From The Architects

Protecting Google Cloud VMware Engine

(GCVE)
Using Veeam

Design Recommendations

Patricio Cerda

Content Table

Introduction

Many companies are planning or are already in the process of migrating at least a part of their workloads to the cloud. There are multiple strategies to accomplish this and one of those strategies is running VMware workloads natively in a cloud provider.

Google Cloud VMware Engine is one of the options that you have, to run VMware workloads natively in a public cloud, in this case in Google Cloud, and leverage all the benefits of a cloud infrastructure.

Of course, as with any other environment, you need to protect your workloads running in Google Cloud VMware Engine (GCVE). The goal of this document is to provide recommendations and best practices to protect GCVE workloads using Veeam solutions.

What is Google Cloud VMware Engine

Google Cloud VMware Engine is a fully managed service that lets you run the VMware SDDC platform in Google Cloud. VMware Engine provides you with VMware operational continuity so you can benefit from a cloud consumption model and lower your total cost of ownership. VMware Engine also offers on-demand provisioning, pay-as-you-grow, and capacity optimization.

The VMware environment runs natively on bare metal infrastructure in Google Cloud locations and fully integrates with the rest of Google Cloud. Google manages the infrastructure and all the necessary networking and management services so you can consume the VMware platform efficiently and securely.

Using the same VMware-based infrastructure and operations stack, Google Cloud VMware Engine delivers a consistent operational experience that allows you to seamlessly migrate and manage workloads using the same VMware technologies and tools you are familiar with today. The best part is you also continue to use the same knowledge and skills you've obtained in your on-premises environment in the public cloud.

Main use cases for Google Cloud VMware Engine are:

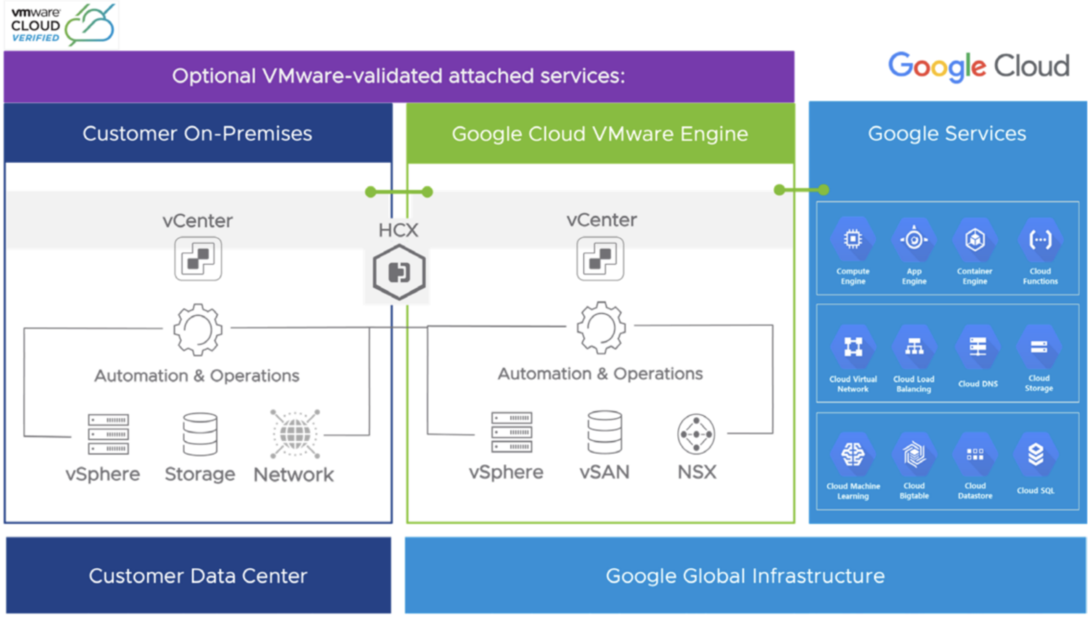
- Accelerate Cloud Migration
- Disaster Recovery
- Make IT Operations more efficient.

Components and Architecture of GCVE

Google Cloud VMware Engine includes the following VMware components, which make this solution fully compatible with your existing VMware tools, processes, and skills training.

- **VMware vSphere**: Virtualization platform to run Virtual machines alongside with modern containerized apps.
- **VMware NSX-T**: Software-Defined Network (SDN) solution that allows to connect and protect apps across the data center, multi-cloud, bare metal and containers.

- **VMware HCX**: VMware solution that seamlessly extend your on-premises environments into cloud, which allows to streamline the application migration, workload rebalancing and business continuity across data centers and clouds.
- **VMware VSAN**: Software-Defined Storage solution that provide storage resources improving scalability, agility and performance, while reducing costs and complexity compared with traditional hardware-based storage solutions.



All these solutions enable your team to manage workloads without disrupting existing policies, such as those related to networking, security, data protection, and auditing.

Protecting GCVE Workloads with Veeam

Why we must protect GCVE workloads

Google Cloud provides a cloud platform that allows you to create and deploy applications. However, as mentioned in the Google Shared Responsibility Model, you are still responsible for protecting and securing your own data.

This basically means Google is responsible for all the managed services provided by Google Cloud and the underlying infrastructure, but they are not responsible of the data hosted on their services/infrastructure.

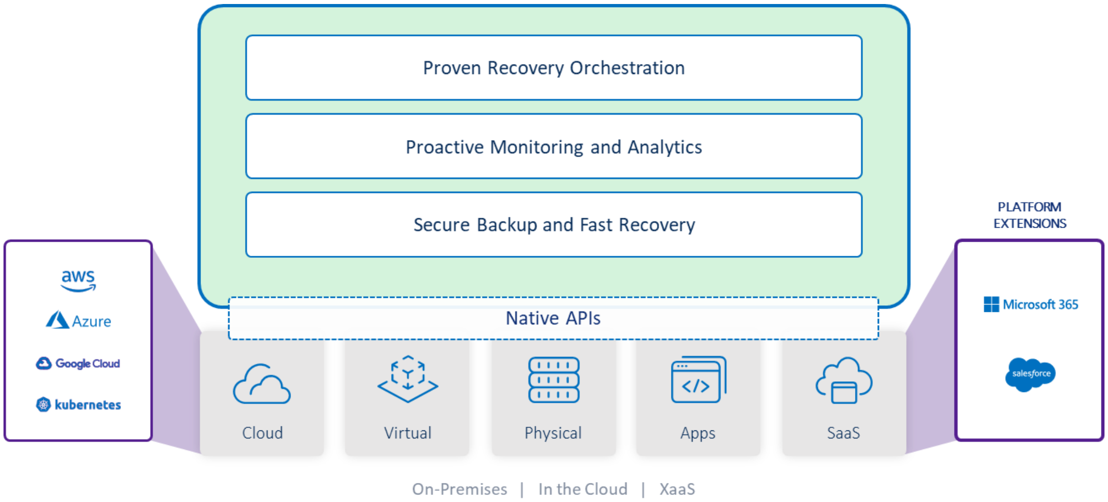
This applies of course to every workload running on GCVE, which are basically VMware virtual machines.

How to protect GCVE workloads

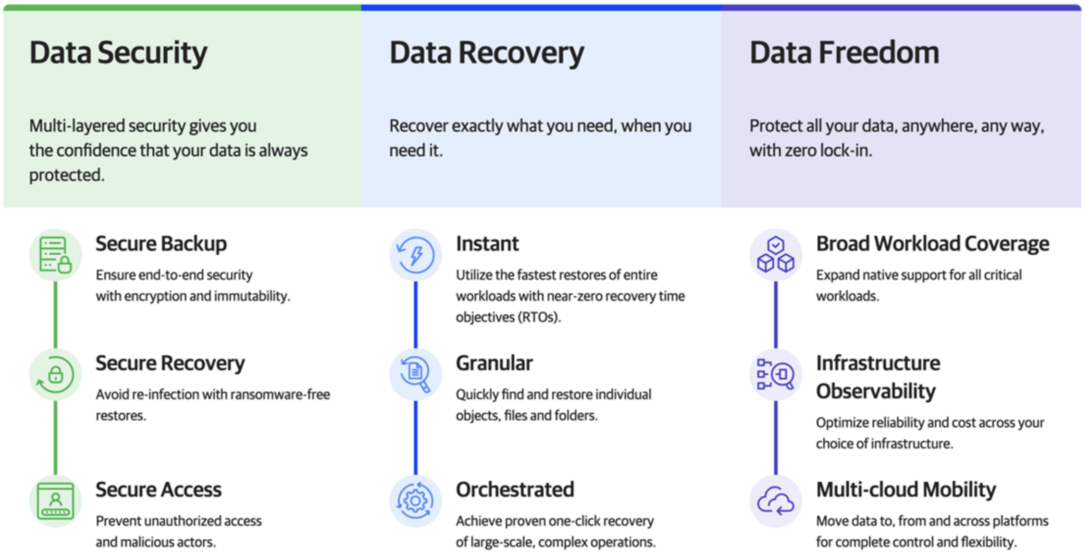
The best way to protect GCVE workloads is using a backup solution focused on protecting VMware virtual machines, like **Veeam Backup and Replication v12**.

Veeam Backup & Replication is the single backup, recovery and data security solution for all your workloads both on-premises and in the cloud. As the foundation of Veeam Data Platform, Veeam Backup & Replication delivers simple, flexible, reliable and powerful data protection. Eliminate downtime with **instant recovery** and stay safe from cyberthreats with native immutability and tested backups, all from one software-defined, hardware-agnostic solution.

