Filestore, GCS to GCS 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	<p>**Protect up to 8000 snapshots per project or 375 requests/minute</p> <p>***Avg. throughput of 684GB/hr (backup) / 781GB/hr (restore) observed on a c2-standard-4 equivalent instance***</p> <p>Plan for up to 100TiB of written streamed backup data, 5% streamed per day.</p>

** Commvault software creates 5 snapshots at a time by default, to avoid exceeding **GCP Resource Usage**

*** Avg. throughput used tuned configuration with **ReadAhead=256**, **WriteBehind=256** for optimal transfer speed.

Seed – Cloud Controller

Commvault can deploy a minimal Google Compute instance to perform power management of Google Compute Instance MediaAgents. This is referred to as a **Cloud Controller**. The Cloud Controller is a Compute Instance 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

Pro-Tip
Commvault recommends enabling your cloud controllers in **server groups**, distributed across availability zones or regions for improved resilience of **cloud power management**.

Commvault Google Cloud Storage.

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 the **GCP Machine Type Recommendations** 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 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.

GCP 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 Google Compute instance. Hosts deduplication cloud libraries for backups	
Instance type	Least cost n2-standard-8 – 16 (snapshot) n2-highcpu-16 – 32 – (streamed) n2-standard-8 – 32 (mixed workload)	Best price/performance n2d-standard-8 –16 (snapshot) n2d-highcpu-16 – 32 – (streamed) n2d-standard-8 – 32 (mixed workload)

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 Google Compute Instances and other 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 **GCP Machine Type Recommendations** 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 GCP. These best practices build on the [Google Cloud Architecture Framework](#) and Commvault Cloud-Architected recommendations with a focus on Commvault data management components and configuration.

Compute

Best Practices

- CV-GCPCOMP-BP01 – Always use the smallest Google Compute instance size recommended by Commvault, then scale to meet business RPO and RTO requirements (see [Sizing Guidelines](#))
- CV-GCPCOMP-BP02 – Use Linux-based resources by default for reduced cost, unless prevented by workload protection requirements.
- CV-GCPCOMP-BP03 – Use **GCP ARM** based Access Nodes and/or MediaAgents for best price-performance for cloud-based data management
- CV-GCPCOMP-BP04 – Use GCP AMD(d)-based resources as a least-cost x86_64 alternative when available
- CV-GCPCOMP-BP05 – Use Intel Xeon-based resources as the best price-performance x86_64 for workloads requiring an x86 chipset (e.g. Microsoft O365 protection).
- CV-GCPCOMP-BP06 – Enable **Cloud Power Management** on all MediaAgents to avoid Compute Engine runtime costs when not being used.
- CV-GCPCOMP-BP07 – Consolidate data management resources until availability or RPO/RTO needs require additional resources (i.e., all-in-one CommServe).
- CV-GCPCOMP-BP08 – 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 GCP).

- CV-GCPCOMP-BP9 – Scale data management and movement resources horizontally (MediaAgent, Access Nodes, Search Engines) for more cost-effective network bandwidth and resilience.
- CV-GCPCOMP-BP10 – Always deploy Virtual Server Agent (VSA) and Cloud Apps resources to MediaAgents to provide self-contained recovery grids.
- CV-GCPCOMP-BP11 – 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-GCPCOMP-BP12 – Use **GCP Committed Use Discounts (CUDs)** for Commvault infrastructure where available.

Anti-patterns

- Deploying Commvault infrastructure based on static t-shirt sizes that do not reflect your workload or recovery service levels.
- Selecting Google Compute instances that do not have the minimum specification required by Commvault software.
- Selecting burstable instances (E2) for Commvault deduplication or network streaming compute-intensive workloads.
- Selecting Google Compute 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 **GCP Placement Policies** for Commvault infrastructure with an expected performance increase. Commvault does not require Placement policies for distributed MediaAgent grids or Access Node groups.
- Mixing Compute 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.

Additional Resources

- **Google Compute Virtual machine instances**
- **Cloud instance rightsizing**

Database

Best practices

- CV-GCPDB-BP01 – Utilize a **logical data dump** of **MySQL** database full backups to provide the ability to perform recovery in-place to the original MySQL instance.
- CV-GCPDB-BP02 – Use the Virtual Server Agent and MySQL Agent to do **full** backups of **MySQL** data
- CV-GCPDB-BP03 – Use the Databases Guided Setup in Command Center to backup **MySQL** data.
- CV-GCPDB-BP04 – Add IAM principal with Cloud SQL Viewer permission for **full** backups of **PostgreSQL** data.
- CV-GCPDB-BP05 – Use the Databases Guided Setup in Command Center to backup **PostgreSQL** data.
- CV-GCPDB-BP06 - Complete the Databases **Guided Setup for Google Cloud Spanner**.

