Usage of Veeam for Backup, Replication and Restore of user workload(s)

Revision history

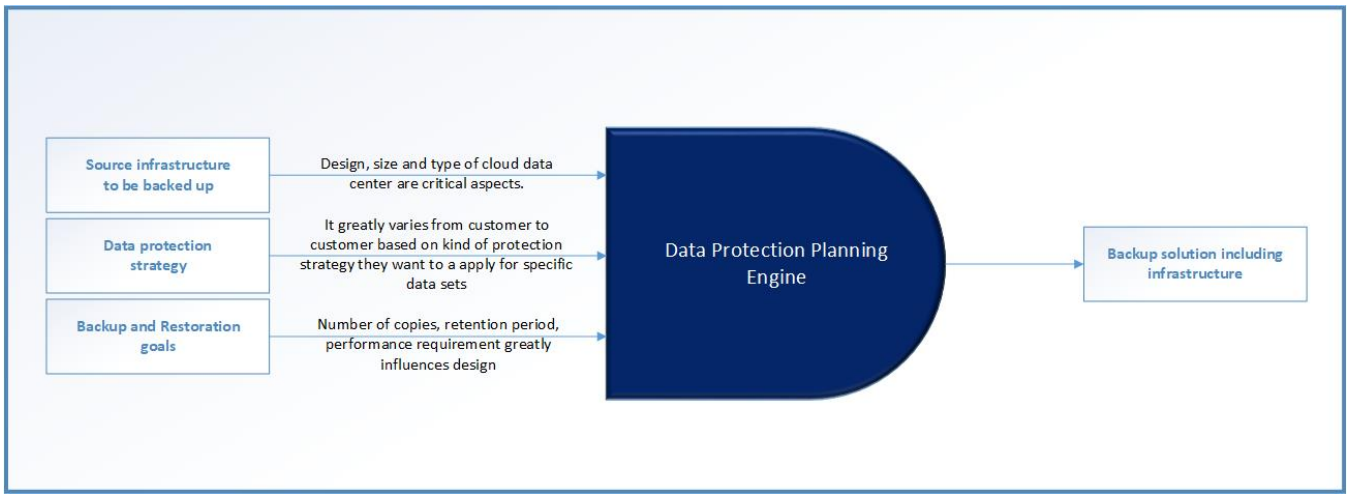
Revision	Author	Date	Description
0.1	Jyoti Ranjan	09-August-2019	Conceptualized the structure of document.
0.5	Jyoti Ranjan	09-August-2019	Copied the content was lying in my laptop but was not able to maintain it because of JIRA not being created
0.6	Jyoti Ranjan	09-August-2019	Updating integration aspects of Agena's with Veeam.
0.85	Jyoti Ranjan	09-August-2019	Added more diagrams.

Introduction

Agena VMaaS cloud platform aims to provide Infrastructure as a Service (IaaS) which has all benefits of cloud consumption with embedded HPE value add. It provides end user to provision virtual infrastructure like a cloud in a way which ahs tenets like robustness, agility, pay-as-you-drink, scalable on demand. For every cloud data center, the user workload is utmost important and so is ensuring its availability in case of data corruption or site failure. It is important to provide security of data is in motion or at rest. To achieves this, it requires comprehensive understanding and design thinking on data protection mechanism before concluding any backup, DR or data management solution. The backup, restore and replication is only one aspect of data protection system. In this document, we are not going to primarily focus on backup aspects and hence the reader is expected to ingest the content keeping this perspective unless it is stated explicitly.

The backup solution is highly influenced by customer's source infrastructure and backup requirements to be achieved. The following points greatly influence the design of Backup solution which is pictorially depicted as well:

1. Data protection strategy.
2. Capacity, consumption and design of source infrastructure to be backed up.
3. Backup and restoration goals



Considering the heterogeneity and diversity of data protection system as well as the need to focus on backup solution, it is very important to define what is covered and what is not covered as explicitly. See for details on it, The focus of this document is not to define complete data protection solution but to define a backup and restoration solution which can be used at the time of fail-over or disaster recovery.

