Why and How to Use the Ceph Storage Appliance as the Repositories of Veeam Backup & Replication

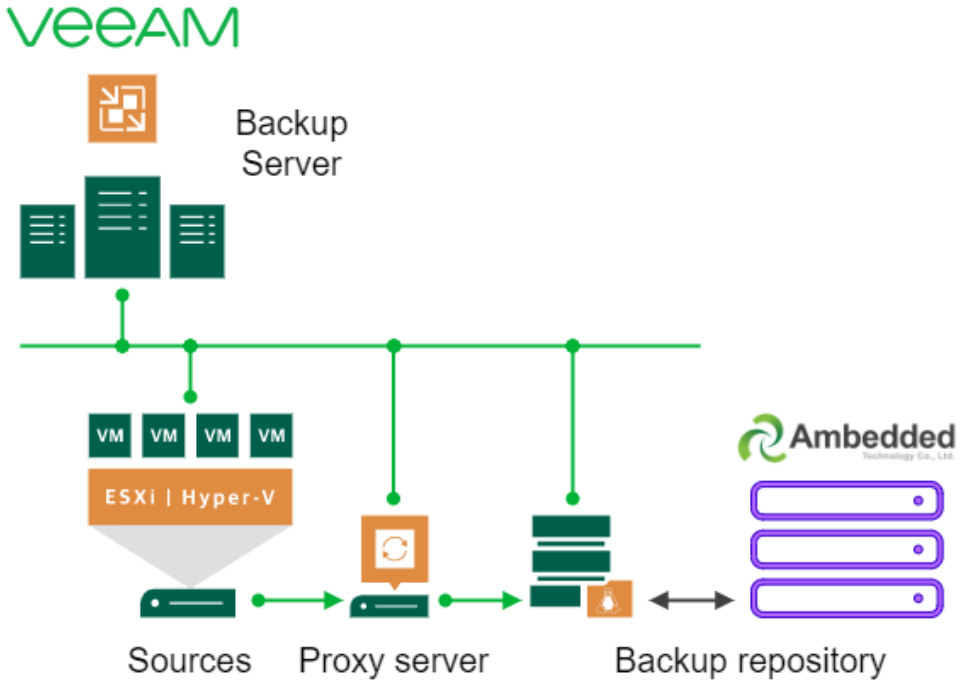
Cooper Su & Aaron Joue
Ambedded Technology, 2020 August

Introduction

This solution guide explains how to use the Ceph software-defined storage as the backup repositories of Veeam backup and replication.
Ceph is a scalable distributed software-defined storage. It features high availability, scale-out, and there is no single point of failure. Ceph supports object storage, block storage, and the POSIX file system all in one cluster. According to backup requirements, customers can select different storage protocols to support the needs of various storage backup strategies.

Ambedded Ceph Appliance Mars 400 is a turnkey solution for enterprise software defined storage based on Ceph. It integrates the Arm microserver, Ceph and the Ambedded Ceph management software UVS manager as a storage appliance.

Ambedded Mars 400 Ceph storage appliance brings the following benefits to customers.



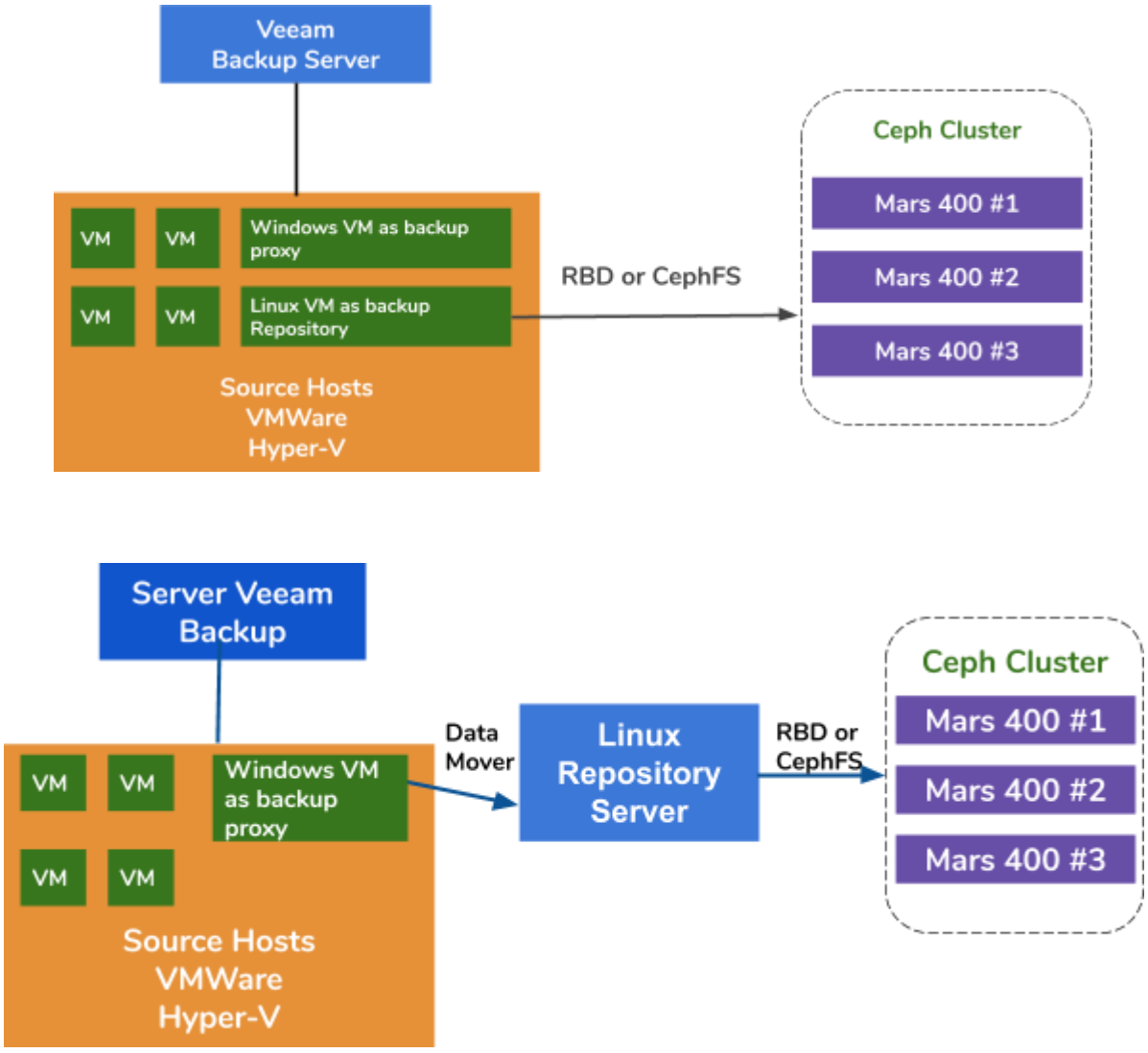
- **Capital Expenditure reduction**: Tuned Enterprise-class open-source software-defined storage appliance.
- **High Available:** Mars 400 cluster does not have a single point of failure and it has the smallest failure domain.
- **Unified Storage:** It supports S3 compatible object storage, shared file system, and virtual disk for local & remote site backup.
- **Unlimited scale-out**: You don't have to worry about forecasting the future capacity and throughput required. You can expand the scale on demand.
- **Operational cost reduction**: Self-healing & self-manage.
- **Web-based user interface**: Easy to manage the storage cluster.
- **Energy-saving**: 100 Watts Low power consumption

