Integration guide for CommVault's Backup solution

Revision history

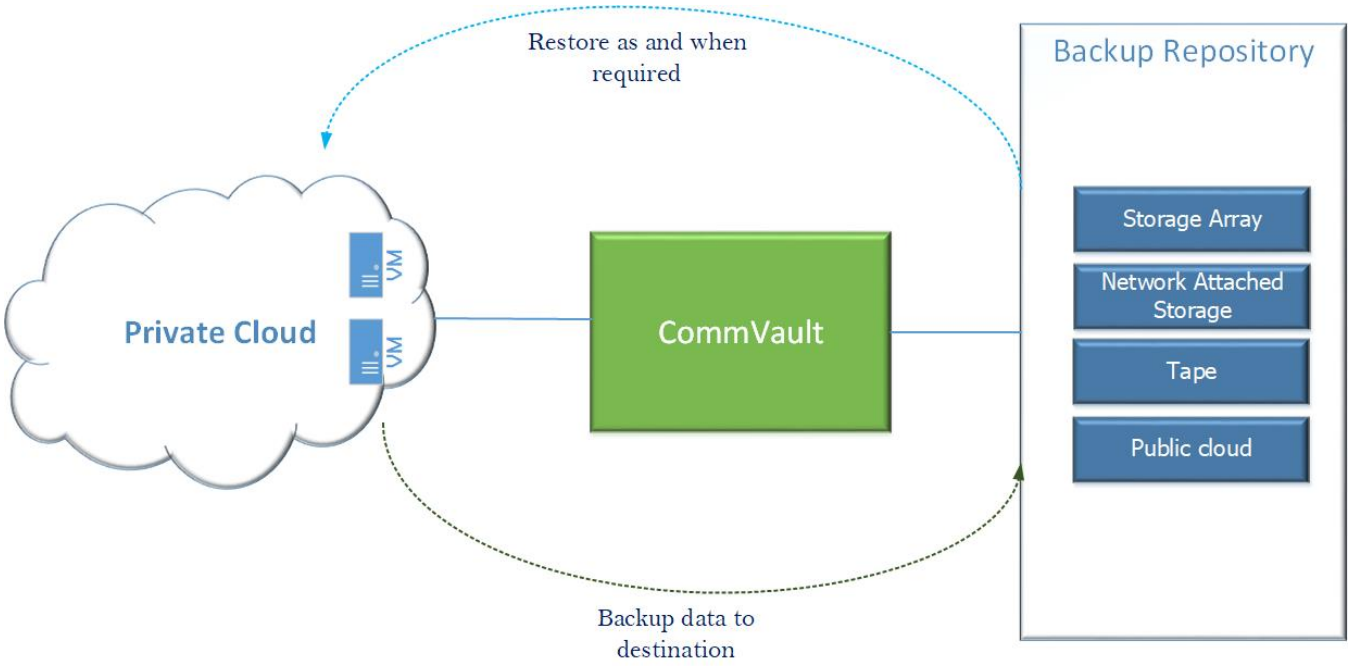
Revision	Author	Date	Description
0.85	Jyoti Ranjan	22-September-2019	Drafted the investigation guide for CommVault with most of content leveraged from Investigation report on Commvault for Manged Backup of Private Cloud

Introduction

GreenLake Hybrid Cloud (GLHC) cloud platform is a flagship managed cloud offering by HPE. Gemini is one of the product line up which aims to support **VMaaS** (VM as a Service) in private cloud space with the robustness, simplicity, scalability and pay-as-you go model. In other words, the goal of the “VMaaS” offering is to reduce friction between on-prem infrastructure and customers need of public cloud like experience using on-prem hosted service. The solution will allow customers to deploy various “services” (virtual machines, networks, storage etc.) through a self-service portal. The services will run on infrastructure residing on the customer premise or in a co-location facility. In order to simplify overall user experience, HPE will manage and operate certain aspects of the solution. Customers will no longer have to deal with the complexity of managing and integrating cloud services.

As part of managing end user's virtual infrastructure, it is paramount to have a data availability strategy for cloud as things fail in production environment. Backup is of the critical aspect of data protection strategy. It is important to perform backup of user's virtual infrastructure and ability to restore it to point in time when last backup was taken. Also, it needs to be ensured that application workload gets minimal impact when backup is being carried our and restoration of failed entities (VM or virtual disk or VM files) is carried out as soon as possible. One of the way to achieve is to integrate Agena cloud with third party backup solution like Veeam, CommVault etc. In this section, we go deeper into aspects needed to integrate CommVault Backup & Replication with Agena cloud without too much focus on specific of CommVault feature. If you are looking for CommVault features in more detail, please see CommVault documentation.

The below picture depicts very simplified view of connecting Private Cloud, CommVault Backup solution and storage entities used to host backup data.



Approach for backup solution belong to primarily two category:

- **Manged Service.** In this case, backup solution is offered by backup service provider and backup user interface is not exposed to end user
- **Self-managed Service approach.** In this case, backup solution runs like a service where API or endpoints are exposed to end user running workloads.