Scope

The data protection strategy is bigger story than backup, replication and restore. In crude terms, the focus of backup is just to keep a offline cope of data which can be used at the time of data corruption or failure for restoration purposes. Many companies (like Veeam, Commvault) are really focusing on building products that offer “Resiliency” and “Business Continuity” which include both traditional backup and disaster recovery features. Also, many products may carry the same features but some will do it better than others. On the other hand, any data protection system is very much tied to customer requirement as well as source infrastructure, it is not possible to cover all permutation and combination of backup solution architecture without specific details with great degree of accuracy. So, it is very important to specifically detailed out what is covered and what is not in this document.

What is not in scope?

- No complete data protection strategy.
- No DR plan.

What is in scope?

- Understand backup and restoration solution provided by Veeam
- Defining backup and restoration solution for Agena cloud using Veeam
- As there is NO customer specific details are available, some assumptions have been made on various aspects. Those assumption greatly influence the architecture, design and configuration of backup and restoration solution. See below.
- Focus on following tenets (implicitly if not stated explicitly):
    - Robustness of backup infrastructure to avoid losing data in case of failure
    - Scalable along with source infrastructure
    - Adherence to 3-2-1 rule (if applicable)
    - Minimal impact on user workload during backup and restore operation.
    - Low RPO and RTO time
    - Designed to strive for low cost

Assumption

Source infrastructure	• Hypervisor = ESXi • Maximum number of ESXi hosts = 24 • Maximum number of VMs = 2000 • Storage media used for VM = Nimble • Maximum amount of raw storage = 100 TB

Backup infrastructure	• Use Veeam product suite for backup solution • Backup storage media = Nimble
Backup and restoration goals	• Retention period = 7 days • Number of backups per VM = 3 • Incremental backup will be taken instead of full backup always to save space • Focus is on better performance of restoration than backup. • No strict compliance to 3-2-1 rule

Design aspect of 'Data Protection System' for user workload(s)

The steps listed are based on author's understanding of data management and protection. The author has tried to present information as succinct as possible. It is implicitly assumed that below steps might need refinement for specific customer based on their data protection strategy for different types of data sets.

- List source infrastructure details
  - Type of hypervisor
  - Composition (number of hosts)
  - Number of VMs
  - Storage media used for VMs
  - etc
- Define data protection strategy (source: Tom Petrocelli)
  - **Backup and recovery.** Goal is to safeguard data by taking backup of data to be used in case of data corruption or failure.
  - **Remote data movement.** The real-time or near-real-time moving of data to a location outside the primary storage system or to another facility to protect against physical damage to systems and buildings.
  - **Storage system security.** Security data in rest as well as in-motion.
  - **Data Lifecycle Management (DLM).** Tiring and accessibility of data based on its age.
  - **Information Lifecycle Management (ILM).** A comprehensive strategy for valuing, cataloging and protecting information assets. It is tied to regulatory compliance as well.
- Define backup and restoration goals
  - How much data loss can the business afford (RPO – Recovery Point Objective)? How many backups we do want to maintain?
  - How quickly do the applications need to be up and running (RTO- Recovery Time Objective)?
  - How long does the data need to be retained?
  - How do we want to simplify backup and restoration mechanism? For e.g. do we need 1-click restoration?
- Understand specific backup solution and its fitment to source infrastructure
  - Feature
  - PoC
  - Licensing strategy
  - Direction in which product is headed
- Customize above steps for specific customer requirement (optional)

Evaluation of Veeam

Veeam is Backup, Replication and Recovery software. Veeam does not install any agent on VM being backed up.

Veeam Backup & Replication performs backups at the image-level using APIs available from the underlying hypervisor. It has no direct visibility of the file structure after backup is finished. It is possible to Use File Level Recovery (FLR) wizard or Enterprise Manager to mount VMs from within a backup file and access/restore VM guest files. It provides self-service portal which can be used to configure backup jobs and remotely manage it. It is highly recommend to design Backup Infrastructure keeping in mind the number of source infrastructure being backed up so that there is no scarcity of resources affecting running VM workloads at the time of backup or restore is being carried out.