In this article, we use Ceph block storage and Ceph file system as the backup repositories and compare their backup job durations of backing up virtual machines from Hyper-V and VMWare.

Backup Solution Architecture

The architecture of backing up virtual machines on VMWare and Hyper-V are similar. Veeam uses data movers to transfer data from source hosts to backup repositories. The data movers run on the proxy server and the repository server. To use Ceph as the backend storage of a backup repository, you can mount RBD or CephFS on a Linux physical server or virtual machine as the **repository server**.

If the proxy and repository servers are virtual machines inside the hypervisor cluster, you can get the benefit of network-free high-speed data transporting between VM disk, proxy server, and the repository server. The best configuration of a large hypervisor cluster is to deploy one proxy server VM and one repository server VM on each VMWare host. Otherwise, you can deploy one backup proxy VM on every VMWare host and one off-host repository host to remove the workload from your production VMWare.



There are three ways to use the Ambedded Ceph appliance as the repositories for Veeam Backup and Replication. CephFS and RBD block devices can be used as the on-premises backup repository. The S3 object storage can be used as the capacity tier for a remote location.

In this article, we explain how to set up the Ceph RBD block device and the CephFS file system as the backup repository of Veeam for backing up virtual machines and files.

The test reports can be the references for choosing the Ceph RBD or CephFS for your backup.

Ceph uses multi-replication or erasure coding to prevent data lost against storage server failure. As the backup applications do not require high IOPS, erasure coding is the most cost-effective solution for data backup. Data stored in a replica 3 pool consumes 3 times of original data size in the storage. However, the erasure code pool (K=4, M=2) consumes

only 1.5 times size of its original size and keeps the same data durability. In our test, we use the erasure code k=4, m= 2 for the RBD, and CephFS data pools.

Test Environment

Ceph Cluster -

- Three Mars 400 with 3x monitors, 20 OSDs, and 1x MDS (metadata server)
- Each Ceph daemon runs on one dual-core Arm A72 microserver
- Operating System: CentOS 7
- Ceph software: Nautilus 14.2.9 Arm64
- Network: 4x 10Gb network per Mars 400

Veeam Backup & Replication 10, Version: 10.0.1.4854

Veeam Backup Server

- CPU: Intel Xeon E5-2630 2.3GHz DUAL
- DRAM: 64GB
- Network: 2x 10Gb sfp+ bonding
- Disk: 1TB for system, 256GB SATA3 SSD for volume
- Windows Server 2019

Veeam Proxy Server

- collocate with Veeam Backup Server

Repository Server

- Virtual Machine
    - CPU: 4 cores 2.3GHz
    - DRAM: 8GB
    - Network: bridge
    - Disk: 50GB virtual disk
    - OS: CentOS 7.8.2003

- Baremetal Server
    - CPU: Intel Xeon X5650 2.67GHz DUAL
    - DRAM: 48GB
    - Network: 2-port 10Gb sfp+ bonding
    - Disk: 1TB for system
    - OS: CentOS 7.6.1810

Hyper-V Host

- CPU: Intel Xeon E5-2630 2.3GHz DUAL

- DRAM: 64GB
- Network: 2-port 10Gb sfp+ bonding
- Disk: 1TB for system
- Windows Server 2019