GreenLake will support first approach. Our goal is to leverage customer deployed backup infrastructure. The focus of section is on Managed Service mode but being managed by customer. In this case, we will not be exposing Commvault API or Commvault UI. All integration aspects of backup and restoration activities will be performed by HPE but bringing up and management of backup infrastructure will be done by HPE customers. Customer can configure same infrastructure among tenants and have unified jobs for backup scheduling. Or, customer can choose application category based backup infrastructure and different jobs for backup scheduling. In other words, customer is responsible for setting the backup policies according to the agreed service level agreement (SLA), and for performing all restore operations after receiving a request from the their end user or on failure of system. Rest of document provides more insight keeping this perspective in mind.

Assumption

Commvault integration significantly gets impacted with the size of infrastructure it has to deal for VM backup. Larger the infrastructure, the deployment architecture should have multiple instances of Veeam specific component. This document does not focus on specific sizing aspect but assumes following parameter as reference for rest of contents in document.

Private cloud 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 • Nimble has been used as primary storage media as well primary backup. iSCSI protocol is used to connect ESXi host with primary storage media.
Backup infrastructure	• Customer is responsible for backup job configuration instead of usage of backup service by direct tenant. • Storage array (e.g. HPE 3PAR or HPE Nimble) has been used as primary backup. Usage of storage instead of tape provides foster faster restoration process and is typically used for secondary backup now a days. Storage-level snapshots is used as VMware hypervisors are relieved of the resource usage due to long-lasting VM snapshots and their consolidation (delete) that occurs at the end of the backup operations. • The environment is configured to use only primary backup. No secondary backup is configured.

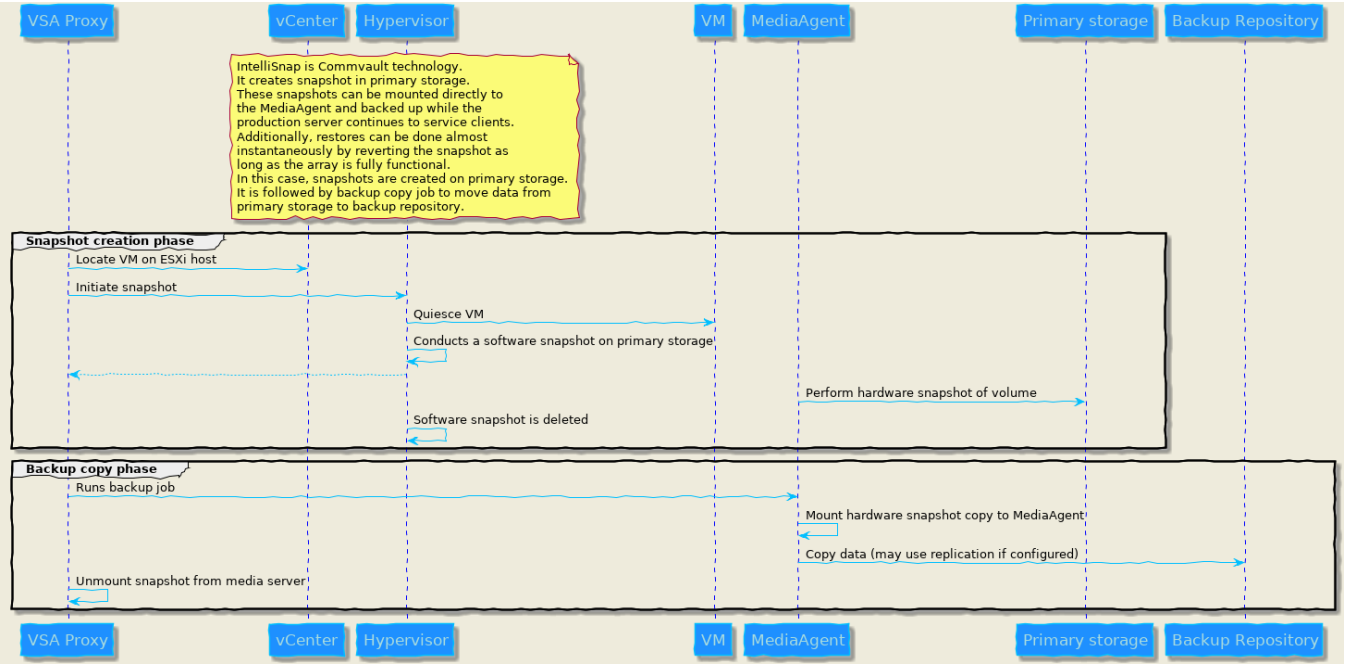
Supported use cases

CommVault provides a comprehensive backup solution. It supports multiple hypervisors, multiple backup media and public cloud. This document focuses on backup and restoration of VM deployed in private cloud only and does not address remote site backup, disaster recovery or backup to public cloud. As of now, the following use cases are intended to be supported:

- Backup of VMware objects
  - VM
  - Virtual disk
  - Guest files
- Live VM recovery
- Live Mount
- Live Browse or Live File Recovery
- VM recovery
  - Full VM
  - Virtual disk only
  - Guest files

Using IntelliSnap feature

IntelliSnap backup enables to create a point-in-time snapshot of the data used for backups. An effective way to back up live data is to quiesce it temporarily, take a snapshot, and then resume live operations. **IntelliSnap** backup works in conjunction with storage arrays to provide snapshot **functionality** for backup. It is very common method used for backup if storage array is used in VMware eco-system. IntelliSnap feature does require creation of hardware snapshot in primary storage. So, there must be provision for extra capacity to host snapshot. The IntelliSnap workflow looks like as depicted below.