Use cases

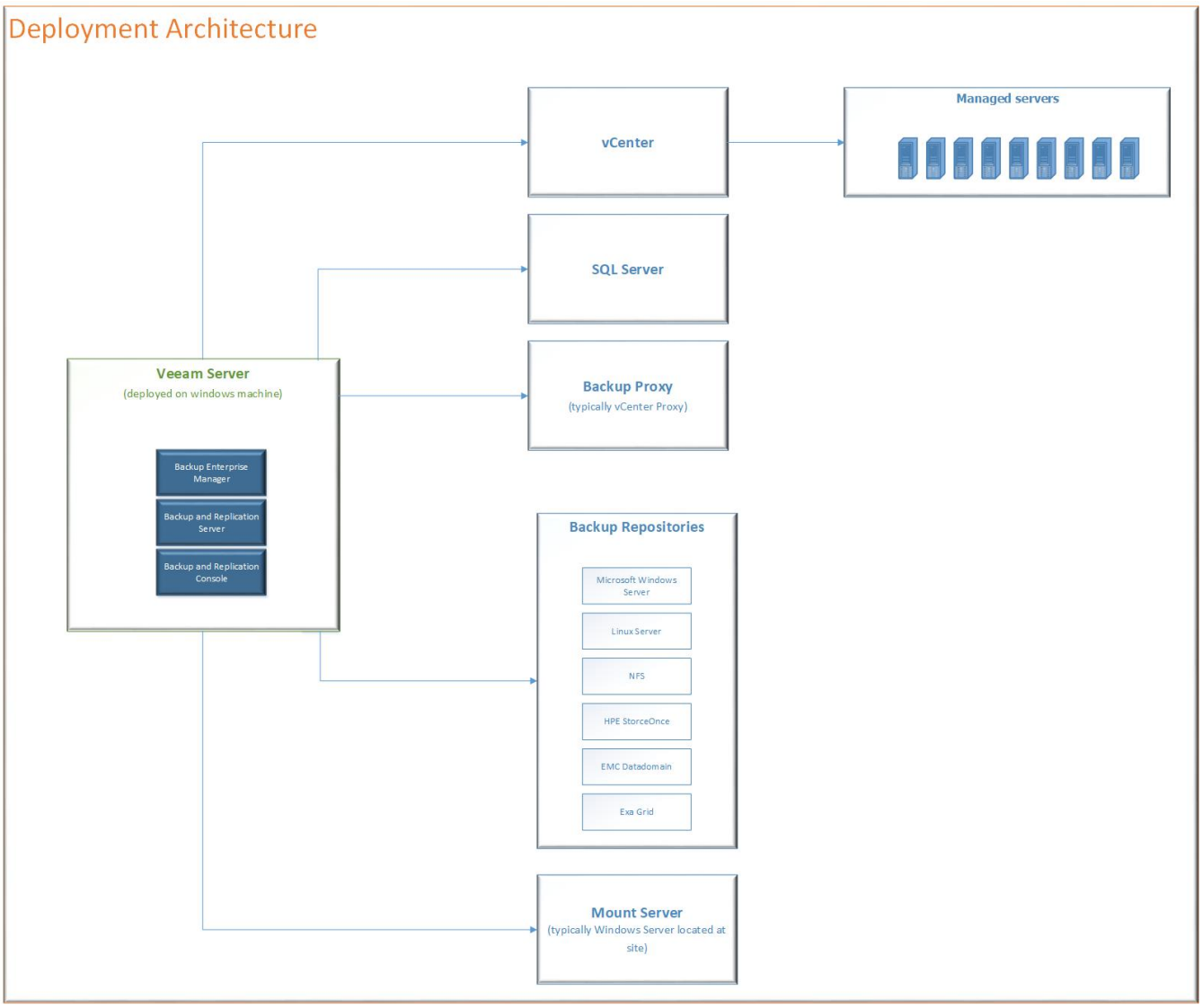
Veeam is a comprehensive backup and restore suite which allows integration with various third party component. The following use cases can be addressed.

1. VM use cases
   - Recover entire VM to the Original or Different host
   - Quickly restore to user by starting a VM directly from a backup file on regular backup storage