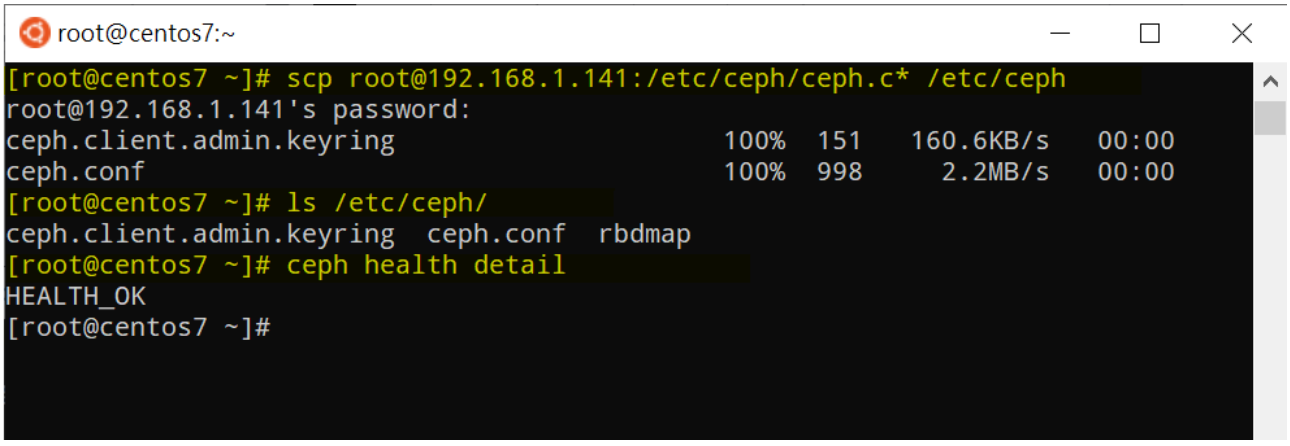
VMWare Host

- CPU: Intel Xeon E5-2630 2.3GHz DUAL
- DRAM: 64GB
- Network: 2-port 10Gb sfp+ bonding
- Disk: 1TB for system
- ESXi 6.5

Network: 10GbE switch

Implementation

Before deploying a repository server, please make sure your repository server has the Ceph configuration file and it can connect to the Ceph cluster. If you didn't have the ceph commands, please install the "ceph-common" or follow the tutorial.
https://docs.ceph.com/docs/nautilus/install/get-packages



Setup Ceph RBD as the Backup Repository

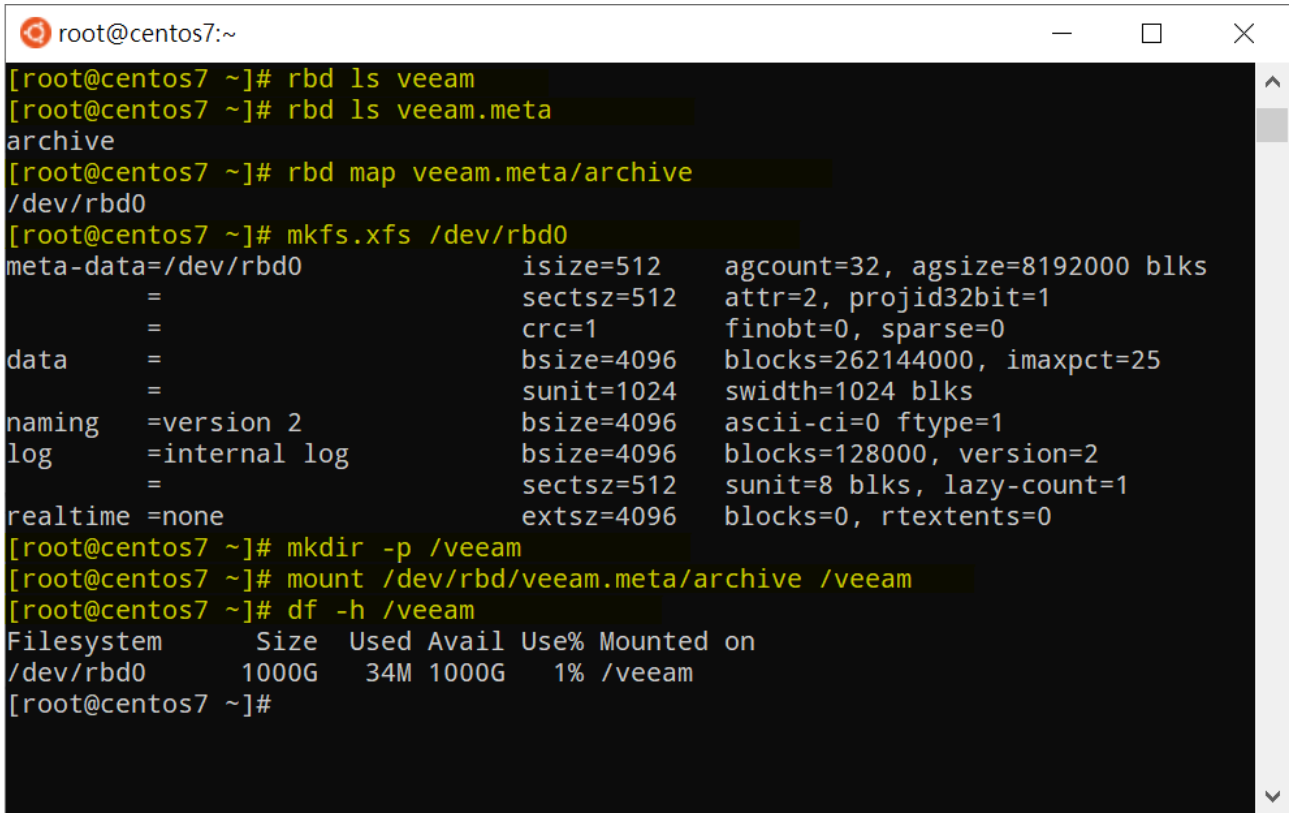
Ceph RBD Protocol (Block Device)

We use a Linux VM to mount the Ceph RBD to be a virtual block device and use this Linux VM as the Linux repository server.
Before creating the erasure code pool, we have created an erasure code profile which configures the data chunk and coding chunk as k+m=4+2. After creating the erasure code pool for RBD use, there are two pools created. The erasure code data pool is for storing image data and another replica 3 pool is used to store the metadata. Then we can create