Veeam Backup & Replication is part of the new Veeam Data Platform, a complete solution for delivering Data Security, Data Recovery and Data Freedom to your entire hybrid environment.


Proven Recovery Orchestration
Proactive Monitoring and Analytics
Secure Backup and Fast Recovery
Native APIs
aws
Azure
Google Cloud
kubernetes
Cloud
Virtual
Physical
Apps
SaaS
PLATFORM EXTENSIONS
Microsoft 365
salesforce
On-Premises | In the Cloud | XaaS

Veeam Data Platform is based on three foundational pillars as described in the image bellow:



Data Security

Multi-layered security gives you the confidence that your data is always protected.

Secure Backup
Ensure end-to-end security with encryption and immutability.

Secure Recovery
Avoid re-infection with ransomware-free restores.

Secure Access
Prevent unauthorized access and malicious actors.

Data Recovery

Recover exactly what you need, when you need it.

Instant
Utilize the fastest restores of entire workloads with near-zero recovery time objectives (RTOs).

Granular
Quickly find and restore individual objects, files and folders.

Orchestrated
Achieve proven one-click recovery of large-scale, complex operations.

Data Freedom

Protect all your data, anywhere, any way, with zero lock-in.

Broad Workload Coverage
Expand native support for all critical workloads.

Infrastructure Observability
Optimize reliability and cost across your choice of infrastructure.

Multi-cloud Mobility
Move data to, from and across platforms for complete control and flexibility.

In the next section we will discuss the design recommendations and considerations when using Veeam Backup and Replication to protect workloads running on GCVE.

GCVE Network Considerations

Connecting GCVE with Google Cloud network

Connecting the GCVE deployment with the Google Cloud network allows to integrate the VMware infrastructure with Google Cloud and consume Google Cloud services. For instance, by using this integration you could complete the following tasks:

- Providing connectivity between VMware workloads and Google Cloud Engine VMs
- Connecting to Google Cloud Storage from GCVE
- Monitoring all your public, private, and hybrid applications by using Cloud Monitoring
- Importing data from databases into BigQuery for analytics

Connectivity with Google Cloud Storage is of high importance when we design a Veeam solution to protect VMware workloads running in GCVE.

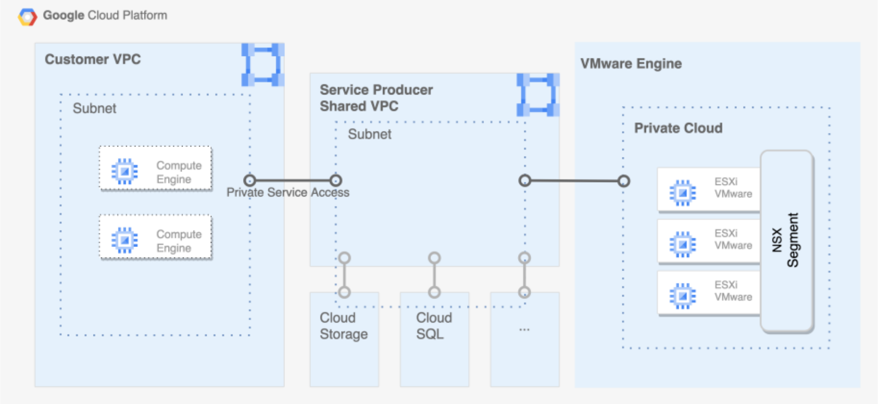
Private Service Access

Private services access is a private connection between your Google Cloud Virtual Private Cloud (VPC) network and networks in VMware Engine.

Private services access enables the following behavior:

- Exclusive communication by internal IP address for virtual machine (VM) instances in your VPC network and VMware VMs. VM instances don't need internet access or external IP addresses to reach services that are available through private services access.
- Communication between VMware VMs and Google Cloud-supported services (like Google Cloud Storage) which support private services access using internal IP addresses.
- Use of existing on-premises connections to connect to your VMware Engine private cloud, if you have on-premises connectivity using Cloud VPN or Cloud Interconnect to your VPC network.

You can set up private services access independently of VMware Engine private cloud creation. The private connection can be created before or after creation of the private cloud to which you want to connect your VPC network.

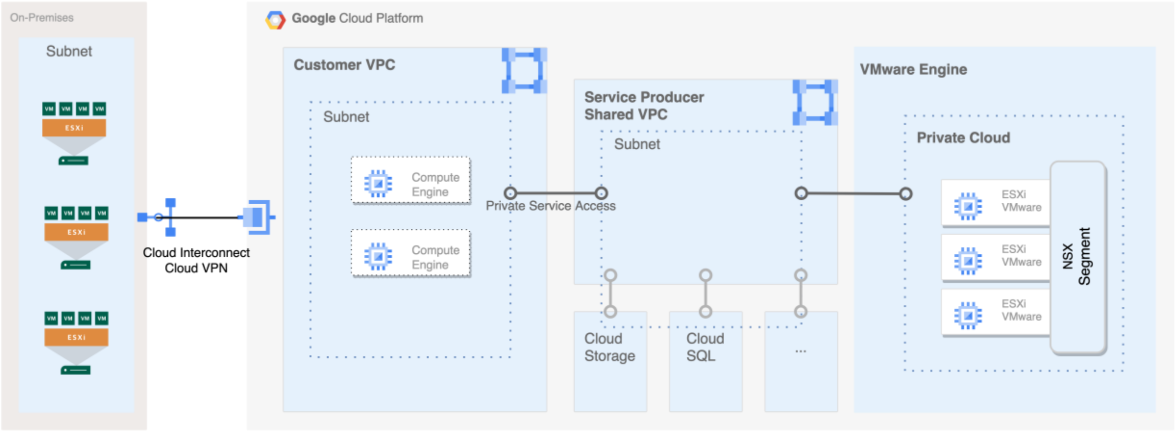


Connecting GCVE with on-premises infrastructure

For Hybrid solutions, when using Google Cloud VMware Engine alongside with a traditional VMware on-premises infrastructure, you must provide proper connectivity between GCVE and your on-prem network.

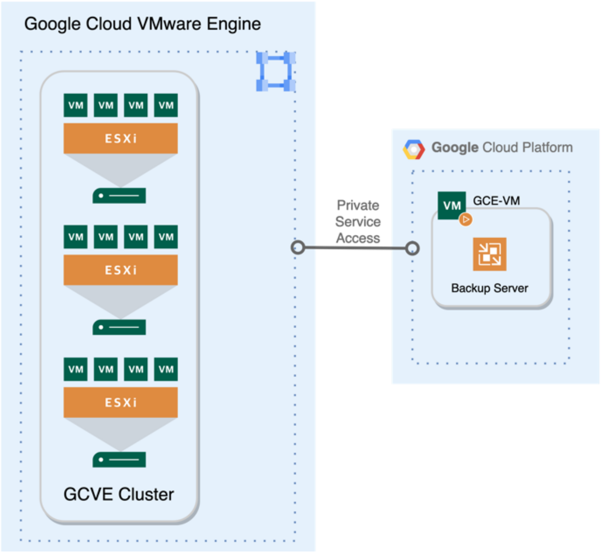
A Hybrid solution could be useful when migrating workloads from on-premises to GCVE, also to balance the workloads between multiple sites, or even for Disaster Recovery purposes, using replication solutions to replicate VMs from on-premises to GCVE or vice versa.

In order to have a hybrid solution, it's necessary to connect GCVE with the on-premises network, which can be done by using **Google Cloud Interconnect**, or **Google Cloud VPN**.



Design recommendations for Veeam components

Veeam Backup Server



In general, according to the Best Practice guide, the backup server should be placed in the main data center, alongside the infrastructure to be protected, to leverage quick response times and local management traffic, which would mean deploying the Veeam Backup Server as a GCVE VM. However, you could benefit of leveraging Google Cloud to deploy Veeam Backup Server as a Google Compute Engine (GCE) VM instead.

By using Private Service Access as mentioned earlier, all communications between GCVE and Google Cloud is optimized, so response times shouldn't be an issue. In addition, by using a GCE VM, you can make Veeam Backup Server independent of the protected infrastructure. This could be important for Disaster Recovery scenarios, where the entire GCVE cluster has failed, as by having Veeam Backup Server running in a separated infrastructure, you will still have access to all backup data. This also applies to scenarios using replication, as mentioned later in this document.

Therefore, it is recommended to deploy Veeam Backup Server as a Google Compute Engine VM. The machine must run Microsoft Windows and meet all Veeam requirements.

Sizing

For Veeam Backup Server sizing, there isn't any different recommendation for GCVE than for other VMware environments. You can follow the sizing recommendations for Veeam Backup Servers provided in the Veeam Best Practices Guide, or you can even use the the Veeam Size Estimator (VSE), where you can get the sizing for most of the Veeam Backup & Replication components.

Proxy Servers

The backup proxy is the component that reads the data from the source infrastructure, elaborates, and transfers it to the final destination: either the backup repository or the DR infrastructure (in case of replication). Proxies are the work horses and are critical components to achieve good backup and restore speeds.

Veeam Proxy Server design depends upon the underlying storage technology used by VMware, especially when choosing the proper Transport mode and the Proxy placement.

Transport Mode

When we work with GCVE, we must consider that the VMware infrastructure is using VSAN to provide storage resources for all the Virtual Machines, and therefore we should follow the design recommendations to use Veeam with VSAN.

For VMware VSAN, the recommendation is using Virtual Proxy Servers, which will use the **Hot-Add transport mode** (also known as Virtual Appliance mode), as this mode provides the most efficient communication between the Proxy Server and the VSAN datastores.