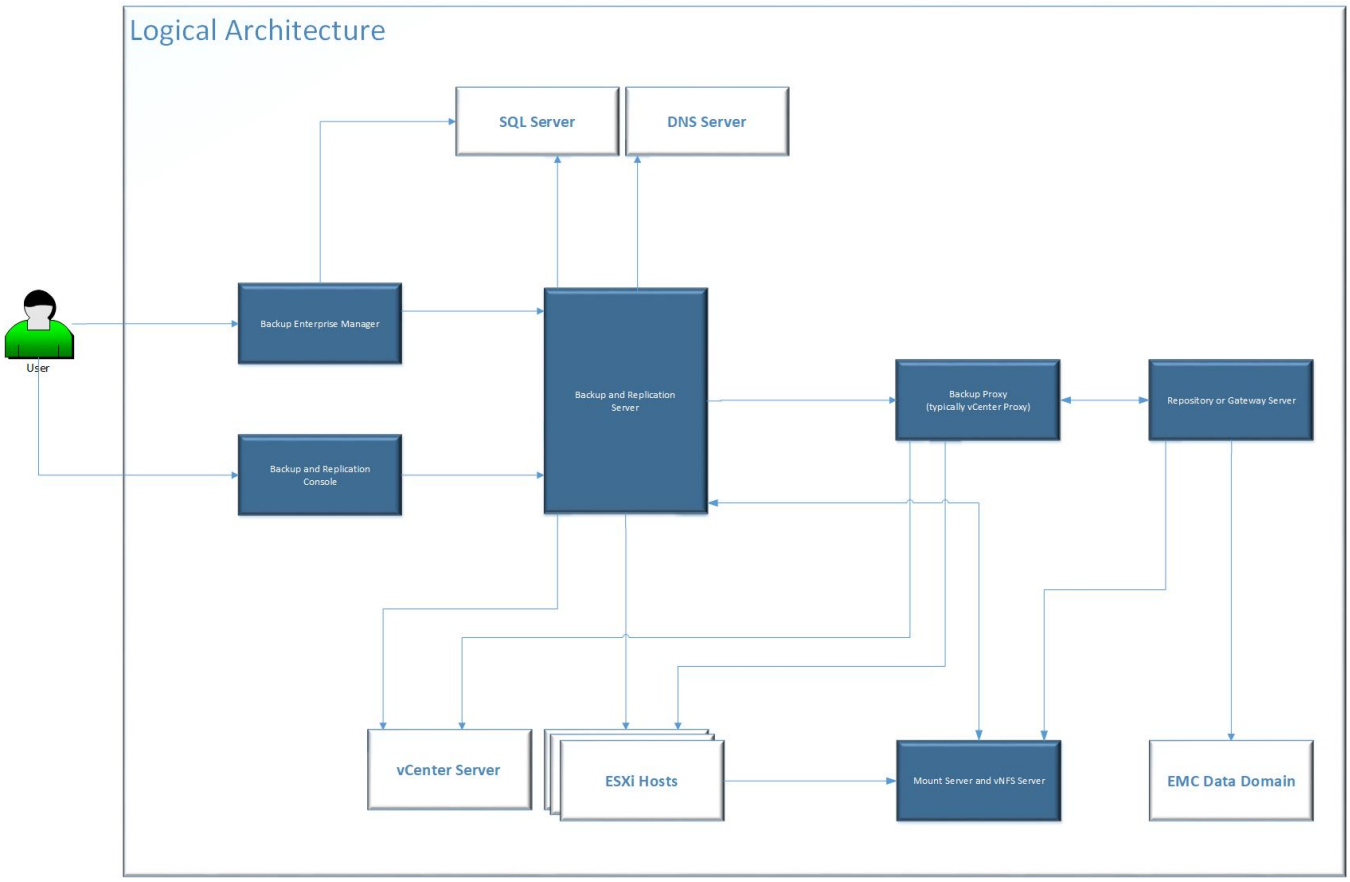
- Recover individual VM files (such as VMX) and virtual disks (vhd, vhdx or vmdk)
- Full VM recovery
- Instant VM recovery
- VM file and virtual disk recovery
- Restore to cloud

2. File level recovery
   - Recovery files from 19 common file systems used by Windows, Linux, Unix, MacOS, Novell, Solaris
3. Application aware backup (for specific enterprise application)
   - Microsoft Active Directory
   - Microsoft Exchange
   - Microsoft SQL
   - Oracle
   - Microsoft SharePoint
4. Replication
   - Replicate image
   - Use case: high availability or off-site for disaster recovery
   - Move production site to disaster recovery site
   - Mostly used for DR testing
   - Create backup from repositories without affecting workload
   - Facilitate data-center migration with zero-data loss
   - Get replicas offsite up to 50x faster and save bandwidth
   - Enterprise edition supports built-in WAN acceleration to Veeam Cloud Connect targets only
   - Enterprise plus edition supports built-in WAN acceleration to any target
   - Image based replication
   - Fail-over and failback
   - Replication from a backup
   - Planned fail-over
   - 1-click fail-over orchestration
   - Built-in WAN acceleration
5. Search and restore
   - User should be able to perform advanced search and restore