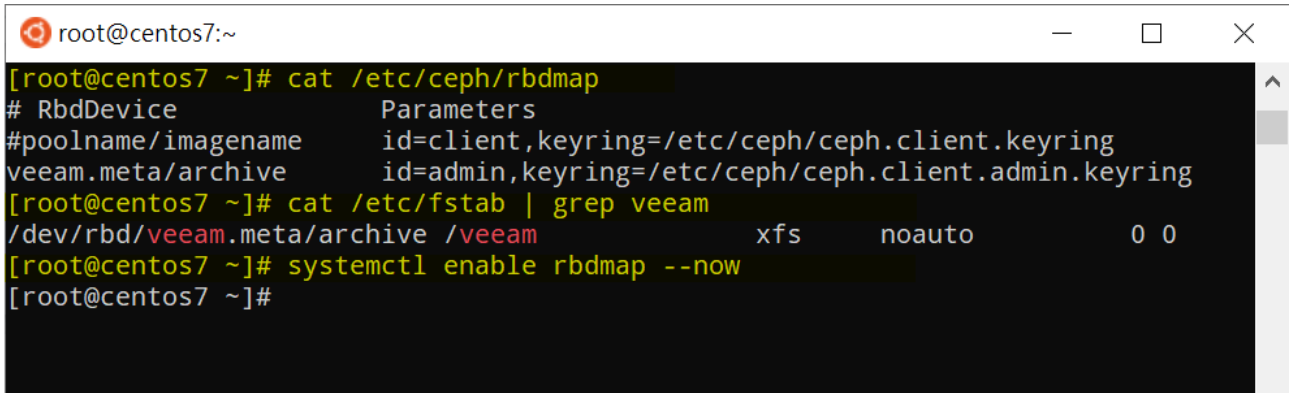
the RBD image by using the erasure code pool. All of the above procedures can be done easily by using the Ambedded UVS manager.

Now, we have to perform “map”, “format” & “mount” these three procedures before we can store data in the repository server.

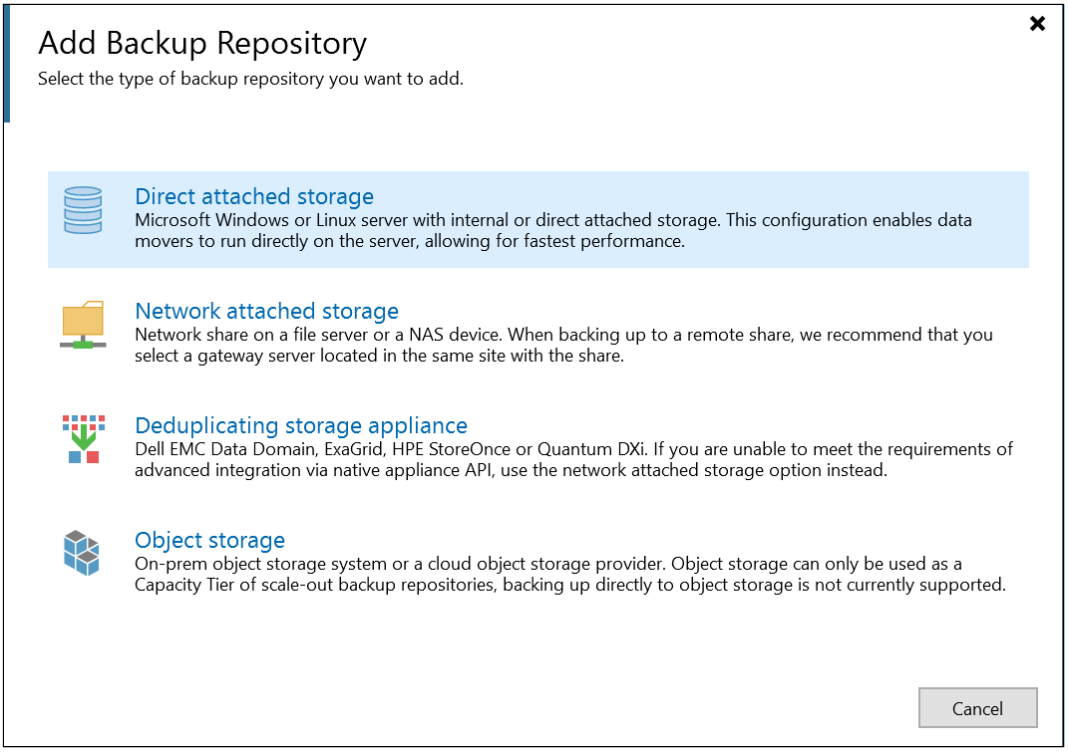
We shall use the metadata pool for the RBD images mapping. To mount RBD images automatically after the server boot, we could use the **rbdmap** service and **fstab**.