Proxy Placement

About where the Proxy Server should be placed, it is recommended to have the Proxy Servers as close to the source data as possible with a high bandwidth connection. The main reason is that the proxy needs to read 100% of the source data from the datastore before optimizing and reducing it for further transfer.

In the GCVE context, we have already mentioned that the Proxy Server should be virtual and using the Hot-Add transport mode. Therefore, and considering the placement recommendation, the Proxy Servers should be deployed in the same VSAN cluster provided by GCVE. If you have multiple VSAN clusters in GCVE, you should deploy at least one Proxy Server per VSAN cluster.

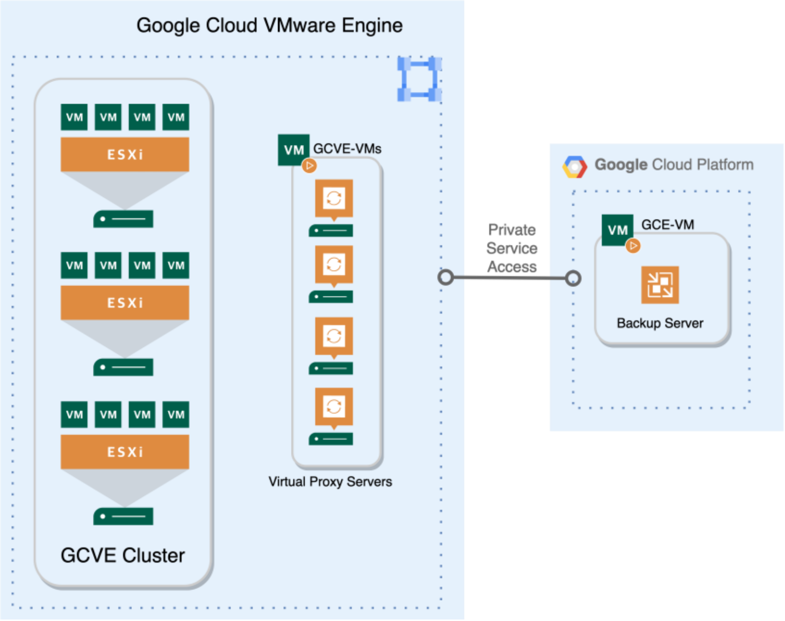
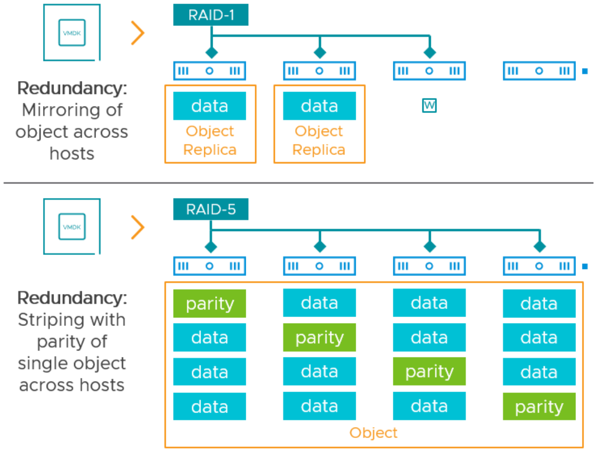


Image 01: Virtual Proxy Servers running in GCVE

Sizing

Getting the right amount of processing power (compute resources) is essential to achieving the RTO/RPO defined by the business. You can follow the sizing recommendations for Proxy Servers provided in the Veeam BP Guide, or you can even use the Veeam Size Estimator (VSE), where you can get the sizing for most of the Veeam Backup & Replication components.

One of the major concerns when using Veeam to backup VMware VMs on VSAN, is the **VM data placement** in the VSAN datastore, and the inter-ESXi traffic through the VSAN network. As you can see in the image, depending on the Storage Policy used in VSAN, every VM component, including Virtual Disks (vmdk) will be distributed among multiple ESXi hosts, whether by creating multiples replicas of every object (Mirroring/RAID-1) or by using Striping with Parity (Erasure Coding – Similar to RAID-5)



So, how Veeam will choose the proper Proxy Server to backup every VM in a VSAN cluster?

According to the official Veeam documentation: *If you have several Backup Proxies on a VSAN cluster, Veeam will choose the most appropriate Backup Proxy to reduce the backup traffic on the VSAN cluster network. To choose a Backup Proxy, **Veeam checks the HDDs directly attached to every ESXi host** and calculates the amount of VM data on these HDDs. The preference is given to the ESXi host that has a direct access to an HDD with the maximum amount of VM data. This approach helps reduce workload on the ESXi I/O stack during data transport*.

In other words, Veeam is smart enough to select the proper Proxy in a VSAN cluster, to reduce the traffic in the VSAN Network. However, at this point another question could arise: how many Proxy Servers should I deploy in a VSAN cluster?

The answer has several parts:

- First, we must complete the Proxy Server sizing as mentioned before, to calculate the total amount of compute resources required for the Proxy role, according to the amount of data to be protected. For instance, let's say we need **32 CPU Cores and 64GB of RAM** for the Proxy role to protect all our workloads in 8 hours.
- Second, for virtual proxy servers, it is recommended to configure multiple proxies with maximum **8 vCPUs** to avoid co-stop scheduling issues. So, in our example, we should have **at least 4 Proxy Servers**, each of them with 8 vCPU and 16GB of RAM, to provide all the compute resources required.
- Third, in previous releases, the Veeam recommendation was to provide one Proxy Server per VSAN Node (ESXi host) to prevent having excessive network utilization during backup operations. However, for most larger environments that utilize 10GbE connections throughout the environment, **there should be plenty of throughput** to allow for fewer proxies. Even more, according to GCVE documentation, in a GCVE cluster you can get up to 100Gbps of bandwidth for east-west networking, so it shouldn't be a problem to have enough throughput for backup and restore operations without having one Proxy per VSAN node.

So, summarizing, it's not mandatory to have 1 Proxy server per VSAN node. Therefore, the number of Proxy Servers to be deployed depends entirely on the amount of **compute resources** required, and the maximum recommended **size per Virtual Proxy**, as mentioned before. Nevertheless, if you prefer to have one proxy per VSAN node, that's also entirely possible.

Backup Repositories

A backup repository is a storage location where Veeam keeps backup files, VM copies and metadata for replicated VMs. It is necessary to provide Backup repositories with enough capacity, redundancy, and performance to keep the backup files, and meet the SLA requirements.

The Veeam Backup Repository can be located wherever the environment allows it. Of course, it's never recommended to use the same Storage resource used for Production workloads as a Veeam Backup Repository.

With Veeam Backup & Replication V12, there are several storage options to be used as backup repository:

- Direct Attached Storage: Virtual or physical servers based on Windows or Linux.
- Network attached storage: Network shares using SMB (CIFS) shares or NFS shares.
- Deduplicating storage appliances
- Object Storage: AWS S3 Buckets, Azure Blob, Google Cloud Storage, S3 compatible.

When using GCVE, the VMware cluster will use VSAN datastore as the storage solution for all workloads in the cluster. As mentioned before, it's not recommended to use the same Production storage to keep the backup files, so this VSAN datastore can't be used as a Backup Repository.

In the GCVE environment, we don't have other storage solutions, but we can leverage Google Cloud resources to provide the components required for a Veeam Backup Repository, alongside with storage resources provided by other cloud providers as AWS and Azure. There are three main options to be used as a Backup Repository in GCVE: Block Storage with **Direct Attached Storage (DAS), NFS/SMB Shares and Object Storage**.

NOTE: In this document, regardless that all three storage options will be explained, we will focus on designing a data protection solution using **Object Storage** for backup repositories.

It is also important, when designing the Data Protection solution, to follow the **3-2-1** rule recommended by Veeam. The 3-2-1 rule means:

- Maintain at least **3 copies** of your data. It means that in addition to your primary data, you should also have at least two more copies/backups.
- Store the backups on **2 different** media.
- Keep at least **1 of the copies** at an offsite location

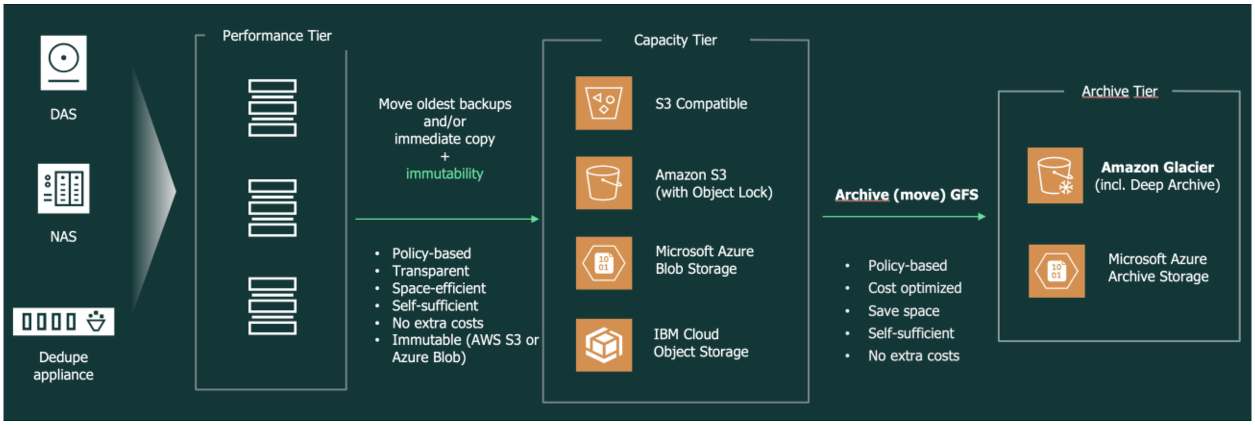
In this section we will discuss multiple options for Backup Repositories, including recommendations to follow the 3-2-1 rule.