Deployment architecture (simplified view)



Logical Architecture

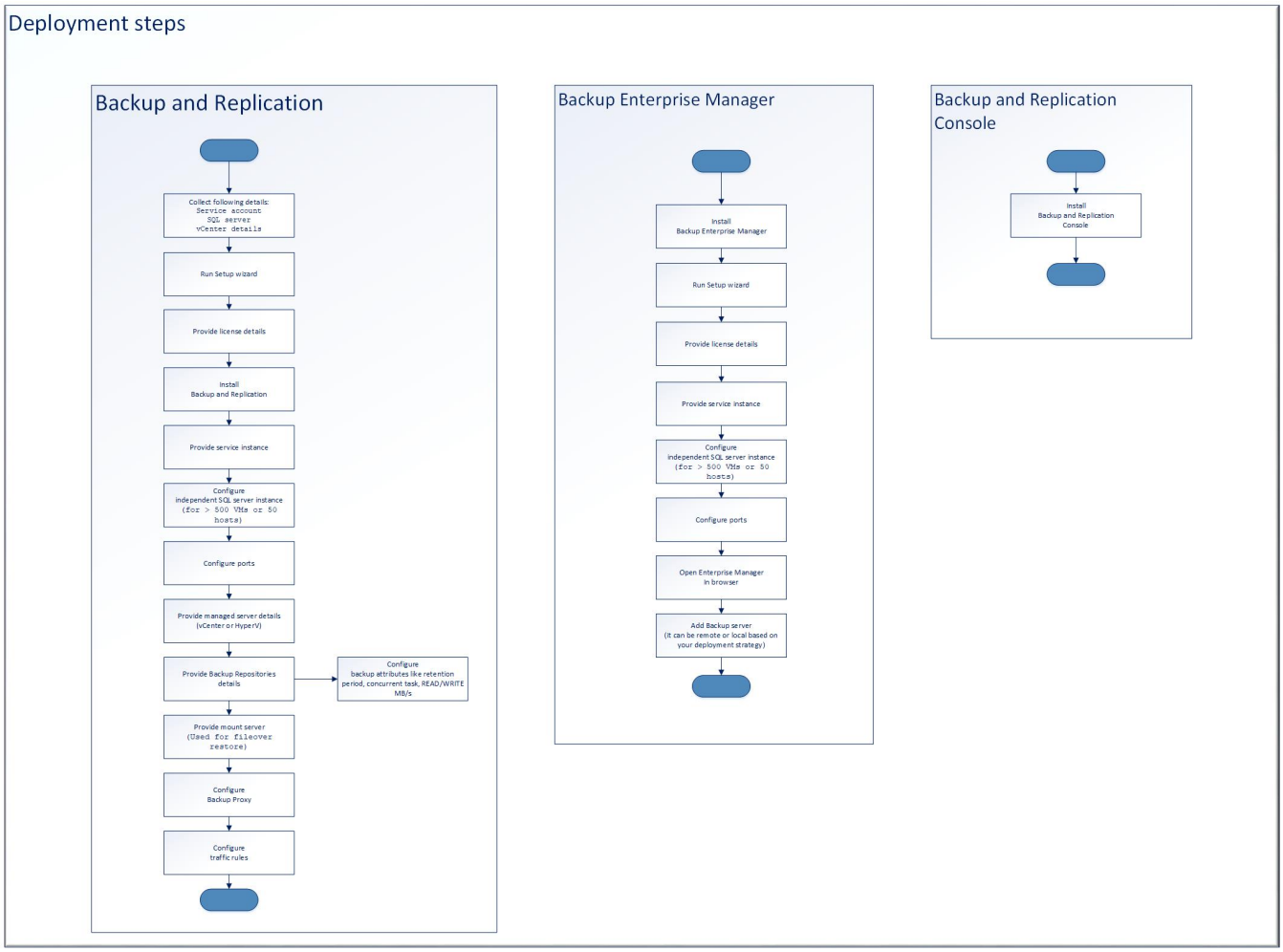


Deployment sequences

The below steps lists how we can deploy and configure Veeam solution for backup of user workloads running for source virtual infrastructure. It does not detail each and every steps very minutely as those can be found Veeam documentation. The purpose is to give coarse steps which can be exercised with the assistance of Veeam product suite which is self-explanatory up to great extent.