Deployment architecture

CommVault requires its backup infrastructure (services, network, port) to be configured so that it can connect to vCenter to perform backup of VMs over host management network and store it to destined backup repository. In context of Agena, CommVault backup infrastructure can be possibly deployed in following ways:

1. **Organic deployment.** Customer deploys CommVault on infrastructure used to user workloads or VMs.
2. **Inorganic deployment.** Customer has pre-existing deployment of CommVault and he or she wants to integrate with Agena cloud.

In this document, the scenario (2) is explored, where customer is responsible for deploying and managing CommVault components.

There are plenty of permutation and combination in the way CommVault and Backup devices can be integrated which can vary based on:

- Source storage media used for VMware datastore
- Destination backup media
- Organization of VMware resources like VMs, virtual disk etc
- Backup policy
- Backup restoration need and restoration mechanism applied at the time of recovery
- Data protection strategy
- etc.

It is not possible to explore and illustrate all permutation and combination in this document. This document aim to illustrate how the most common use cases can be realized by following steps detailed below. For others, refer CommVault documentation.

Assumptions

1. Customer is responsible for bringing up backup infrastructure.
2. The environment is configured to use only primary backup. No secondary backup is configured.
3. Nimble has been used as primary storage media as well primary backup. Usage of storage instead of tape provides foster faster restoration process and is typically used for secondary backup now a days.

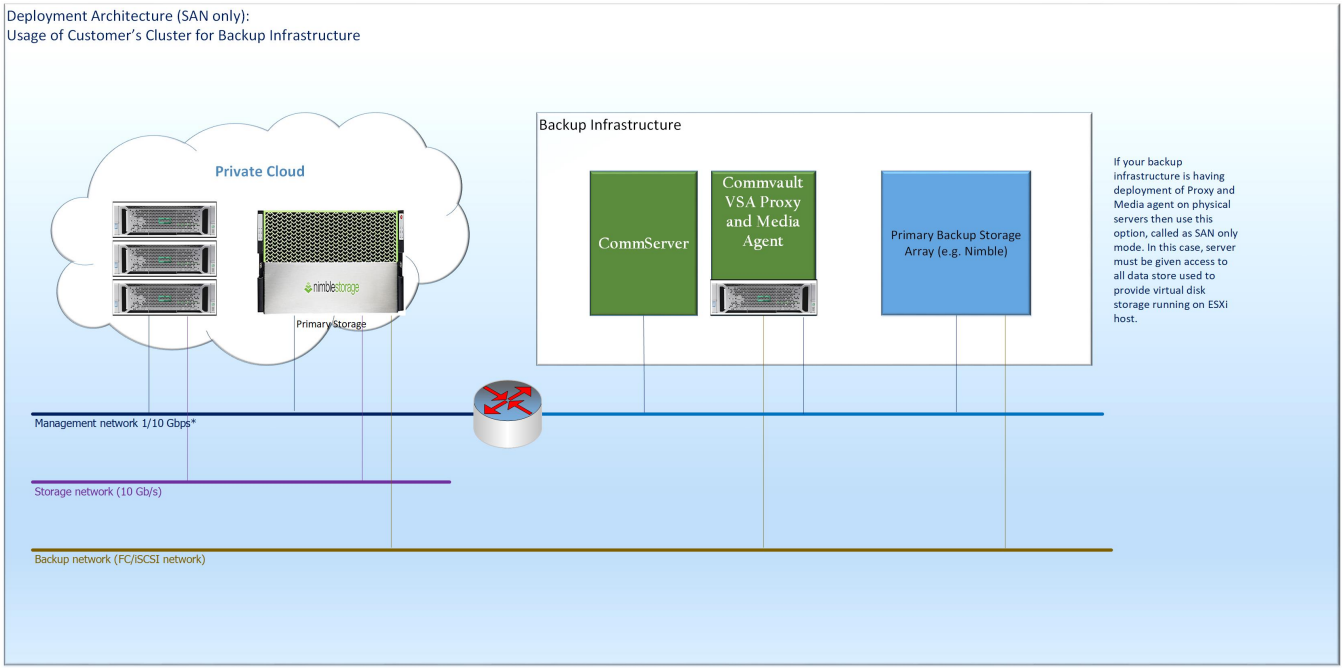
Deployment architecture choices

We strongly recommend to use Storage array for primary backup instead of tape etc though Agena cloud. Based on the awareness that Agena cloud uses Nimble storage array and we are going to adapt pre-deployed backup infrastructure by customer, these are following choices we have:

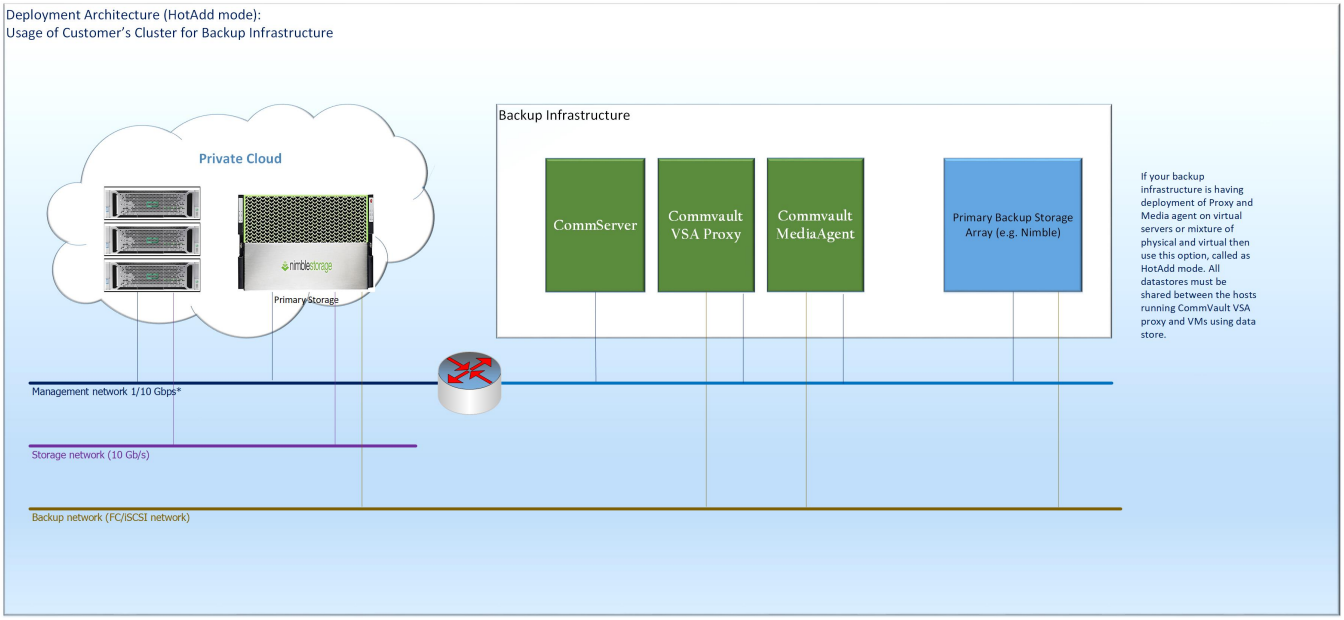
- SAN only mode
- HotAdd mode
- Network mode (should be used only if tape is used as primary backup)

The network diagram for each use cases are given below. But, it is worth to note that Agena cloud supports only SAN only mode or HotAdd mode because Network mode needs connectivity between ESXi hosts and VSA proxy.

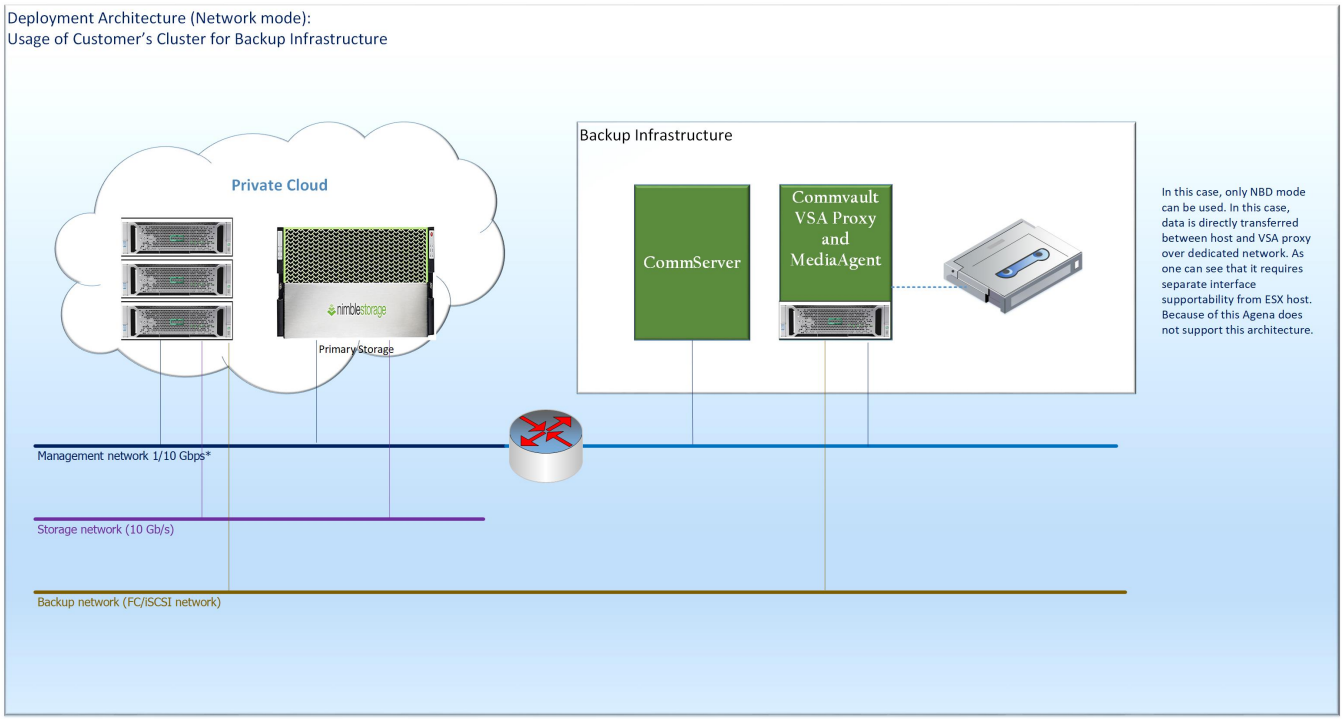
Deployment architecture using SAN only mode