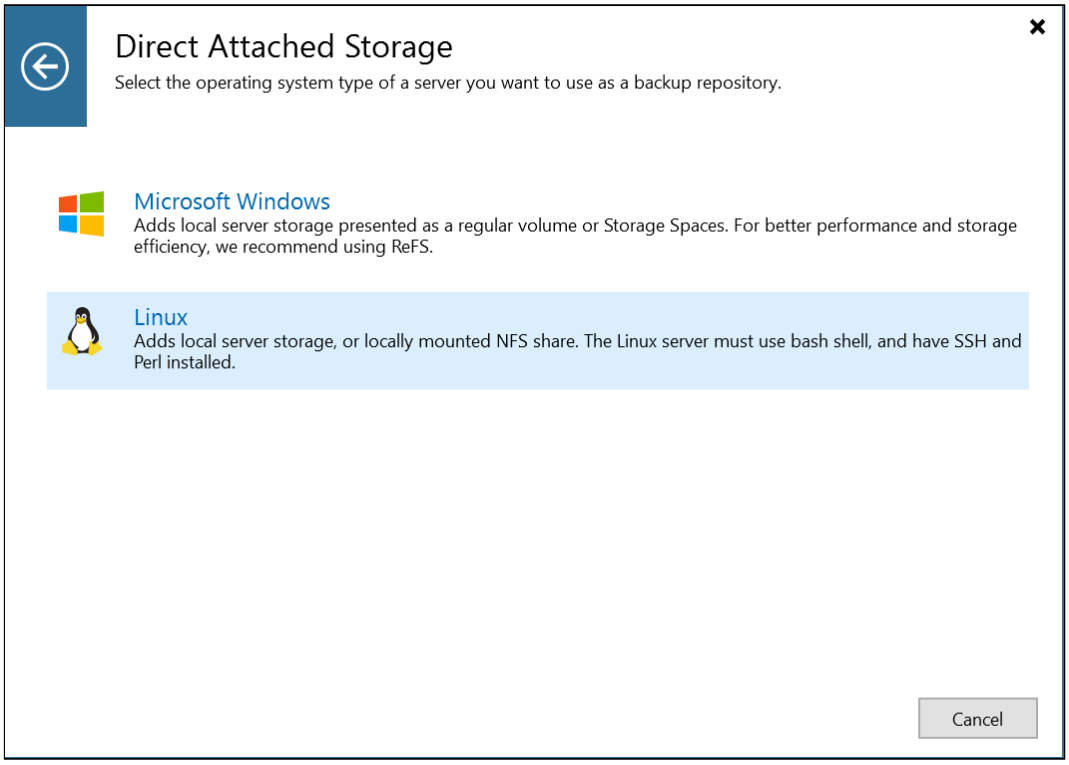


Finally we have to edit the configuration to make the mount persistent. Edit the /etc/ceph/rbdmap file. The first column is the target pool/image name, and the second column is the authentication. We also need to assign a mount point on the fstab file, and set the mount option “noauto”. Otherwise, you may get stuck in the computer boot up.