1. Preparation phase
    - Windows machine (VM or baremetal?)
    - Windows service account
    - SQL server instance
    - vCenter details
    - Backup solution details (like Microsoft Windows Server or HP StoreOnce)
    - Mount server details
2. Deployment and configuration phase for Veeam
    - SQL server instance
    - Port configuration for Catalog service port, Backup service port and Secure connection port
    - Provide managed servers details: ESXi or HyperV
    - Configure Backup repositories
        - Microsoft Windows Sever
        - Linux server which can Direct or internal attached server or NFC mount point
        - Shared folder (common for mediocre use case)
        - De-duplicating storage appliance like EMC Data Domain, Exa Grid, HP StoreOnce etc
        - Concurrent task settings (a critical parameter based on your system configuration in terms of resources, number of disks)
        - Read and Write in MB/s
        - Provide repository name
        - Choose the type of repository
        - Provide backup configuration attributes
        - Provide mount server details
        - Remove default backup repository which is local folder in Veeam server
    - Start the setup wizard (like windows installer)
    - Provide following
    - Configure backup proxy
3. Configure jobs for VM backups as per the customer workload.

The same is pictorially represented below!



Pros and cons

The below section captures author review of Veeam subjected to limited time spent on this.

Pros	• Intuitive and self-service portal • Rich feature set • Very good integration with ESXi and HyperV • Supports public cloud integration • Good supportability for HPE Storage System like HPE Store Once, HPE Nimble.
Cons	• Does not support KVM virtual infrastructure • Scale and performance can be eyebrow raising if system has greater than 2000 VMs (to be discovered more) • Data optimization during backup and restoration is average. I feel that HPE SimpliVity can lead here significantly.

Integration of Agena with Veeam for user workload(s)

This section is very much focused on specific aspects of of usage of Veeam for backup solution of Agena cloud.

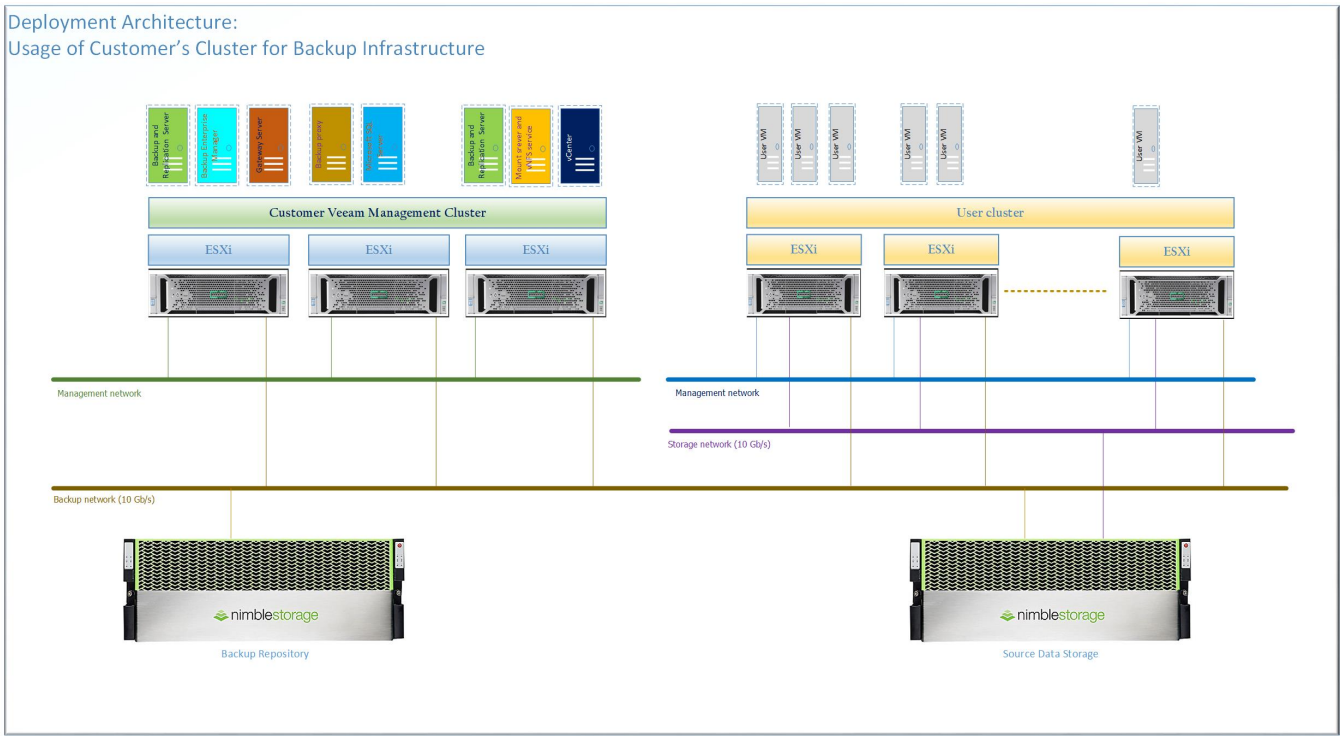
1. Deployment architecture
2. Deployment and configuration steps for backup operation
3. Recommended backup infrastructure