Anti-Patterns

- Migrating to GCP MySQL cloud databases and expecting the same data protection practices used on-premises to work without modification.
- Migrating to GCP MySQL cloud databases and expecting the ability to download (backup) and apply transaction logs to your database for roll-forward and roll-back activities.

Additional Resources

- **Cloud Spanner Best Practices**
- **GCP General Best Practices for MySQL**
- **GCP General Best Practices for PostgreSQL**

Management & Governance

Best practices

- CV-GCPMGMT-BP01 – Centrally manage your environments with **GCP Organizations** to simplify permission management to ensure workload account resources are protected with Commvault.
- CV-GCPMGMT-BP02 – Use **Google Cloud Landing zones** to help deploy, use, and scale Google Cloud services.

- CV-GCPMGMT-BP03 – Build a **Resource hierarchy** to manage Commvault resources.
- CV-GCPMGMT-BP04 – Use **project and folder resources** to manage Commvault resources.
- CV-GCPMGMT-BP05 – Configure **GCP Budget alerts** for all GCP accounts that are part of the GCP organization.
- CV-GCPMGMT-BP06 – Use **GCP Deployment Manager** to automate the deployment and configuration of your Commvault all-in-one CommServes from **GCP Marketplace**.
- CV-GCPMGMT-BP07 – Enable **audits** in your Organization, all GCP accounts, and all Regions, written into GCS buckets in a separate GCP account. ⓘ Note: Don't forget to protect your GCS stored audits.
- CV-GCPMGMT-BP08 – Use **GCP Committed Use Discounts (CUDs)** to reduce the cost of baseline Commvault management resources. (ⓘ Note: Commvault reduces your backup compute runtime costs by using power management for access nodes and media agents)
- CV-GCPMGMT-BP09 – Use **GCP Labels** to identify and protect workloads following your business policy, risk-classification, and data-classification

Anti-Patterns

- Operating your organizations with a single GCP account for shared services and protected workload accounts.
- Placing Production, Pre-Production, and Sandbox workloads in the same GCP projects, account and/or networks.
- Placing GCP audit logs in the workload account that generates them, allowing for workload owner log modification or deletion.
- Attempting to perform backup and recovery with Commvault without enabling all the required IAM actions.

Additional Resources

- **Create, edit, or delete budgets and budget alerts**
- **Deployment Manager sample templates**
- **GCP Labels and Tags** (Note: Commvault tracks GCP Labels)

Storage

Design Best Practices

- CV-GCPSTG-BP01 – Use **Balanced persistent disks** for initial builds of Commvault resources for best performance and cost.
- CV-GCPSTG-BP02 – Use **GCP Machine Type Recommendations** to obtain persistent disk
- CV-GCPSTG-BP03 – Use **Persistent disk snapshots** as the *primary recovery point* for low-RTO, rapid recovery of Google Compute instance workloads. Enable snapshot management by **enabling IntelliSnap®** (not default) on your GCP VM Groups in the Commvault Command Center console, via API, CLI, or SDK.
- CV-GCPSTG-BP04 – Use the **Google Cloud Storage Nearline** access storage class for Commvault cloud storage storing backup copies with a data retention period of averaging monthly, 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-GCPSTG-BP05 – Use **Google Cloud Storage Coldline** access storage class for Commvault cloud storage storing backup copied with a data retention period averaging 90 days with low access levels (per quarter for example).
- CV-GCPSTG-BP06 – Use Automatic **Synthetic Full schedules**, to minimize the amount of GET retrieval activity from Cloud storage.
- CV-GCPSTG-BP07 – (Optional) Enable Commvault FIPS-140-2 compliant **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

① Note

Be aware that **object 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.

has a significant CPU load impact.

Anti-Patterns

- Using GCS archive tier to store primary backup data with a high likelihood of recall.

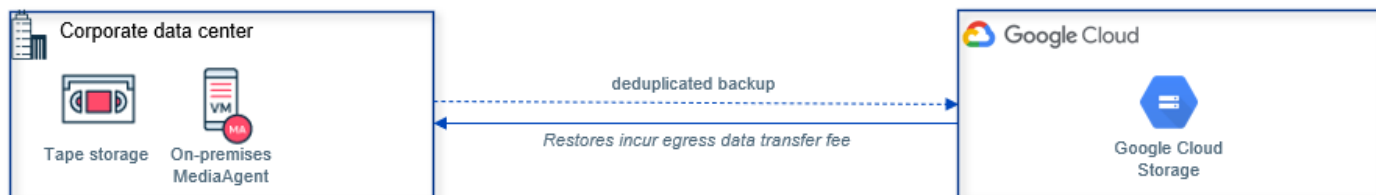
- Using Commvault **extended retention rules** on Cloud Copies, which store long-term retention data alongside near-term backup data.
- Using an on-premises or edge-based MediaAgent to write Primary backup copies to Google Cloud Storage. Any restores will incur Google Cloud Storage data egress costs.