Deployment architecture using HotAdd mode

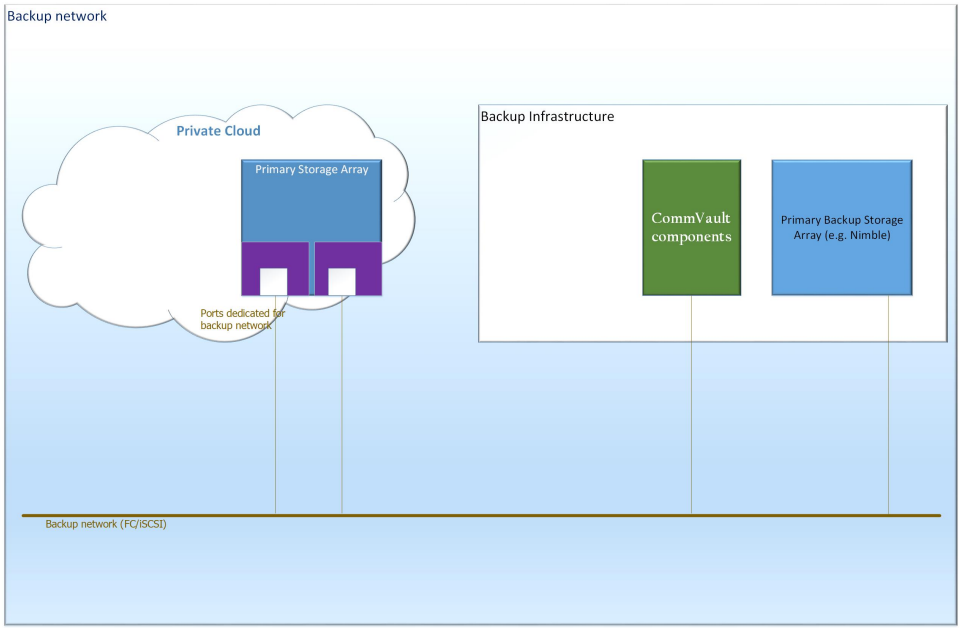


Deployment architecture using tape library



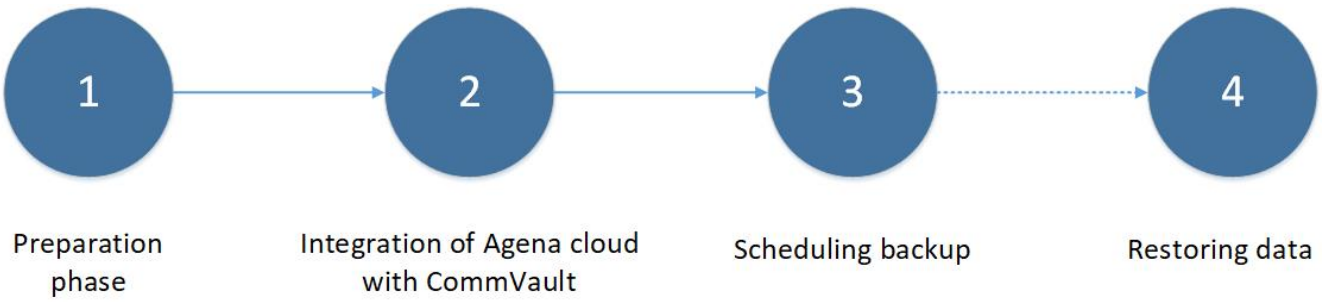
Network details

Network	Purpose	Integration specific information

Backup network	It is used to segregate data traffic from backup traffic.	Private cloud should facilitate an independent network termed as backup network) which facilitated FC or iSCSI connection am of SAN only and HotAdd mode. The Network mode is not supported. It is owned by customer. 
Management network	It is the data center management network used to connect vCenter with ESXi hosts.	As the expectation is to use pre-deployed backup infrastructure used by customer, a routing mechanism needs to be devised w management network. It is owned by customer. **Note:** We have not been able to find reference architecture where vCenter management network is not shared with Veeam bac two network should work fine. This piece needs to be tested in lab.
Storage network	It is used to connect ESXi host with primary storage.	It is owned by HPE.