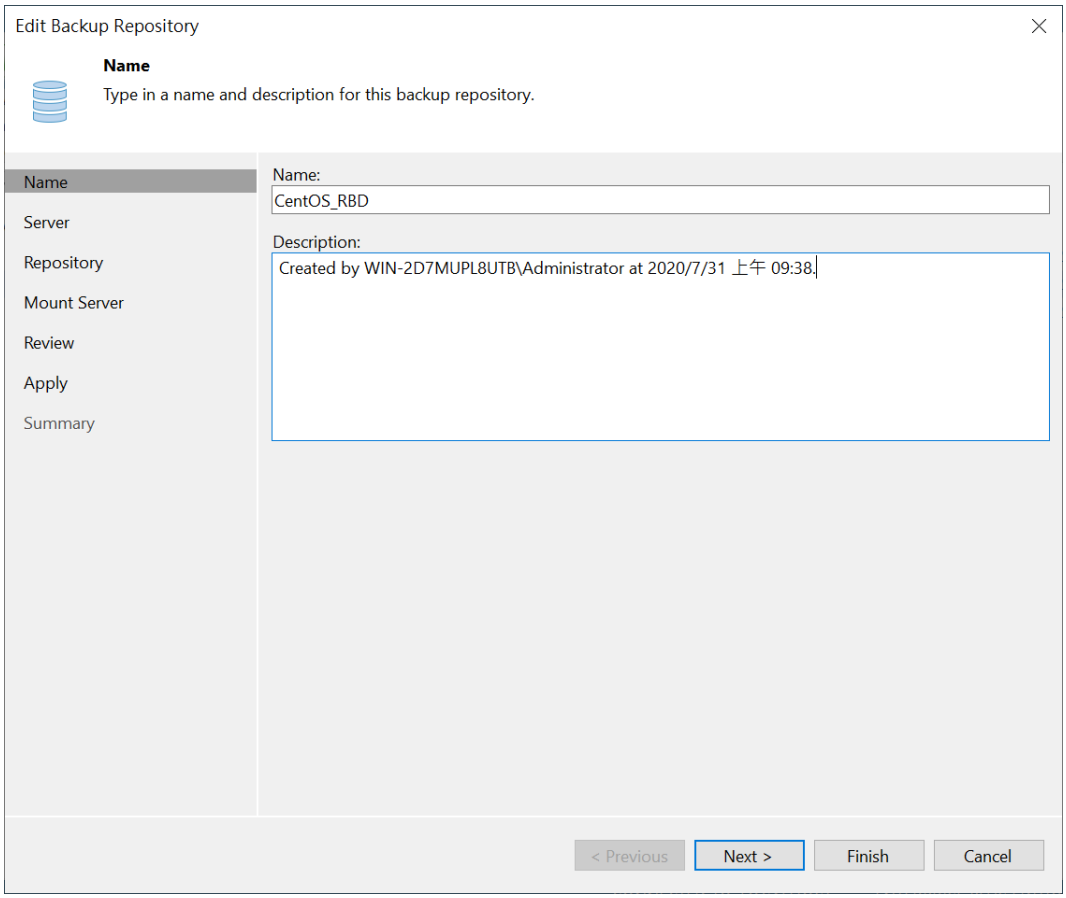


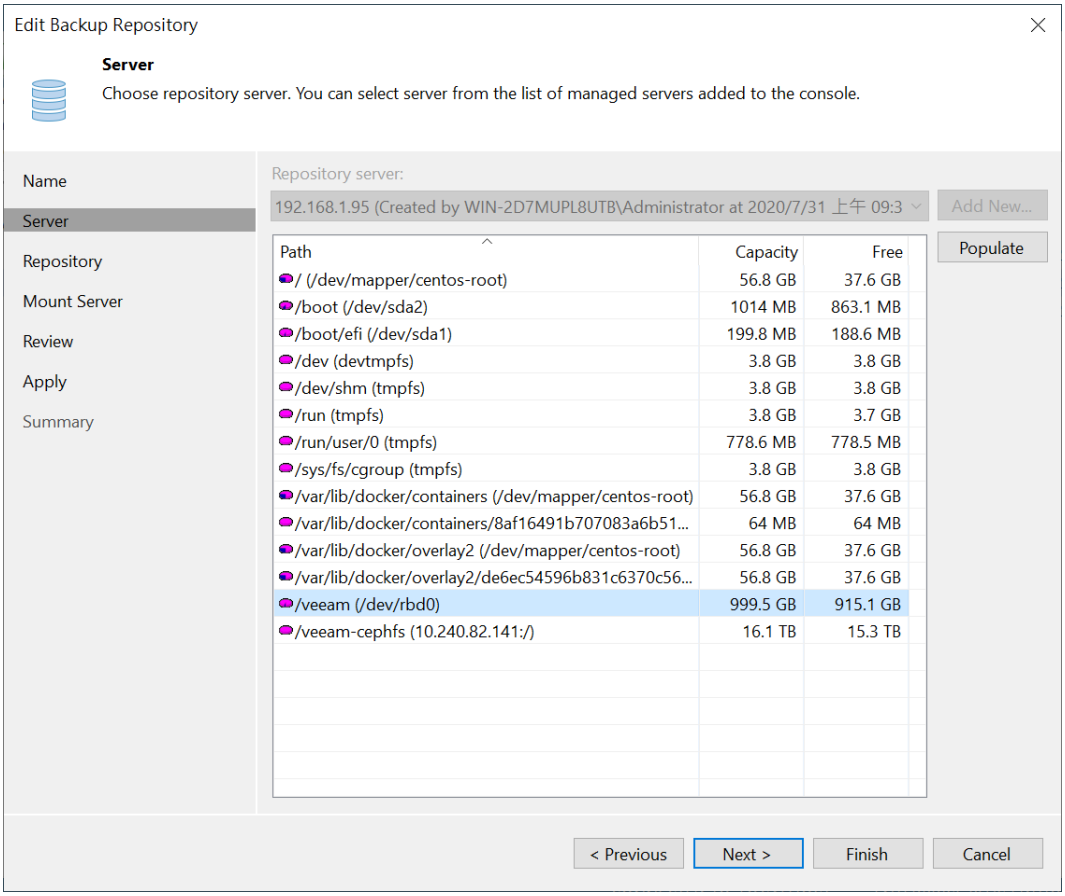
Add Ceph RBD as the Direct Attached Linux Backup Repository