Patterns

The following section provides common patterns employed when building your Commvault data management platform in GCP. Review patterns along with **Best practices**.

Backup on-premises directly to Google Cloud Storage

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.

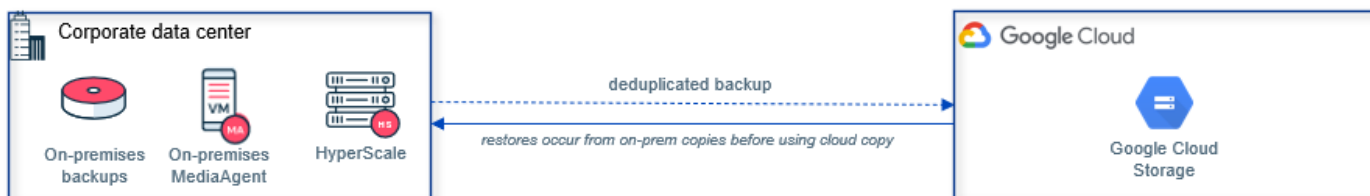


- Short-term operational recovery copies are held offsite in Google Cloud Storage (GCS).
- Leverages on-premises compute infrastructure (MediaAgent, MA) to optimize (deduplication, compress) data before the transfer.
- Allows businesses to completely remove secondary storage infrastructure from on-premises 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 Google Cloud Storage 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.



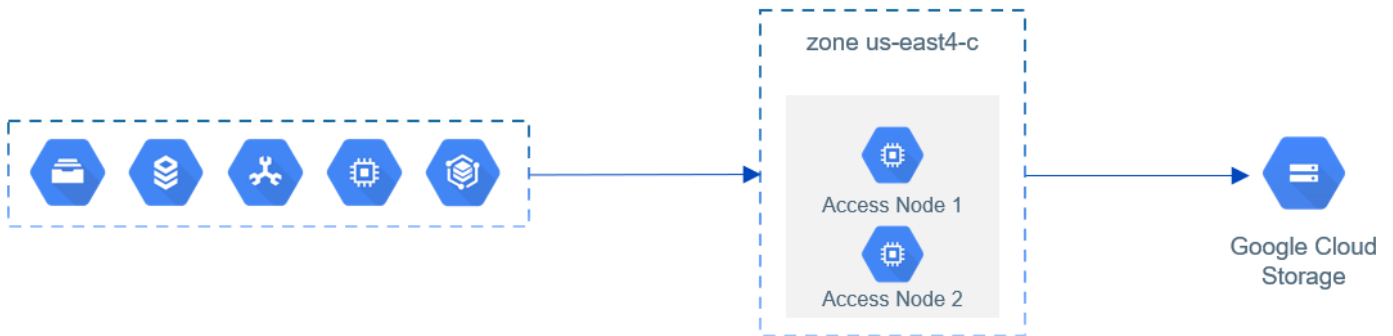
- Short-term operational recovery copies are held onsite disk/object/tape library (e.g. 7 - 30 days).
- Long-term operational and disaster recovery copies are held in the Google Cloud Storage.
- Leverages on-premises compute infrastructure (MediaAgent, MA).

- No object storage 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 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-balance and failover between healthy components in the event of one or more nodes becoming unavailable.