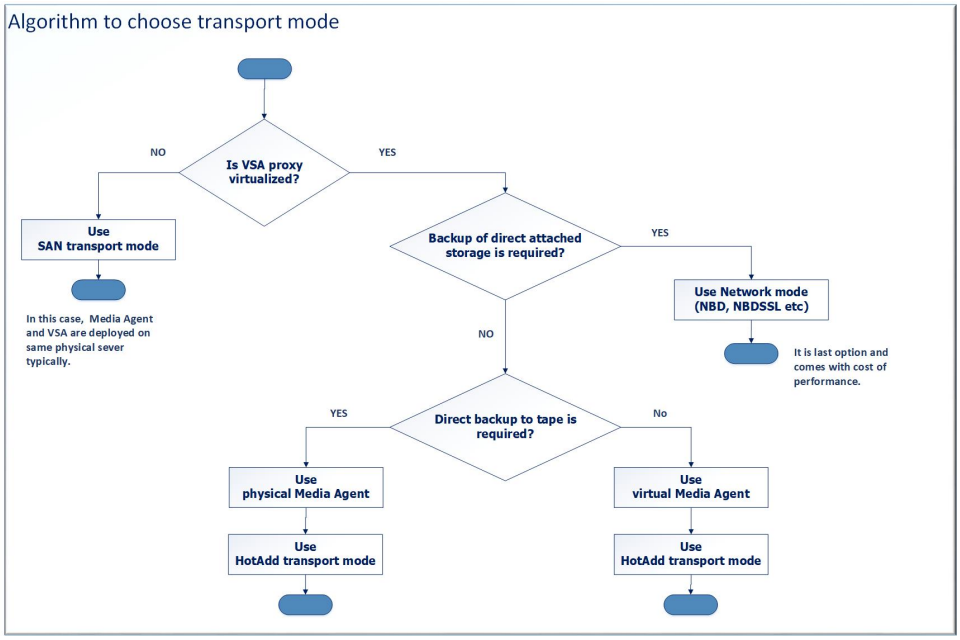
Integration steps

The integration activity can be categorized in four groups primarily as mentioned below.



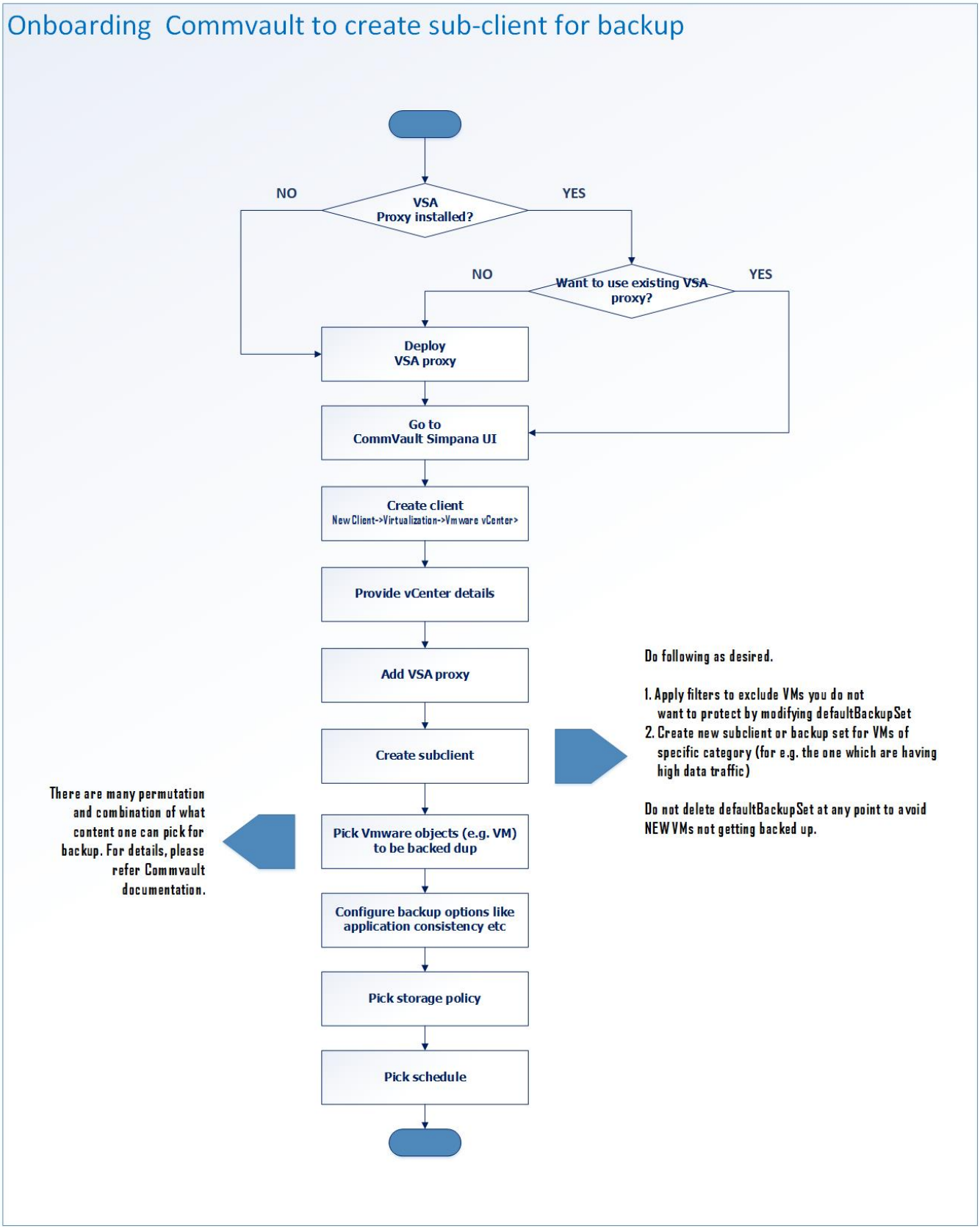
Preparation phase

In this phase, we primarily collect following sets of information which will be needed during integration phase.