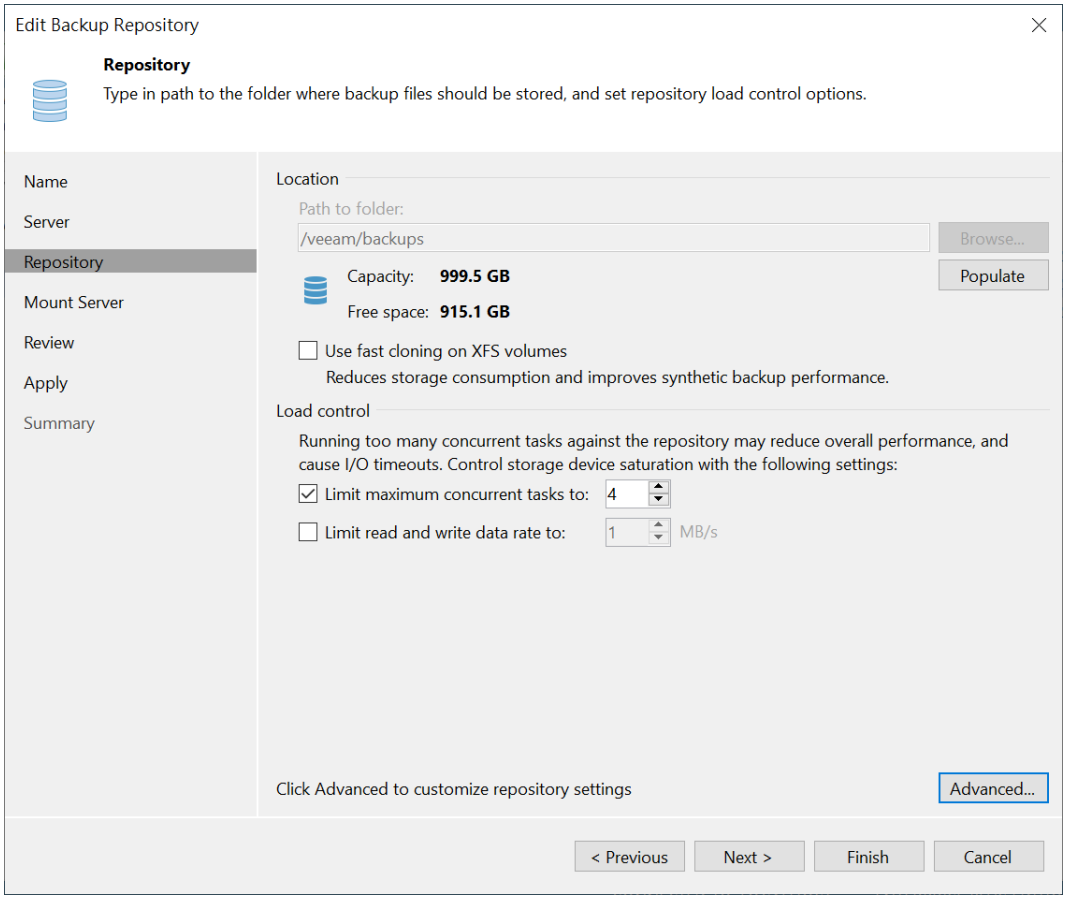


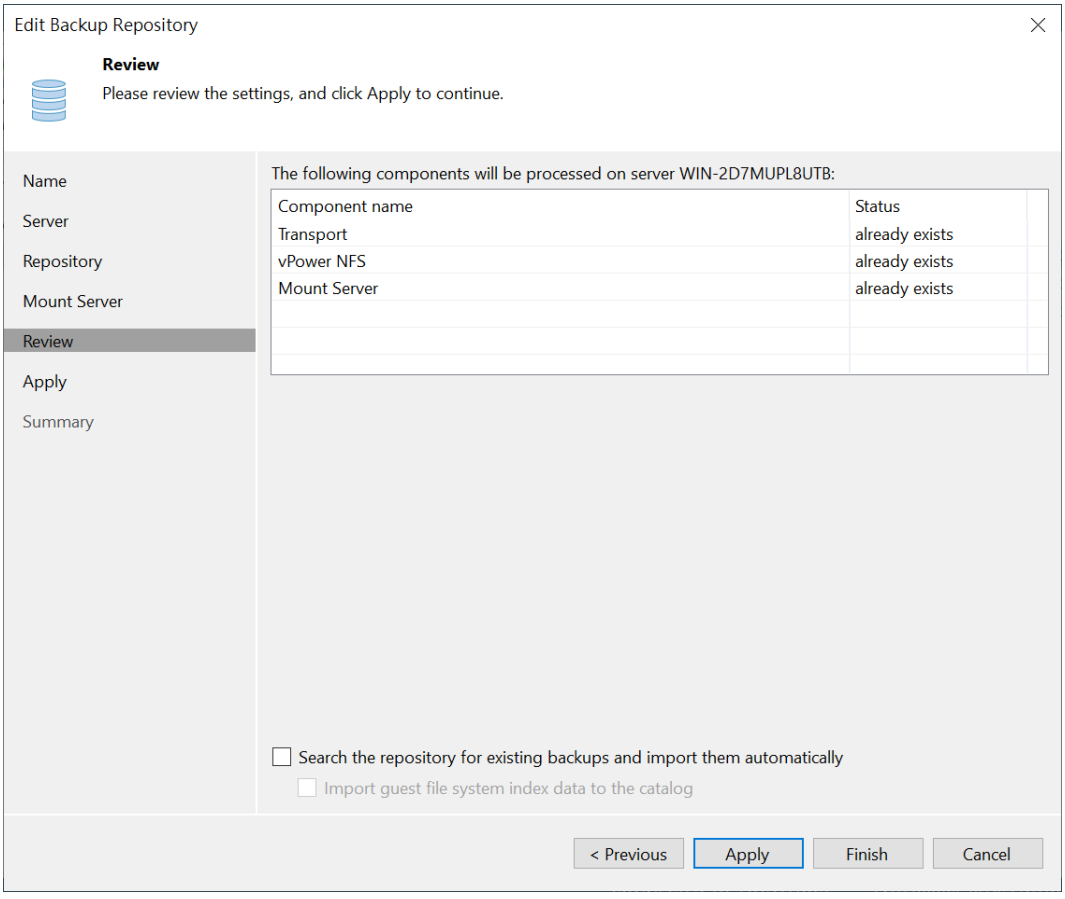


Edit Backup Repository


Edit Backup Repository
Name
Type in a name and description for this backup repository.
Name
Server
Repository
Mount Server
Review
Apply
Summary
Name:
CentOS_RBD
Description:
Created by WIN-2D7MUPL8UTB\Administrator at 2020/7/31 上午 09:38.
< Previous
Next >
Finish
Cancel


Edit Backup Repository
Server
Choose repository server. You can select server from the list of managed servers added to the console.
Name
Server
Repository
Mount Server
Review
Apply
Summary
Repository server:
192.168.1.95 (Created by WIN-2D7MUPL8UTB\Administrator at 2020/7/31 上午 09:3
Add New...
Populate
Path
Capacity
Free
/ (/dev/mapper/centos-root) 56.8 GB 37.6 GB
/boot (/dev/sda2) 1014 MB 863.1 MB
/boot/efi (/dev/sda1) 199.8 MB 188.6 MB
/dev (devtmpfs) 3.8 GB 3.8 GB
/dev/shm (tmpfs) 3.8 GB 3.8 GB
/run (tmpfs) 3.8 GB 3.7 GB
/run/user/0 (tmpfs) 778.6 MB 778.5 MB
/sys/fs/cgroup (tmpfs) 3.8 GB 3.8 GB
/var/lib/docker/containers (/dev/mapper/centos-root) 56.8 GB 37.6 GB
/var/lib/docker/containers/8af16491b707083a6b51... 64 MB 64 MB
/var/lib/docker/overlay2 (/dev/mapper/centos-root) 56.8 GB 37.6 GB
/var/lib/docker/overlay2/de6ec54596b831c6370c56... 56.8 GB 37.6 GB
/veeam (/dev/rbd0) 999.5 GB 915.1 GB
/veeam-cephfs (10.240.82.141:/) 16.1 TB 15.3 TB
< Previous
Next >
Finish
Cancel