Deployment architecture

There are various ways we can deploy and manage Veeam product suite in association with GLHC components. The few of choices are:

1. Usage of 'Private Cloud Gateway' for backup infrastructure.
2. Usage of 'Agena cloud infrastructure' for Veeam infrastructure with the cross cluster management design.
3. Usage of 'Customer provided backup infrastructure'.

Author has been informed that customer is responsible for hosting Veeam components. So, only choice (3) is explored here. The below architecture diagram illustrates we customer leverages their Veeam component deployment in association with Agena cloud. Here we have assume Nimble as primary backup media (for illustration) so that we backup, replication and restore operation can be done with better speed.



Deployment and configuration steps for backup operation

It is assumed that customer is responsible for deploying backup infrastructure which is aware of source cloud infrastructure (i.e. Agena cloud). He or she can follow the steps mentioned above and configure jobs as per their

current user workload and application grouping.

Recommended backup infrastructure and its configuration

1. General recommendation
    - Define your data protection strategy
    - Category workloads in the terms of medium, high and critical workload
    - Try to adhere to 3-2-1 rule
2. Recommendation for Backup infrastructure
    - Do not deploy backup infrastructure in source cloud infrastructure being backed up
    - Deploy multiple Backup and Replication server if numbers of VMs are greater than 500
    - Use independent instance of Microsoft SQL Server
    - Use Nimble as primary backup storage. It will help to meet better RTO.
    - [Optional] For critical workload, use primary secondary storage of lower cost.
3. Backup job configuration
    - Number of backups
    - Retention period of 10 days
    - Choose your backup strategy. It is recommended to go for 'Forever Forward Incremental Backup'
    - Configure Job backup with following attributes:
    - [Optional] For backup of secondary backup, use backup copy job in the chain of backup operation.

More activities

As per author, the story US421 is necessary for but not sufficient for covering all aspects of providing advisory document for GLHC customer for backup and restoration process. There are other activities like PoC, scale and sizing, integration with public cloud (if desired) etc needs to be covered. Effectively, these can be translated into user stories for further work on it if we intend to provide refined document to customer. At this point of time, the following activities have been done:

- Evaluation of Veeam product suite
    - Feature
    - Logical architecture
    - Deployment architecture
- Define backup and replication architecture for Agena Cloud
- Document procedure to integrate Veeam with GLHC

It will be good to carry out more steps like PoC, Detailing scale and sizing aspect, Integration with HPE de-duplication solution HPE StoreOnce, Integration with public cloud (if desired) etc.

Summary

In a nutshell, Veeam provides reach feature suite for backup, replication and restoration of VM workloads. It has pretty well adaptability and supportability. Also, it can be integrated with different backup repositories like Windows Server, Linux Server, Nimble, EMC Data Domain, Exabyte, HPE StoreOnce etc. The self-service portal really makes admin life simplified. The only aspect which needs to be double clicked is of scale and performance. As per various source of documents, the number of VM counts greater than 2000 seems to be very large infrastructure to be managed for Veeam. And they strongly hint for laying out backup infrastructure and configure it to distribute the load optimally. Otherwise, there is very like hood of performance bottleneck which might affect running workload significantly. Author strongly recommend to carry out scale and performance test in an environment which is closer to user production deployment in terms of size as well as performance expectation.