Scale-Out Backup Repository

One of the options you have when designing Veeam Backup Repositories, is the use of Scale-Out Backup Repository (SOBR).

A SOBR is a repository system with horizontal scaling support for multi-tier storage of data. A SOBR consists of one or more backup repositories or object storage repositories called **performance tier**, and can be expanded with object storage repositories for long-term and archive storage: **capacity tier** and **archive tier**:

- **Performance tier** is the level used for the fast access to the data. It consists of one or more backup repositories or object storage repositories called performance extents.
- **Capacity tier** is an additional level for storing data that needs to be accessed less frequently. However, you still can restore your data directly from it. The capacity tier consists of cloud-based, or on-premises object storage repositories called capacity extent.
- **Archive tier** is an additional level for archive storage of infrequently accessed data. Applicable data can be transported ether from the capacity or archive tier. For restore from the archive tier, data must undergo preparation process.



In this document we would include the use of SOBR as a recommendation to provide the required capacity for backups using different options for storage, and to facilitate following the 3-2-1 rule.

For more information about Scale-Out Backup Repositories, including all requirements and limitations, please refer to the official Veeam documentation.

Repository Design with Block Storage

By using Google Cloud Engine (GCE) it is possible to create Virtual Machines with persistent storage, which of course can be used as a Backup Repository. These GCE instances can be based on Linux or Windows, where the Linux option is the most recommended to also provide Immutability by configuring a Linux Hardened Repository.

Even more, it is entirely possible to configure a Scale-Out Backup Repository (SOBR) to use the GCE Instance and its persistence storage (block storage) as a Performance Tier, and then offload the backups to a Google Cloud Storage Bucket using Capacity Tier.

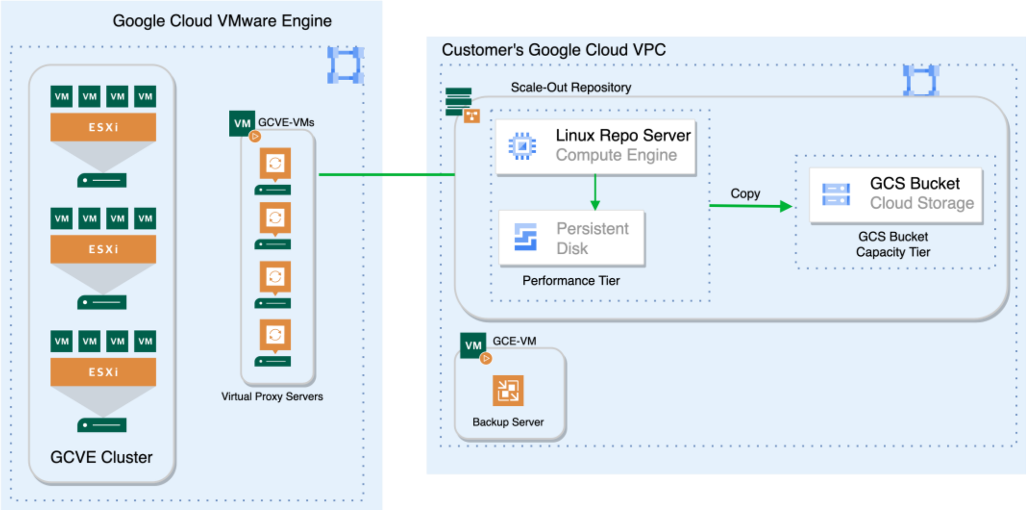


Image 02: SOBR with Block Storage using GCE Instances and Persistence Storage

By using SOBR with Capacity Tier, and the Copy option, it will be possible to follow the 3-2-1 rule recommended in the Veeam Best Practice Guide, to have multiple copies of your data.

NOTE: It's important to highlight that for Capacity Tier we have more options than just using a Google Cloud Storage bucket. We could use AWS S3, Azure Blob or any other S3 compatible bucket. Of course, when using GCVE, for performance and costs reasons, the most recommended option is stick with GCS buckets.

Repository Design with Object Storage

Starting with Veeam Backup & Replication V12, it is also possible to send the backups **directly to an Object Storage**, without the need of sending the backup to a block storage first. This provides a great option when designing the Backup Repository to protect workloads running in a GCVE cluster, as it's not necessary anymore to use a GCE instance with persistence disk (block storage) as we described earlier. We can send the backups directly to a GCS Bucket.

Now, at this point it's also important to highlight that, so far, Veeam Backup and Replication doesn't support using Google Cloud Archive Storage (only Standard and Nearline are supported), so we can't use SOBR with Performance + Archive Tier for GFS and archiving purposes. Still, using a SOBR with **Performance** and **Capacity** Tier is entirely possible, and allows to follow the 3-2-1 rule.

There are other options, too. As you can see in the image below (Image 03), it is entirely possible to use Scale-Out Backup repositories by leveraging Object Storage resources from other Cloud Providers. For instance, it is possible to create a SOBR using an AWS S3 bucket for Performance Tier, and then using AWS S3 Glacier for Archive Tier. Same approach is available using Azure resources.

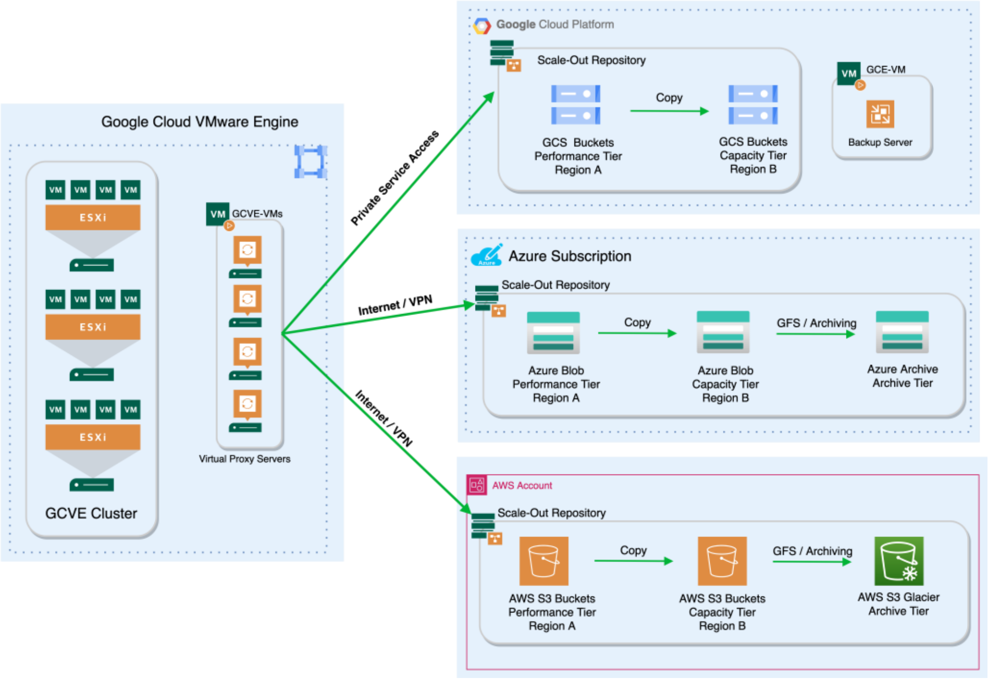

Google Cloud VMware Engine
VM VM VM VM
ESXi
GCVE Cluster
VM GCVE-VMs
Virtual Proxy Servers
Private Service Access
Internet / VPN
Internet / VPN
Google Cloud Platform
Scale-Out Repository
GCS Buckets Performance Tier Region A
Copy
GCS Buckets Capacity Tier Region B
VM GCE-VM
Backup Server
Azure Subscription
Scale-Out Repository
Azure Blob Performance Tier Region A
Copy
Azure Blob Capacity Tier Region B
GFS / Archiving
Azure Archive Archive Tier
AWS Account
Scale-Out Repository
AWS S3 Buckets Performance Tier Region A
Copy
AWS S3 Buckets Capacity Tier Region B
GFS / Archiving
AWS S3 Glacier Archive Tier

Image 03: SOBR using Object Storage buckets from different Cloud Providers

NOTE: When using Scale-Out Backup Repositories with Object Storage buckets only, it's entirely possible to create a SOBR with just Performance and Archive Tier, without using Capacity Tier. This will allow to move the GFS backups to a cheaper bucket, reducing the overall cost of the solution. Please be aware that this option won't allow to follow the 3-2-1 rule, as no extra copy of the data will be created.

Of course, it is always important to take in consideration the network connectivity and the available throughput, to make sure you can complete all backup jobs within the required backup window, and of course, to meet the RTO requirements. For instance, to connect with AWS or Azure resources from Veeam components in GCVE, the communication will go through Internet, and as an option a site-to-site VPN could be used.

In the other hand, to connect the Veeam components in GCVE with a GCS bucket, all you need is to set a **Private Service Access** between GCVE and the Google Cloud VPC. This access allows GCVE to connect with Google Cloud resources without using Internet, which can provide a better throughput.