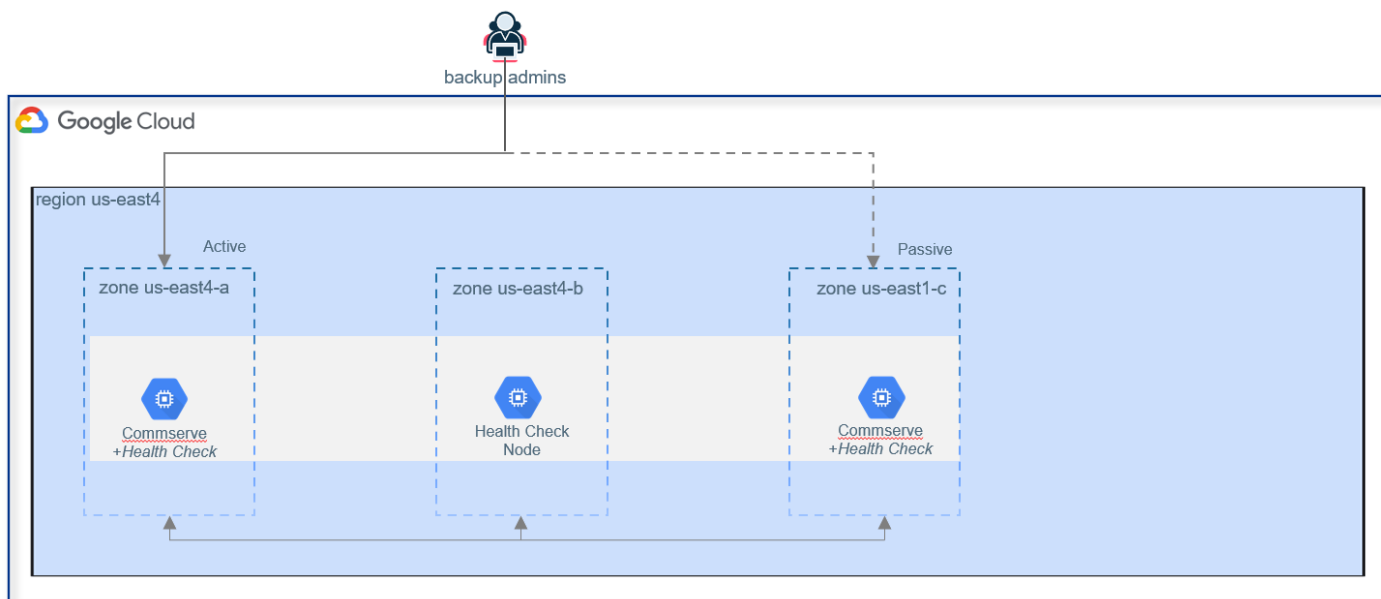


- Access Nodes (AN) are grouped within a zone to provide load-balancing and high availability.
- Co-location of Access Nodes avoids in-region cross-zone data transfer fees.
- This solution protects workloads within the zone at the least cost and may optionally protect other zones with cross-AZ data transfer fees incurred.

This solution provides highly available, least cost, and least latency protection for a single zone.

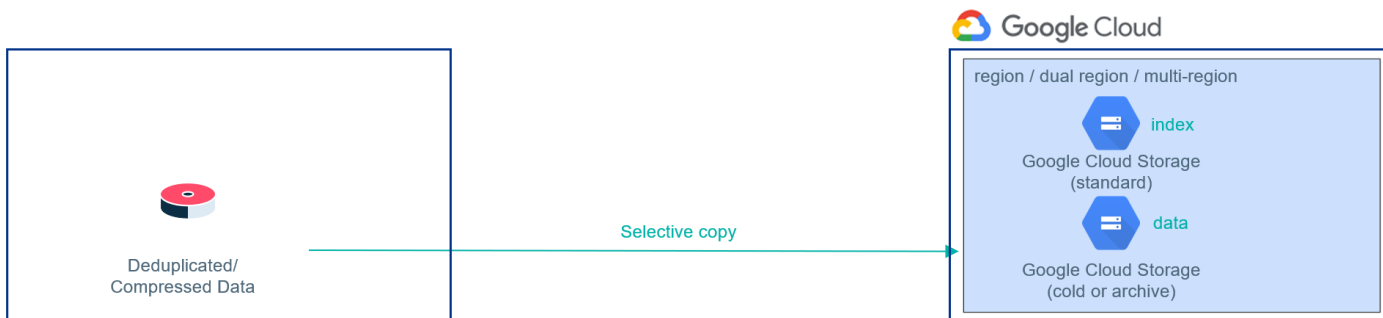
Setup HA/DR for Access Nodes across multiple zones

Commvault CommServe® LiveSync provides the ability to create active:passive deployments that distribute one or more passive CommServe instances across zones (depicted below) or Regions. Commvault implements health checks and automated failover when the Active CommServe becomes unavailable or enters Maintenance Mode.



Archive and deduplicate data to Google Cloud Storage

As businesses look to store more data for future data analytics, visualization, and ultimately for 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.