Nimble storage array	The following set of information is needed only if we are using IntelliSnap feature of Nimble storage array. 1. iSCSI Discovery IP of the subnet that carries the data traffic. 2. User name and password (of administrator role or the power user role) 3. Snapshot properties • Snap Operation Retry Interval (in seconds) • Snap Operation Retry Count • Enable Diagnostic Logging • CHAP Username • CHAP Password
Backup repository	It is assumed that your backup infrastructure is pre-deployed and ready to be used. As we know that Nimble is being used as primary sto backup to your primary backup media in two ways: 1. Nimble storage replication, if your backup repository is also a Nimble storage array 2. Generic backup copy feature supported by Commvault For first case, you need to create the replication relationship between the local and the remote arrays, and set up replication parameters. Feature for more details.
Firewall settings	In an environment with firewalls, the vCenter, ESX servers, Virtual Server Agent, and MediaAgent must be able to communicate with eac the firewall, ensure that the ports for web services (default: 443) and TCP/IP (default: 902) are opened for communication on each of thes VSA with VMware for more details.
vCenter version	One must ensure that appropriate version of vCenter is used before exercising specific feature as mentioned in vCenter Server Versions
vCenter credentials	You can create a separate user account in vSphere for backup and restore operations. The user must have access rights as mentioned i
vSphere license	Depending upon the version of vCenter edition, you can use HotAdd transport mode. Please check your vCenter edition. If not then your
Choose transport mode	Based on your CommVault deployment architecture, you need to choose appropriate transport mode. It is recommended to use auto-sele technical reason, adhere to following decision flow. As Agena uses SAN array for VMDK datastore, one is expected to use SAN only or H backup, replication and restore operation greatly. Refer CommVault documentation for more details. 
Port requirement	• vCenter Port for web service (default: 443) must be opened. If vCenter is configured to use non-default ports, the non-default ports r • ESXi host Ports for web service (default: 443) and TCP/IP (default: 902) must be opened for the vStorage APIs for Data Protection. • vCloud Director (if used) port for the vCloud REST API (default: 443) must be opened

Integration of Agena cloud with CommVaullt

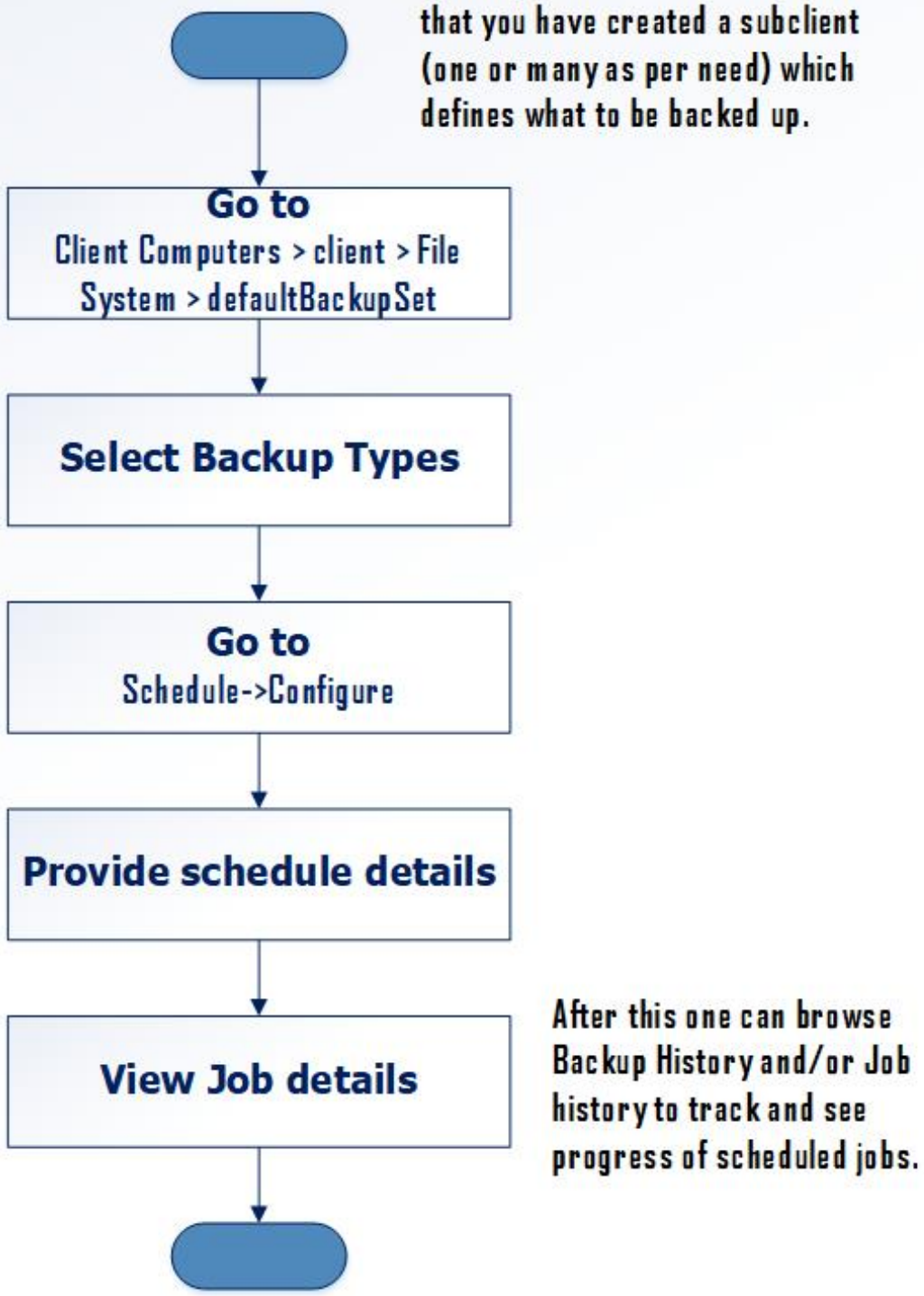
The below workflow tells how to create a sub-client, a term used to define what needs to be backed up. It is typically tied to storage policy. It is advised to have different sub-client for different category of workloads. By category, we mean by high traffic workload, critical workload, normal workload, application specific workload etc. For more details, see Commvault documentation.