IMPORTANT: It's important also to consider the costs of data egress when choosing one of these options. When using GCS buckets and a Private Service Access, all communications will use internal IP addresses, so just VM-VM egress and VM-to-Google services costs must be considered. In case of using AWS, Azure or any other Cloud vendor to provide an Object Storage bucket, you must consider the Internet Egress rates, which usually are quite higher.

Repository Design with Network Attached Storage

The third option available to provide repository capacity in Google Cloud to protect GCVE workloads with Veeam Backup & Replication, is using Network Attached Storage via NFS, in this case using Google Cloud Filestore.

NOTE: Filestore isn't the only option available in Google Cloud to provide a File Share. Another available option is using NetApp Cloud Volumes Service (CVS) for Google Cloud to provide NFS or SMB shares.

Filestore instances are fully managed file servers on Google Cloud that can be connected to Compute Engine VMs, GKE clusters, on-premises machines, and of course to GCVE workloads. Once provisioned, you can scale the capacity of your instances according to need without any downtime.

Important: Filestore uses the NFSv3 file system protocol on the Filestore instance and supports any NFSv3-compatible client.

Just like with Block storage and Object storage, it is entirely possible to configure a Scale-Out Backup Repository (SOBR) to use a Filestore instance (NFS share) as a Performance Tier, and then offload the backups to a Google Cloud Storage Bucket using Capacity Tier.

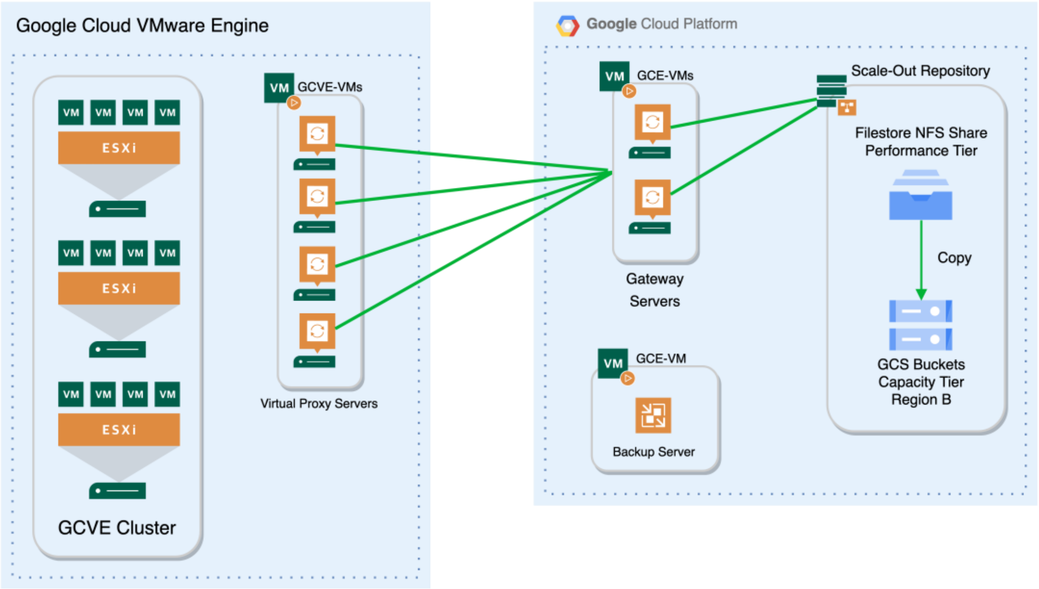


Image 04: SOBR with Block Storage using GCE Instances and Persistence Storage

By using SOBR with Capacity Tier, and the Copy option, it will be possible to follow the 3-2-1 rule recommended in the Veeam Best Practice Guide, to have multiple copies of your data.

NOTE: As mentioned before, it's important to highlight that for Capacity Tier we have more options than just using a Google Cloud Storage bucket. We could use AWS S3, Azure Blob or any other S3 compatible bucket. Of course, when using GCVE, for performance and costs reasons, the most recommended option is stick with GCS buckets.

Repository Design Recommendations

As already mentioned, it is possible to create a Scale-Out Backup Repository (SOBR) whether you are using block storage or object storage.

In a SOBR context, a possible design decision is the option of using **multiple** Object Storage buckets in Performance Tier. By using multiple Object Storage buckets, SOBR can distribute the backups among all these Buckets according to SOBR configuration (please consider that when using Object Storage for Performance Tier, the placement policy will be always Data Locality). This can provide potential performance improvements and can help to overcome Bucket capacity limitations from some vendors.

For GCVE, of course the first option will be leveraging Google Cloud resources for backup repositories. However, as already mentioned, Veeam Backup and Replication doesn't support using Google Cloud Archive Storage (only Standard and Nearline are supported), so we can only use Performance and Capacity Tier when using SOBR. The following image (Image 05) summarizes a SOBR design using Google Cloud Storage:

- One or more GCS buckets in a Google Cloud region for Performance Tier.
- One or more GCS buckets in a different Google Cloud region for Capacity Tier
- Copy mode enabled for Capacity Tier

This design allows to follow the 3-2-1 rule by having multiple copies of the data and having at least one copy in another location/region.

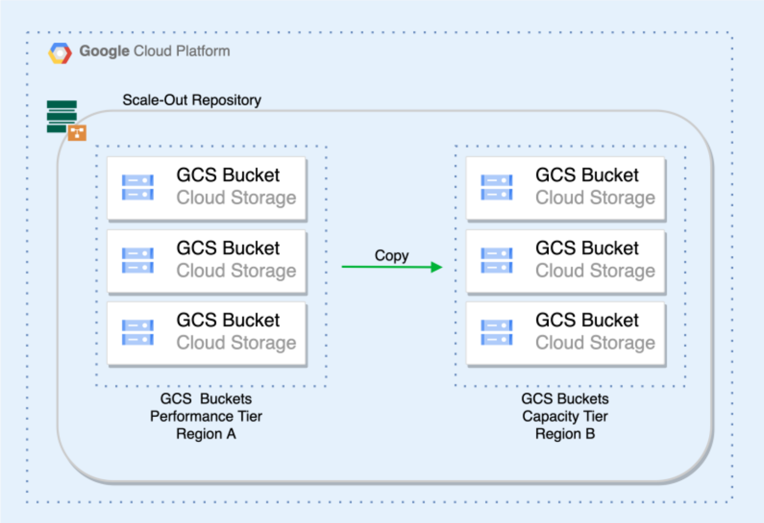


Image 05: Multiple buckets for Performance Tier and Multiple buckets for Capacity Tier – SOBR with Copy Mode.

Using other cloud providers, like AWS or Azure, is also an option we could consider to provide Object Storage Buckets. The design is not so different in compare with the previous example with Google Cloud resources, but with AWS and Azure we can also leverage Archive Tier to offload GFS backups to a cheaper Object Storage bucket for archiving purposes and saving storage costs.

NOTE: Additional connectivity considerations with other cloud vendors will be detailed in the next section about Object Storage Connection.

The following image (Image 06) summarizes a SOBR design using AWS S3 buckets:

- One or more AWS S3 buckets in an AWS region for Performance Tier.
- One or more AWS S3 buckets in a different AWS region for Capacity Tier.
- Copy mode enabled for Capacity Tier
- One AWS S3 Glacier for archiving purposes.

This design allows to follow the 3-2-1 rule by having multiple copies of the data and having at least one copy in another location/region. In addition, GFS backups can be offloaded to Archive Tier reducing overall storage costs.

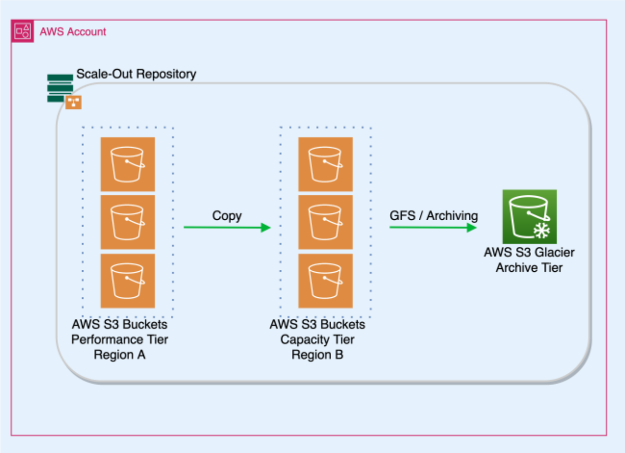

AWS Account
Scale-Out Repository
Copy
GFS / Archiving
AWS S3 Glacier Archive Tier
AWS S3 Buckets Performance Tier Region A
AWS S3 Buckets Capacity Tier Region B

Image 06: Multiple AWS S3 buckets for Performance Tier and Multiple buckets for Capacity Tier – SOBR with Copy Mode.

NOTE: When choosing a different Cloud Provider, like AWS or Azure, it is important to consider the costs of sending the data from Google Cloud to the selected provider (Data Egress). It's also important to consider the performance impact during backup and restore operations.

Sizing

For the sizing of Veeam Backup Repository and the Repository Server, there isn't any different recommendation for GCVE. You can follow the sizing recommendations for [Veeam Backup Repository] and [Repository Server] provided in the Veeam BP Guide, or you can even use the [Veeam Size Estimator (VSE)], where you can get the sizing for most of the Veeam Backup & Replication components.