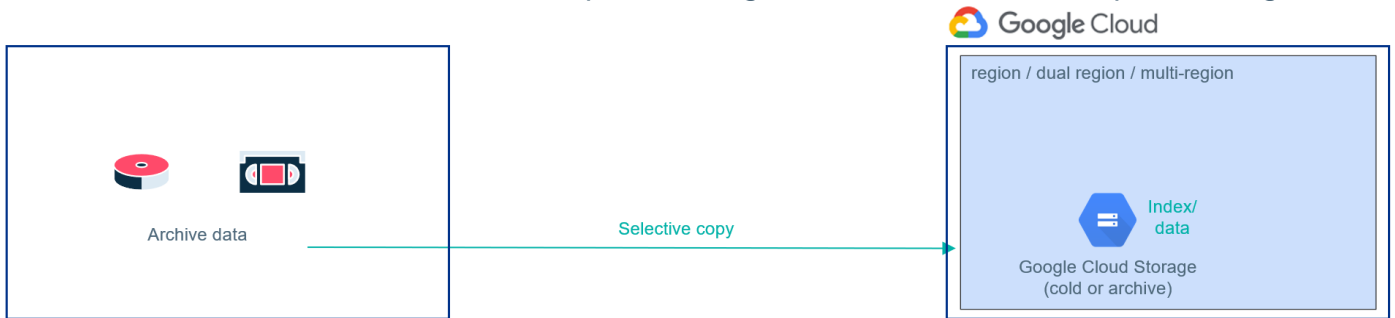


- Archival data represents a subset of backup data residing on-premises or in cloud storage services.
- Archives are typically kept for restoring to their original operational location or a temporary file system for simplified search and retrieval.
- Indexes are stored in frequent access storage classes (standard), and data is stored in an archive access storage class (cold or archive).

- Data may be recalled as a simple restore, by leveraging accessible indexes in frequent access storage classes.
- This Solution provides long-term archive retention with an optimized recall process.

Archive data to Google Cloud Storage

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 Google Cloud Platform 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 GCP cloud.

- Do not perform periodic media data verification

Google Cloud 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 Google Cloud Storage 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. Google Cloud Storage (GCS) is infinitely scalable and ultra-low-cost cloud storage. Commvault will automatically manage and reclaim object storage using micro-pruning but based on the amount of data to be reclaimed, 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 can and will incur ingress and egress fees. See **Google Network pricing**.

- Do not scale vertically when scaling horizontally will deliver the same outcome

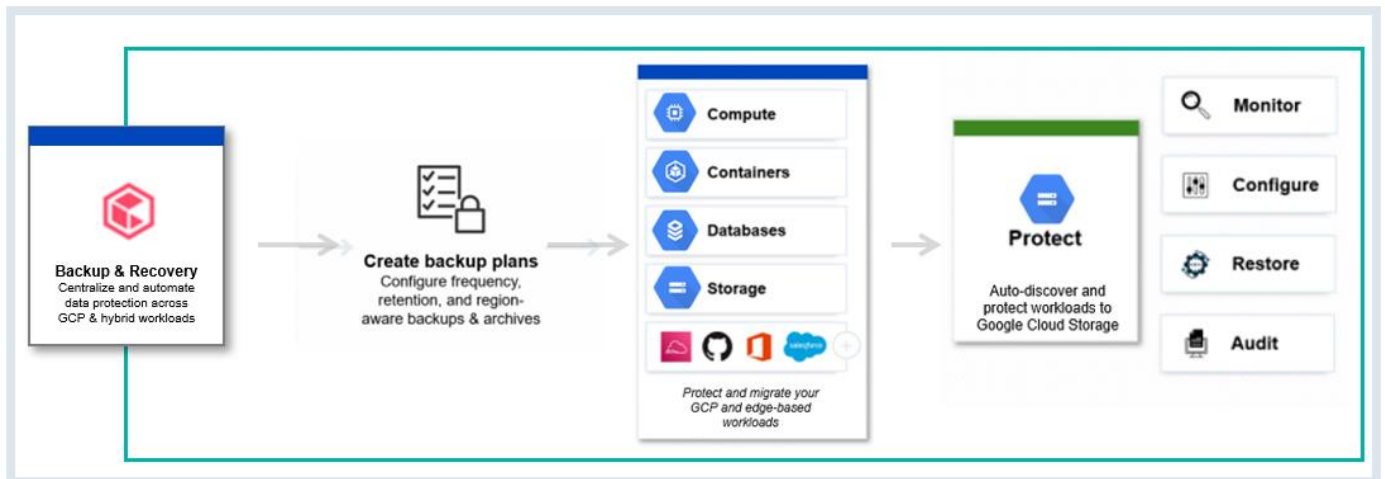
Historically scaling a workload in the data center meant upgrading or life-cycling the compute, storage, and network infrastructure every 3-5-8 years. This often led to an approach of vertically scaling applications within the budget of the individual line-of-business (LOB) and very large vertically 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.