Scheduling backup

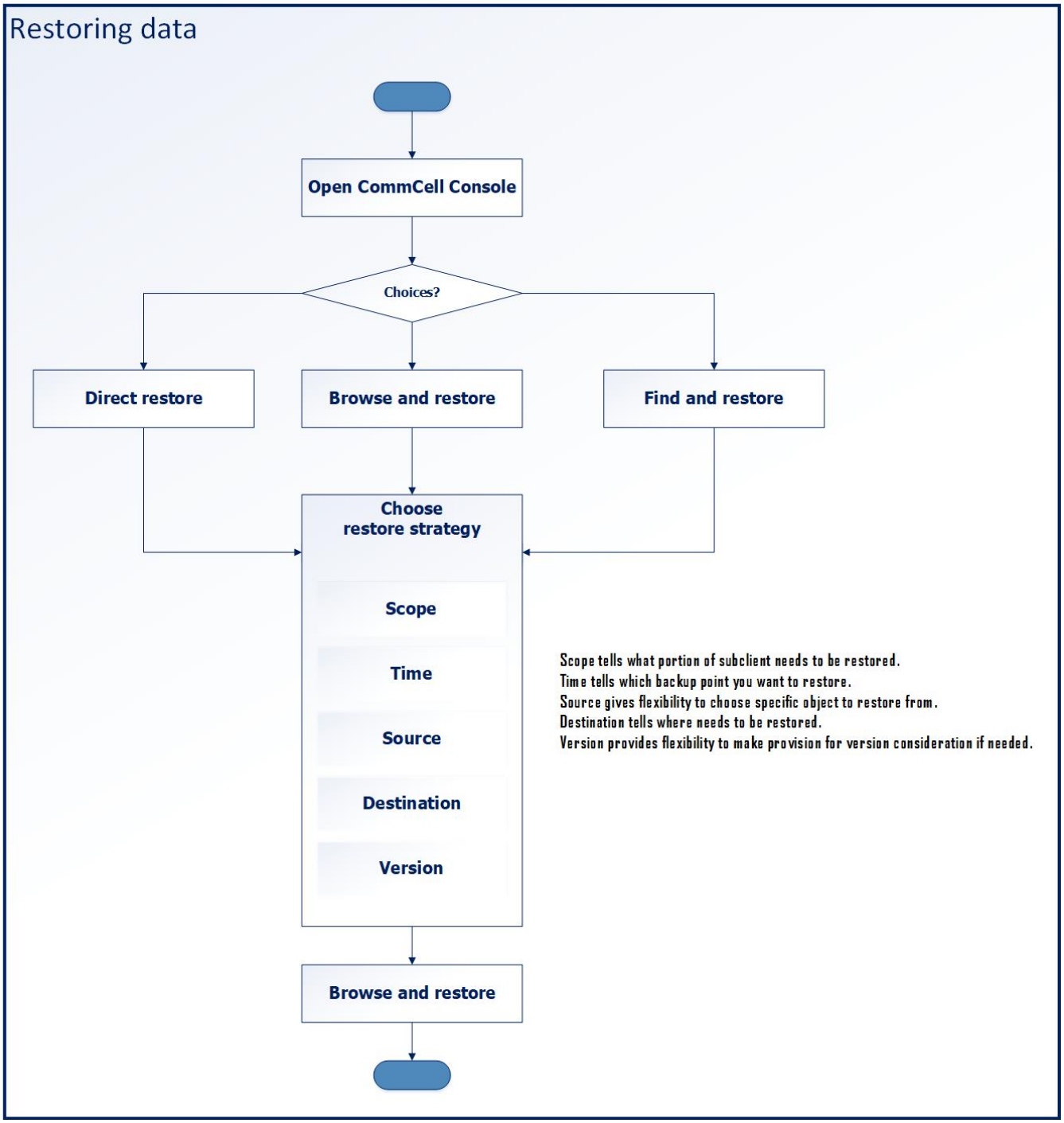
There are many ways to schedule a backup but the most commonly used workflow is depicted below. For others, see Commvault documentation.

Scheduling backup



Restoring data

There are many ways to take a backup but the most commonly used workflow is depicted below. Also, exact steps might vary based on kind of restoration we perform. For e.g. steps for instant recovery will different from normal steps to restore data. For others , see Commvault documentation.



Summary

Comvault Backup, Replication and Restoration provides a rich feature set to solve data availability and protection strategy for Private cloud of varying size. Based on size of private cloud, backup infrastructure for Veeam should be scaled for specific component to avoid sluggish performance. Integration of Veeam involves primarily following steps:

- Collect datasets needed to add and configure Veeam

- Adding Veeam
- Configuring backup job
- On demand, restoring data from backup repository

It is strongly recommended to use storage array as primary backup device and use IntelliSnap feature of CommVault for snapshot management as well as backup operation. Also, the selection of transport mode should be made based on deployment of backup infrastructure at customer site in following decreasing order of preference: SAN copy mode, HotAdd mode and Network mode.