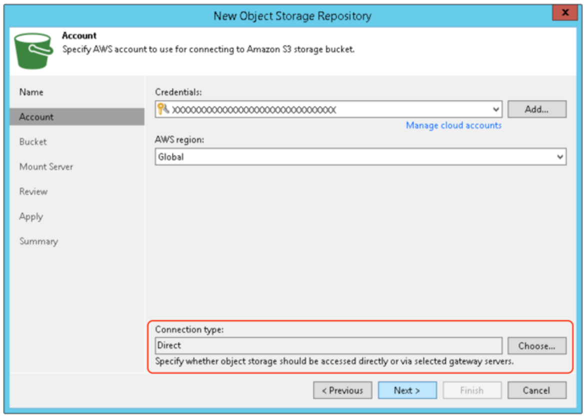
Object Storage Connection

When using Object Storage repositories, we need to specify how the Backup Proxy servers will transfer the data to the bucket. In v12 we have 2 options: **Direct connection** and **Through a Gateway Server**.

Direct Connection

With this option, the Backup Proxy servers are going to be responsible of transferring the backup data directly to the Object Storage repository, without requiring any other Veeam component.

This option is the most effective way in terms of **performance and scalability**. The main reason is that, when running a Backup Job with multiple VMs, it's very likely that multiple Proxy servers will be used to process those VMs. Therefore, each one of those multiple Proxy servers will also be used to transfer the backup data directly to the Object Storage repository,



Account
Specify AWS account to use for connecting to Amazon S3 storage bucket.

Credentials: XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX

AWS region: Global

Connection type: Direct

Specify whether object storage should be accessed directly or via selected gateway servers.

which increase the throughput by running multiple parallel tasks and multiple threads to connect with Object Storage repository.

The following are the pros of using Direct option:

- Scalability by using virtual Proxy servers. Multiples proxy servers can be used when running a Backup Job, allowing a higher scalability.
- High ingest rate by running multiple tasks in parallel during backup and restore operations.
- No bottleneck during backup and restore operations, as Veeam will use multiple Proxy servers to process the backups and to send and receive the backup data to the Object Storage repository.
- Simple design. Without the need of additional Veeam components, you can keep the design simpler.

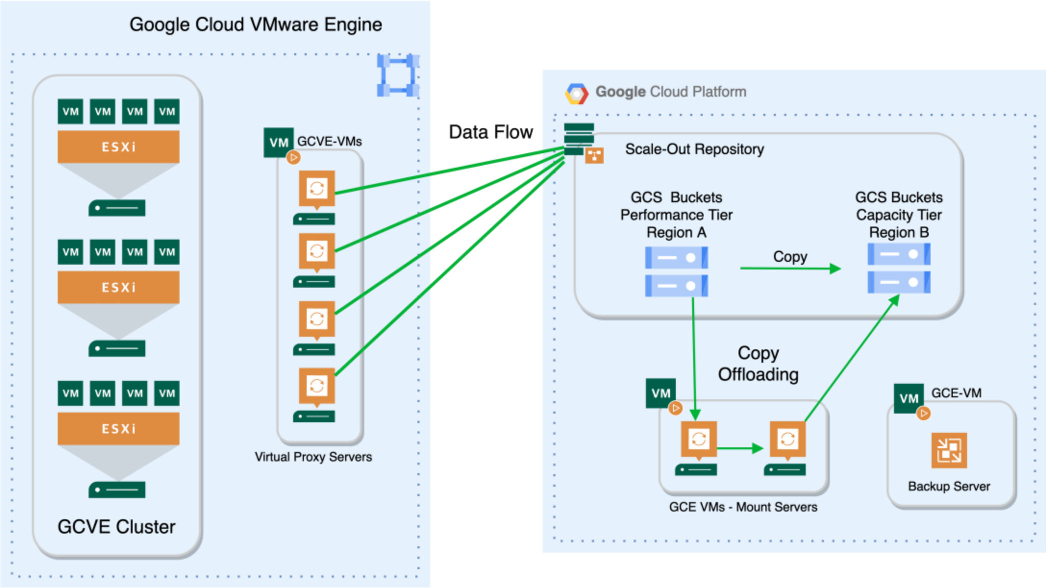


Image 07: SOBR with Capacity Tier and Direct connection in Google Cloud – Offloading using Mount Servers

But of course, we must consider that there are also some cons when using this option:

- By using this option, we are forced to use the **Mount server** associated with the source and target Object Backup repositories during the Offloading process. The Mount server itself could become a bottleneck as only one source Mount Server and one target Mount servers will be used to process the entire Offloading process, which could lead to performance issues.
- The Mount server must use Microsoft Windows, which could lead to extra licensing costs.

At this point is important to provide some recommendations when using Direct Connection:

- As already mentioned earlier, Proxy Servers should be as close as possible to the data to be protected, so in this case, it would be recommended to have the Proxy Servers running as Virtual Machines in GCVE (Image 07).

- [When using Google Cloud Storage] The Mount Servers are MS Windows VMs, and they could run either directly on GCVE or in Google Compute Engine (GCE). In both cases, because of the use of Private Service Access between GCVE and the rest of the Google Cloud infrastructure, the network throughput and costs should be similar (Image 07).
- [When using other Cloud Provider for Object Storage] The Mount Servers should be running in the same Cloud Provider (for instance, AWS) that we are using to provide the Object Storage Buckets. The main reason for this recommendation is to optimize the data traffic when SOBR offload the backup data from Performance Tier to Capacity Tier, as the Mount Server role is the responsible for this operation. As we can see in the image bellow (Image 08), using an AWS example, the Mount Servers are running as AWS EC2 instance, which makes the communication with source and target AWS S3 buckets more efficient, also reducing data transfer costs, as all the data transfer during SOBR offloading occurs within AWS network.
- Microsoft Windows licenses must also be considered for Mount Servers, whether they are running as GCVE VMs or as VM instances in the chosen Cloud Provider.

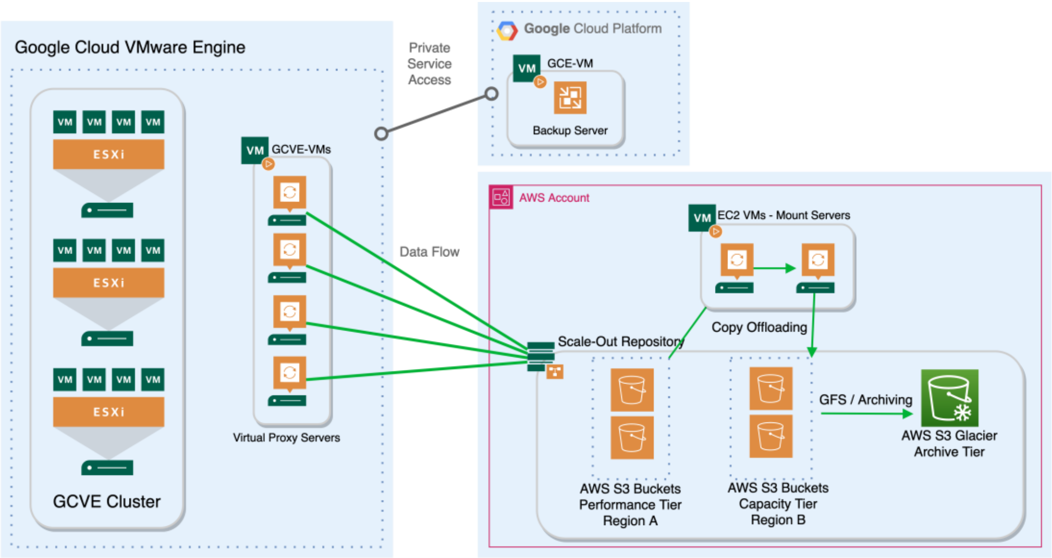


Image 08: SOBR with Capacity and Archive Tier using Direct connection to another Cloud Provider

Through gateway server

A Gateway Server is an auxiliary backup infrastructure component that "bridges" the backup server and backup repository. It can also "bridge" a source backup repository and a target backup repository in case of backup copy jobs or during offloading operations in a SOBR. The Gateway Server role, when used with NFS/SMB File Shares or Object Storage repositories, can be assigned to a Microsoft Windows or Linux machine added to the backup.

For the sizing of Gateway Server, there isn't any different recommendation for GCVE. You can follow the sizing recommendations for Repository Server provided in the Veeam BP Guide (as Gateway Servers and

Repository Servers require similar compute resources), or you can even use the Veeam Size Estimator (VSE), where you can get the sizing for most of the Veeam Backup & Replication components.

By using this option, we can choose one or more Gateway Servers to transfer the backup data to Object Storage. In this case, the data path would be (Image 09):

1. Proxy Server will get the data from production VMware cluster, process it (deduplication, compression and as an option also encryption), and then send the data to one of the selected Gateway Servers.
2. Gateway Server receive the data from the Proxy Server, and then send the data to the Object Storage repository (SOBR).
3. Gateway Servers also will be used for offloading backup data from Performance Tier to Capacity Tier according to SOBR configuration.