Intelligent Data Use-Cases

Data Protection

Commvault Data Protection lets you rapidly recover data cost-effectively and at scale, in GCP or edge-based data centers. Commvault unifies your data protection by protecting your cloud instances, containers, SaaS services, databases, storage, and traditional applications to Google Cloud Storage.

How it Works



Use Cases

- Complete unified backup & recovery

Back up all business data, including Cloud and edge-based Compute, Containers, Databases, Storage, SaaS, and traditional Applications. Recover across accounts and GCP Project, regions, and zones including Google Anthos.

- Anywhere to GCP disaster recovery

Quickly recover mission-critical operations by restoring virtual machines, databases, file systems, and object stores to GCP 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 Google Cloud Storage. Migrate large datasets offline with **Google transfer appliances** in network-constrained edges.

How to get started

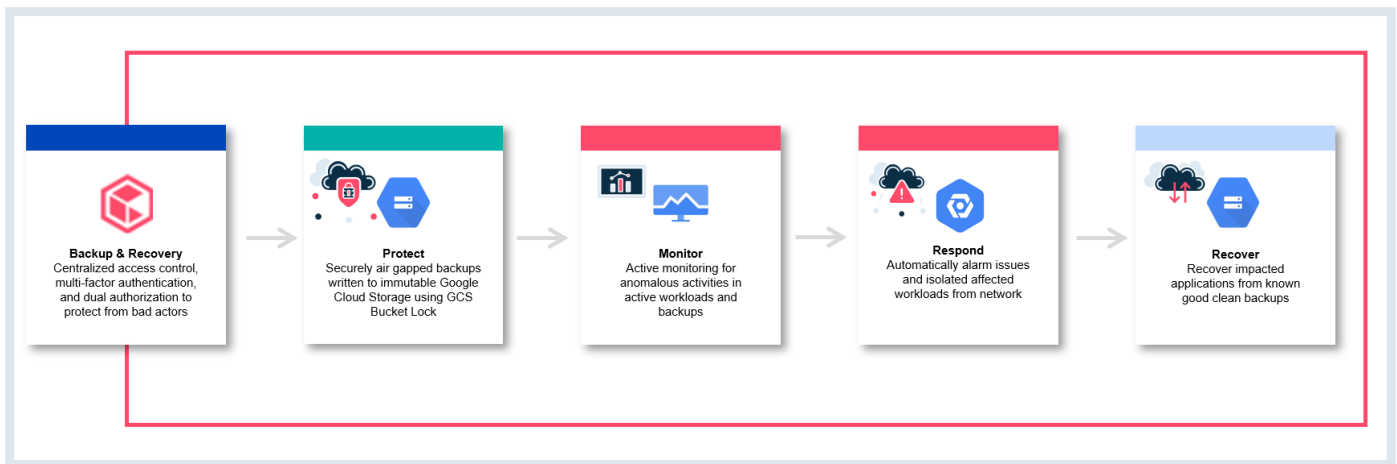
- At least one GCP project configured with a **service account** with access to resources to protect.

- At least one MediaAgent/Access Node (in GCP or on-prem) to optimize and read/write data to Google Cloud Storage.
- A network connection between your workload location and Google Cloud Storage (**GCP dedicated interconnect, GCP Cloud VPN**)
- Refer to docs.commvault.com for a list of all protected workloads, including **GCP resources**.
- ① Note: Remote offices can write directly to Google Cloud Storage without a requirement for storage or tape gateways.

Data Security

Commvault Data Security helps you detect, protect, and recover from ransomware attacks and other data breaches affecting your GCP and edge-based workloads.

How it works



Use cases

- Air gap and harden your backups

Securely air gap your backup copies to mitigate lateral moving threats and prevent modification by writing to immutable **Google Cloud Storage Bucket Lock** buckets.

- Be alerted to anomalous threats

Be automatically notified of anomalous threats across your active workloads and backup data. Automatically respond to alarms by isolating potential threats and locking backups for forensic SecOps investigation.

- Rapidly recover to a known good state

Recover with cloud-scale by recovering infected workloads in highly parallelized restores aimed at recovery from known good backups.

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 encrypting everything which slows attackers who do not have access to your KMS keys.
- Enable business-logic workflows to require **dual-authorization** for high-risk changes or insider threats.
- Enable **Google Cloud Storage Bucket Lock** for business-critical backups requiring added protection from unintended modification.
- Review and implement the **Ransomware Protection Best Practices** from Commvault.
- Review **Mitigating ransomware attacks using Google Cloud** for guidance on securing your GCP resources.

Works with

- Google Compute Engine (GCE)
- Google Kubernetes Engine (GKE)
- Filestore
- Cloud Spanner
- Cloud PostgreSQL, MySQL
- Cloud SAP HANA
- Google Cloud VMware Engine (GCVE)

Protect Google Cloud Databases

SAP HANA support

Commvault supports Google Cloud Snapshots protecting SAP HANA. This feature enables the ability to back up HANA databases from Google Cloud on servers using IntelliSnap to create snapshots of the Data Volume mounts in HANA. These snapshots can be used for recovery of the database in-place or cross server, as well as creating clones of the database to another server in the cloud, and lastly revert in place to quickly revert the data volume from the snapshot for fast recovery of large databases.

Google Cloud instances running Linux versions that support HANA as per SAP guidelines is required. The block devices used for the HANA data volume mount can be a single virtual or persistent disk hosted on the supported Linux filesystem, or a collection of block devices configured as LVM using LVM2 on Linux. If LVM is configured, lvm2-lvmetad service must be disabled as documented in **Commvault Documentation**. The instance must have compute engine access set to Read/Write.

Supported operations include:

- Snapshot backup
- Backup copy
- Restoring from snapshot
- Restoring from backup copy
- Cloning

Cross Project mounting and Encrypted disk is currently not supported.

MySQL / PostgreSQL support

Commvault provides a complete data protection solution for GCP MySQL and PostgreSQL databases by automating backup operations and by providing the following recovery methods:

- Restore from GCP MySQL to GCP MySQL
- Restore from GCP MySQL to an on-premises MySQL server
- Restore from an on-premises MySQL server to GCP MySQL
- Similar combination with PostgreSQL supported

Best Practice: The MySQL server on the Access Node should run the latest release to ensure the proxy server is in sync with the GCP instance. Add the MySQL server instance using Command Center. When you restore an on-premises database to the GCP cloud, the restored database uses the standard tier model.

For more information on Google Cloud MySQL, see [Google Cloud Database for MySQL](#).

For more information on Google Cloud PostgreSQL, see [Google Cloud Database for PostgreSQL](#)

Google Cloud Spanner Support

Google Cloud Spanner is a horizontally Scalable RDBMS (Relational Database Management System) Google Cloud Spanner is especially suited for applications requiring:

- Strong global consistency
- Database sizes exceeding ~2TB
- Many IOPS (Tens of thousands of reads/writes per second or more)

Commvault supports [Google Cloud Spanner](#) agentless backup and restores. This includes entire instances or individual databases. In place or out of place restores are supported.

Requirements:

- An Access Node must have the Virtual Server Agent and Cloud Apps Agent installed to support backup and recovery of Spanner. For Access Node operating system requirements, see [Google Cloud Spanner Access Nodes](#).

- Your Google account must contain Cloud Spanner, cloud storage, and a compute engine. The Dataflow API must be enabled for your account.
- See **Roles and Permissions** to verify the Google Cloud Service Account has the correct permissions to access the Google resources

For more information on Backup and Recovery of Google Cloud Spanner, see **Google Cloud Spanner** in Commvault Books Online.

Google Cloud specific workloads

Virtual machine recovery into GCP VM instances

The Commvault Virtual Server Agent provides the ability to easily perform direct conversion of protected VMs to GCP instances from the following hypervisors:

- VMware
- AWS
- Azure

Backups can be stored either within Google Cloud Storage (GCS), another Cloud Library, or to an on-premises disk library.

This process is used as part of a disaster recovery strategy using GCP as a Cold DR site, or as a migration strategy (Lift-and-Shift). Additional details can be found **Cross-Hypervisor Restores**.

Protecting and recovering active workloads in Google Cloud

Application consistency can be achieved within Google Compute instances with application consistent snapshots.

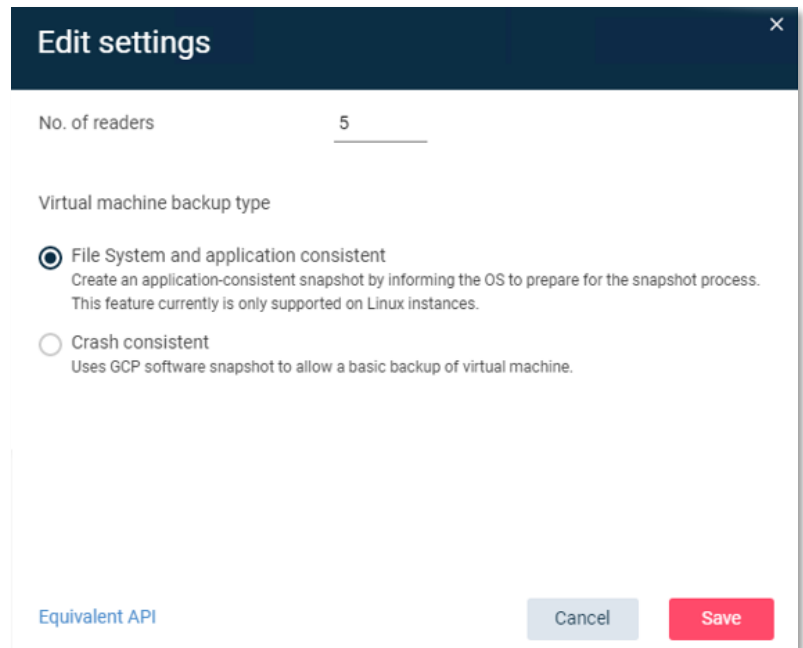
Application Consistent Snapshot support for Linux instances