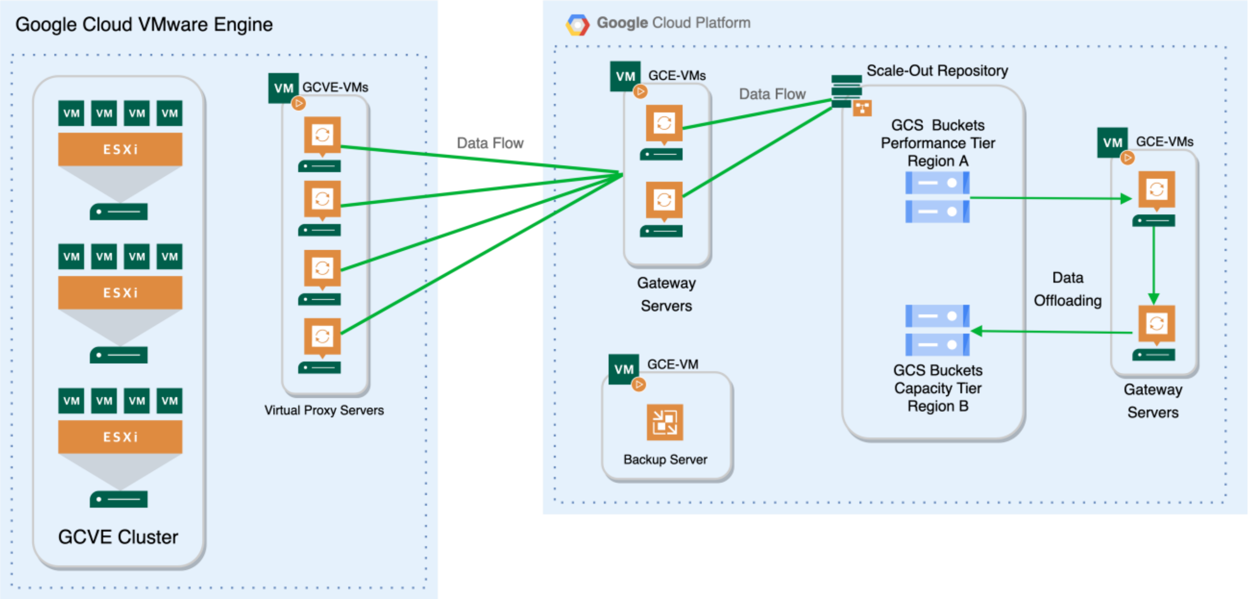


Image 09: Backup Job with Proxy Servers and Gateway Servers running in GCVE

Of course, using this option also has some pros and cons, just like the Direct connection option. The main benefit of using Gateway Servers, is that these can be used **during the Offloading process** (Image 09), avoiding the bottleneck when using Mount Servers with Direct connection, as multiple Gateway Servers can be used concurrently.

About the cons, we can mention the following:

- A more complex design in compare with Direct connection.
- Higher costs and footprint by having more Veeam components (proxy servers plus gateway servers).
- Scalability and performance may not be as good as Direct connection during backup operations.

We can also provide some recommendations when using Object Storage through Gateway Servers:

- To get proper performance and scalability, it's recommended to select multiple Gateway Servers per Object Storage repository, to prevent having a bottleneck, allowing the data transfer through multiple threads, and also to gain high availability for this component.
- Gateway Servers should be located as close to the backup repository as possible. Make sure that the Gateway Servers and Object Storage buckets are located in the same region, in the same Cloud Provider.
- [When using Google Cloud Storage] The Gateway Servers could run either directly on GCVE or in Google Compute Engine (GCE), as long as they are in the same region as the GCS bucket. In both cases, because of the use of Private Service Access between GCVE and the rest of the Google Cloud infrastructure, the network throughput and costs should be similar (Image 09).
- [When using other Cloud Provider for Object Storage] The Gateway Servers should be running in the same Cloud Provider (for instance, AWS) and the same region that we are using to provide the Object Storage buckets.
  - The main reason for this recommendation is to optimize the data traffic when SOBR offload the backup data from Performance Tier to Capacity Tier, as the Gateway Server role is the responsible for this operation.
  - As we can see in the image bellow (Image 10), using an AWS example, the Gateway Servers are running as AWS EC2 instance, which makes more efficient the communication with source and target AWS S3 buckets, also reducing data transfer costs, as all the data transfer during SOBR offloading occurs within AWS network.

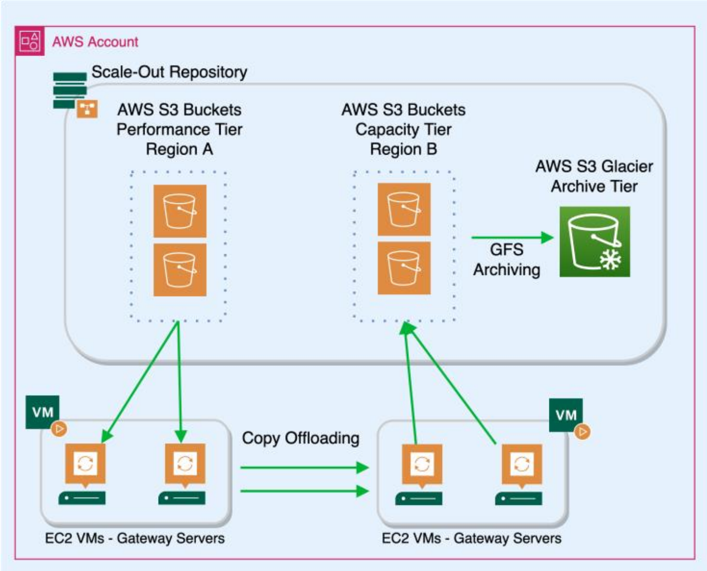


Image 10: SOBR Offloading process with Gateway Servers running in AWS EC2

Permissions

By default, GCVE retains full administrative access to your private cloud environment. When you create a private cloud, a default user, *CloudOwner@gve.local* is created in the vCenter Single Sign-On domain and given *Cloud-Owner-Role* access to manage objects in the private cloud. This default user is granted only with the sufficient administrator privileges to deploy and manage the virtual machines (VMs) in your environment.

To connect the Veeam Backup Server with the vCenter Server instance running in GCVE, we need a user account with proper privileges. The privileges granted to the default GCVE user account aren't enough, so we need to provision a proper account in vCenter Server with the required permissions so Veeam Backup Server can run all the tasks required during backup and restore operations.

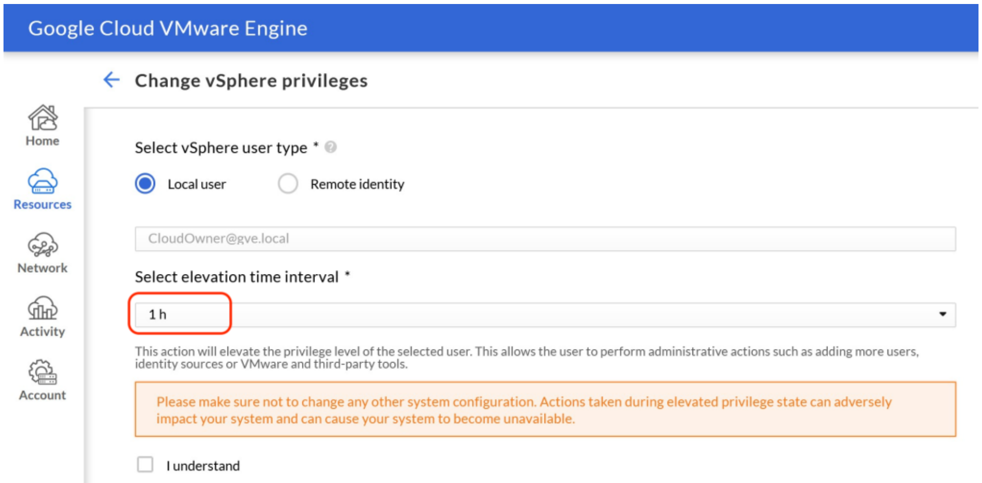
Elevate privileges

To create additional user accounts, manager vCenter Server permissions, and other tasks as adding additional Identity Sources in vCenter Server, you need proper privileges. As already mentioned before, those privileges aren't available by default in GCVE.

However, GCVE allows to temporary "Elevate" the VMware Engine privileges to perform some administrative tasks. Elevation of privileges basically means temporary adding the selected user to the vSphere built-in Administrators group.

As we can see in the image bellow, when elevating the GCVE privileges you can choose to elevate the privileges for the default GCVE local account (*CloudOwner@gve.local*), or for a user account from a remote Identity Source (like Active Directory) already configured in vCenter Server. In addition, you must select the elevation time interval from the list, as this elevation can't be permanent.

NOTE: Choose the shortest time interval that lets you complete the task. If additional time is required, you can extend the privilege elevation time interval afterwards.



Change vSphere privileges

Select vSphere user type *

● Local user ○ Remote identity

CloudOwner@gve.local

Select elevation time interval *

1 h

This action will elevate the privilege level of the selected user. This allows the user to perform administrative actions such as adding more users, identity sources or VMware and third-party tools.

Please make sure not to change any other system configuration. Actions taken during elevated privilege state can adversely impact your system and can cause your system to become unavailable.

□ I understand

Create a user account and role for Veeam Backup Server

Create a Service Account to provide proper access to vCenter Server from Veeam Backup Server. This account could be a local account in the vCenter Server SSO domain, or could be provided through an Active Directory domain or LDAP.

Regardless it is possible to provide Administrator privileges to the mentioned Service Account, for security reasons it's recommended to use the ***principle of least privilege***, providing the minimal permissions required by Veeam.

Therefore, the recommendation is to create a custom Role in vCenter Server, with the required permissions to allow Veeam Backup Server to run all the operations to backup and restore Virtual Machines. A full list of the permissions required by Veeam Backup Server can be found in the official documentation.

Other considerations

It's important to be aware that the use of Virtual Lab with NSX-T is not supported. This means that running SureBackup jobs won't be possible in GCVE as all the network stack is based on NSX-T.

A workaround can be found in the Veeam Community Forums, but you must remember that despite this workaround allows you to use SureBackup jobs with NSX-T, is still a not supported feature yet.

Disaster Recovery with GCVE and Veeam

Using Google Cloud VMware Engine and Veeam is also possible to design a Disaster Recovery strategy to protect the VMware Virtual Machines in case of a disaster, not just using Backups, but also using the replication capabilities available in Veeam Backup & Replication.

There are 2 main scenarios to use replication for GCVE virtual machines: Multiple GCVE Private Clouds in different regions, and a Hybrid Cloud with a GCVE Private Cloud and an on-premises VMware infrastructure.

NOTE: There are other scenarios which are out of the scope of this document like:

- Stretched GCVE Private Clouds
- Stretch on-premises Layer 2 networks to a private cloud using NSX-T
- Extending Networks Using VMware HCX

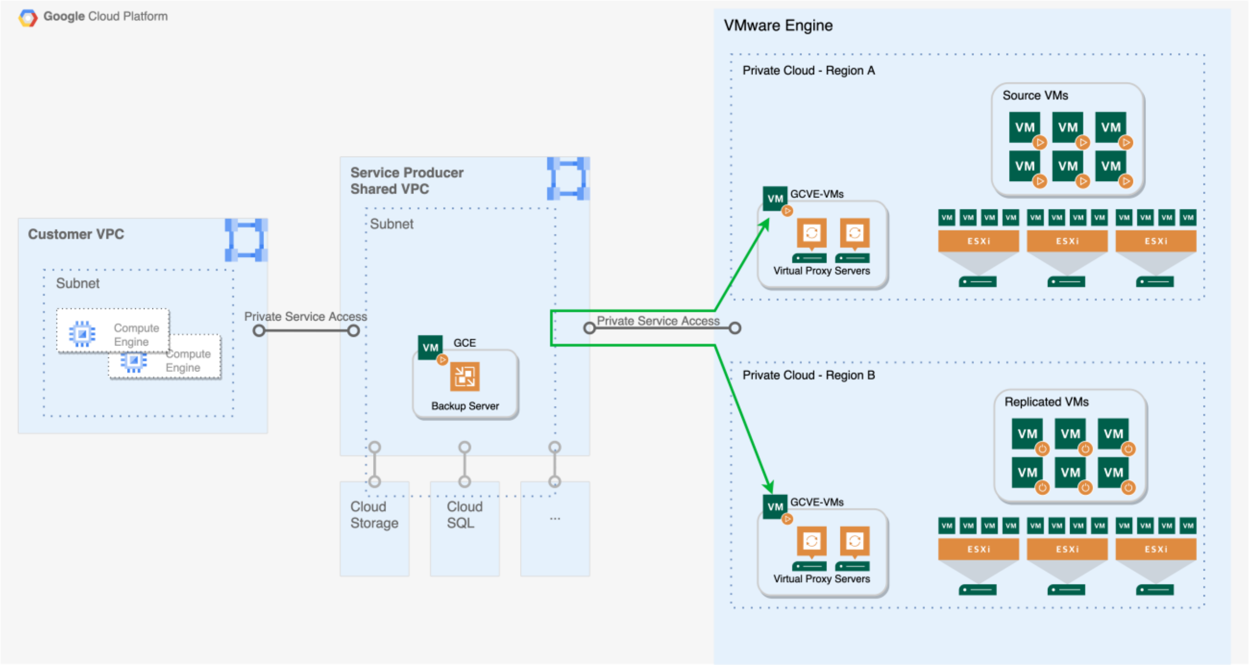
Multi-Region GCVE

In this scenario, we have at least 2 GCVE private clouds deployed in different Google Cloud regions. By having multiple private clouds in different regions allows for expansion, but also provides the resources required in a Disaster Recovery strategy.

As we can see in the image below, we have:

- 2 GCVE clusters in different Google Cloud regions.
- Both GCVE clusters are connected to the same service producer VPC network by using Private Service Access.
    - The use of Private Service Access allows each GCVE cluster to connect with Google Cloud resources without requiring internet access.
    - By having both GCVE clusters connected with the same Google Cloud VPC, also allows the communication between Virtual Machines running in both GCVE clusters.
- In each GCVE cluster we have Veeam Proxy Servers that are going to replicate the required virtual machines.
    - In Region A, Proxy Servers will get the virtual machine data from GCVE cluster.
    - Proxy Servers in Region A will send the virtual machine data to Proxy Servers in Region B.
    - Proxy Servers in Region B will use the data to create the replicated virtual machines in the target GCVE cluster.

NOTE: If private clouds are in different regions, then connectivity goes through the service producer VPC network, which is managed and owned by Google. This is the case as long as the **routing mode is set to global**.

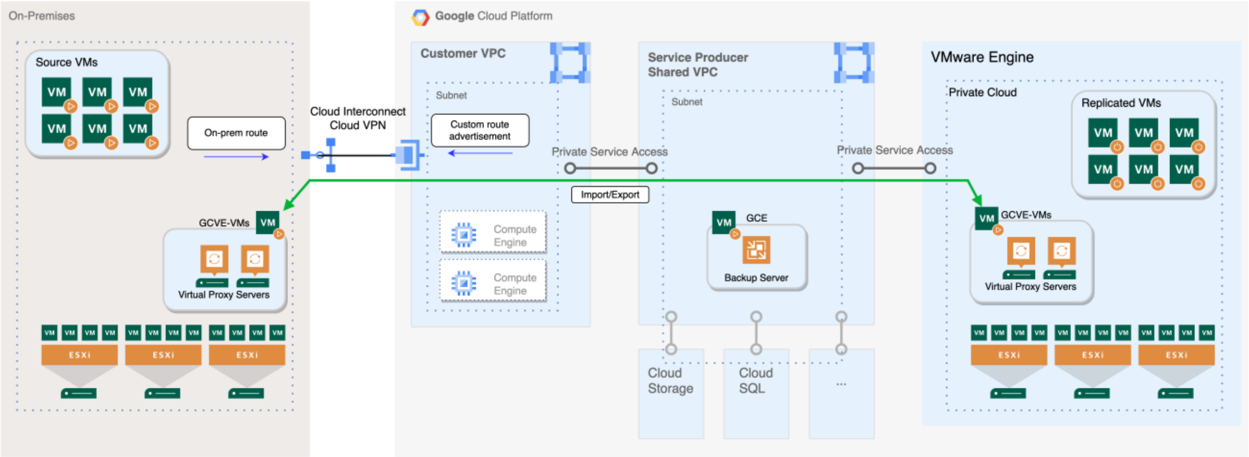


Hybrid solution

In this scenario, we have a GCVE Private Cloud, and an on-premises VMware infrastructure. Both sides are connected by using Google Cloud VPN or Google Cloud Interconnect, allowing to have a Hybrid solution. This hybrid infrastructure allows for expansion, but also provides the resources required in a Disaster Recovery strategy.

As we can see in the image below, we have

- One GCVE cluster.
- One or more on-premises VMware vSphere clusters.
- GCVE cluster is connected to a shared VPC by using Private Service Access. The use of Private Service Access allows the GCVE cluster to connect with Google Cloud resources without requiring internet access.
- Google Cloud VPN or Google Cloud Interconnect is used to connect the on-premises network with the required Google Cloud VPC.
- In the GCVE cluster and in the on-prem VMware infrastructure, we have Veeam Proxy Servers that are going to replicate the required virtual machines.
    - In the On-prem infrastructure, Proxy Servers will get the virtual machine data from VMware vSphere clusters.
    - On-premises Proxy Servers will send the virtual machine data to Proxy Servers in the GCVE cluster. The data goes through the Cloud Interconnect/VPN, the Google Cloud VPC, and then through the Private Service Access to reach the GCVE network.
    - Proxy Servers in GCVE Private Cloud will use the data to create the replicated virtual machines in the target GCVE cluster.



IMPORTANT: To allow a VMware Engine network to reach on-premises networks, you must enable **Import/export custom routes** on the VPC network peering connection associated with the private services access. This enables routes that are advertised from on-premises to the VPC network to be propagated to the VMware Engine region.

- When you use Cloud VPN for on-premises connection to the VPC network, you must add VMware Engine networks to the Cloud VPN tunnel.

- When you use Cloud Interconnect for on-premises connection to the VPC network, you can add custom routes to the Cloud Router that terminates the Cloud Interconnect attachment.

Conclusion

As described in this whitepaper, there are multiple design options to protect GCVE workloads with Veeam Backup and Replication, some of which were explained in detail.

Even when this document is focus on protecting Google Cloud VMware Engine workloads, it is important to highlight that with Veeam you have multiple options to be used as backup repository, not only provided by Google Cloud but also by other Cloud provider.   This provides freedom of choice and of course allows avoiding vendor lock-in.

The bottom line is that Veeam Backup & Replication provides you with simple, flexible, reliable, and powerful data protection solution, to protect all your workloads both on-premises and in the cloud.   All of these with a software-defined, hardware-agnostic solution.

About the Author

Patricio Cerda is presently a Senior Solutions Architect based in Spain, specialized in microservices with focus on Kubernetes and Kasten.   He has more than 20 years of experience in the IT industry, with wide experience working with Veeam, VMware SDDC solutions and AWS before joining Veeam Software in 2021.

About Veeam

Veeam Software is a privately held, U.S. information technology company with a U.S. based leadership team.

Founded in 2006, we focused on simplifying backups for virtual machines. We quickly became the industry leader. Veeam continues to charge forward to innovate the industry so you can own, control and protect your data anywhere in the hybrid cloud.