This utilizes the guest-flush option on Linux hosts. The instance will be quiesced just before a snapshot is taken, which provides a consistent backup. Application consistent snapshots capture the state of the application data at the time of the backup with all application transactions completed and all pending writes flushed to the disk.

To prepare the Linux guest environment for application consistency, custom shell scripts must be executed on each instance before and after the snapshot. See **Creating the Script Files** for details on creating and executing the scripts.

The following GCP permissions are required on the IAM or Service account(s) to perform Application Consistent Snapshots:

- `compute.disks.createSnapshot` on the disk



- `compute.disks.addResourcePolicies` on the disk
- `compute.snapshots.delete` on the snapshot
- `compute.resourcePolicies.create` on the resource policy

To enable application consistency in Command Center, edit the VM Group settings and **File System and application consistent**.

See [Enabling Application-Consistent Backups](#) for more information.

Agentless VM protection (Virtual Server Agent)

The Virtual Server Agent (VSA) delivers an agent-less, block-level capture of Google Compute instances and their attached persistent disks. Restoration options include Full virtual machine recovery and granular-level file recovery. These features provide customers with additional functionality when protecting and recovering IaaS VMs in GCP. Commvault added the ability to backup customer managed encrypted disk using the [Google Cloud Key Management Service](#). These features provide customers with additional functionality when protecting and recovering Google Compute instances.

A GCP Service account and its corresponding JSON file is used to backup and restore GCP instances. Downloading the JSON file is used for service account authentication. For backing up instances in multiple GCP projects, the service account will need access to all the projects where the instances exist, including the projects with access nodes.

For more information on creating the GCP service account see [Creating a Google Cloud Platform Service Account](#).

When to use the VSA for Google Cloud

Agent-less protection approach for Google Cloud instances and file-level data. Agent-in-guest is not required to perform block-level backup and File-level recovery.

When not to use the VSA for GCP

When you require application-consistent backups for applications not supported by application aware VSA – the VSA for GCP approach creates a crash-consistent image of the source instance and its attached block volumes. If you require application consistency, use an agent-in-guest either standalone or in conjunction with the VSA.

Protecting worker/stateless VMs – Worker nodes may generate valued data that is moved to another centralized repository and the nodes themselves do not require protection. It is recommended to instead target that centralized repository for Data Protection instead of the individual worker nodes, whether with VSA for GCP or agent-in-guest, depending on the required level of backup (crash vs. application consistent).

Do you need help?

Do you still need help architecting, designing, and deploying your cloud-native GCP 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.

- 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 complements 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 gcp@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.

Revision History

Version	Data	Changes
2022E	February 2023	<ul style="list-style-type: none">• Commvault Platform Release 2022E (Newsletter)• Added GCP Marketplace information• Added Cloud Shared Responsibility• Added Zero Trust Model• Added Ransomware Protection• Added Design and Best Practices• Added Solution Design tool information• Added Sizing Guidelines• Added Patterns• Added Intelligent Use-Cases
11.26	December 15 th , 2021	<ul style="list-style-type: none">• Updated document from 11.20 to 11.26 functionality• Updated Reference Diagrams• Support for regional disk backup and in-place restores• IntelliSnap Support for GCP Virtualization• Support for disk level filters• GCP to VMware conversion• Support for Spanner
11.20	June, 2020	<ul style="list-style-type: none">• First version

Additional Resources

Community Forum

Commvault Community Forum 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.

Datasheets

- **How to Migrate Legacy Backups to the Cloud**
- **Commvault Complete™ Data Protection**
- **Greater data protection: Immutable backups to the cloud with Commvault**
- **Solution brief: Google Cloud Backup and Recovery**
- **Data management architecture design and implementation**

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.

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.