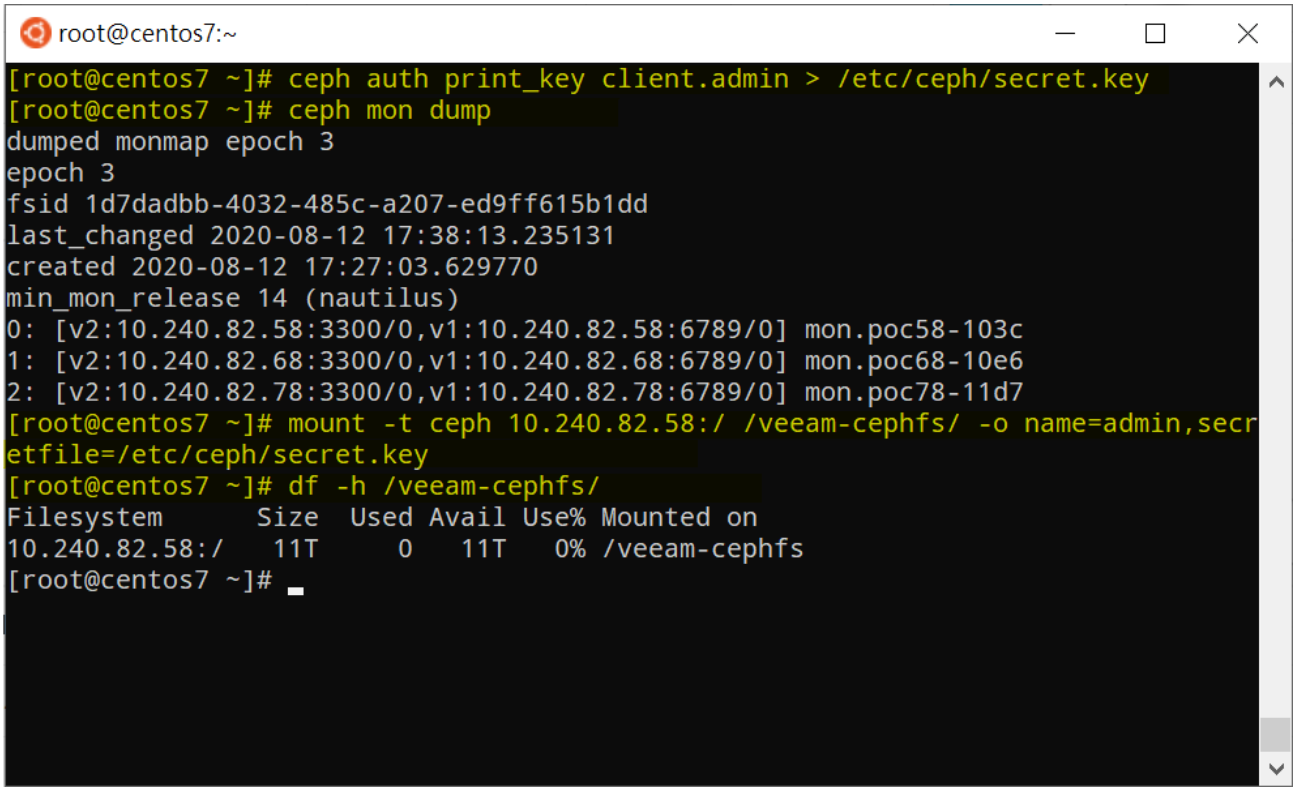

Edit Backup Repository
Repository
Type in path to the folder where backup files should be stored, and set repository load control options.
Name
Server
Repository
Mount Server
Review
Apply
Summary
Location
Path to folder:
/veeam/backups
Browse...
Capacity: 999.5 GB
Free space: 915.1 GB
Populate
Use fast cloning on XFS volumes
Reduces storage consumption and improves synthetic backup performance.
Load control
Running too many concurrent tasks against the repository may reduce overall performance, and cause I/O timeouts. Control storage device saturation with the following settings:
Limit maximum concurrent tasks to: 4
Limit read and write data rate to: 1 MB/s
Click Advanced to customize repository settings
Advanced...
< Previous
Next >
Finish
Cancel


Edit Backup Repository
Review
Please review the settings, and click Apply to continue.
Name
Server
Repository
Mount Server
Review
Apply
Summary
The following components will be processed on server WIN-2D7MUPL8UTB:
Component name
Status
Transport
already exists
vPower NFS
already exists
Mount Server
already exists
Search the repository for existing backups and import them automatically
Import guest file system index data to the catalog
< Previous
Apply
Finish
Cancel

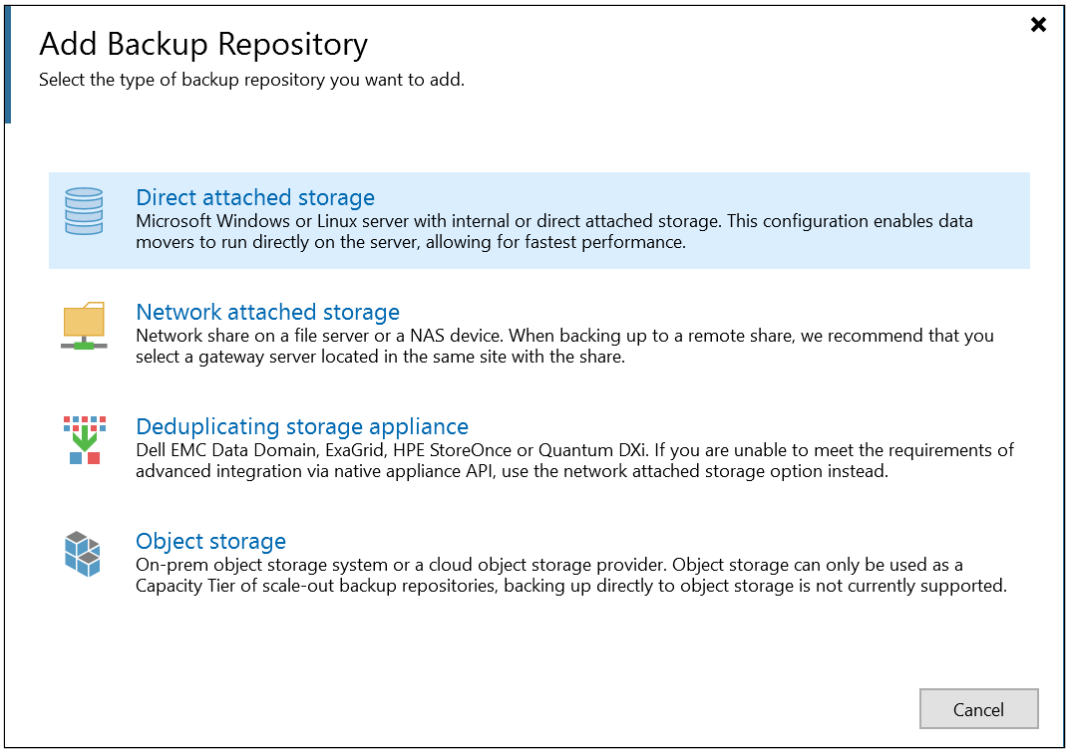
Setup Ceph File System as the Backup Repository

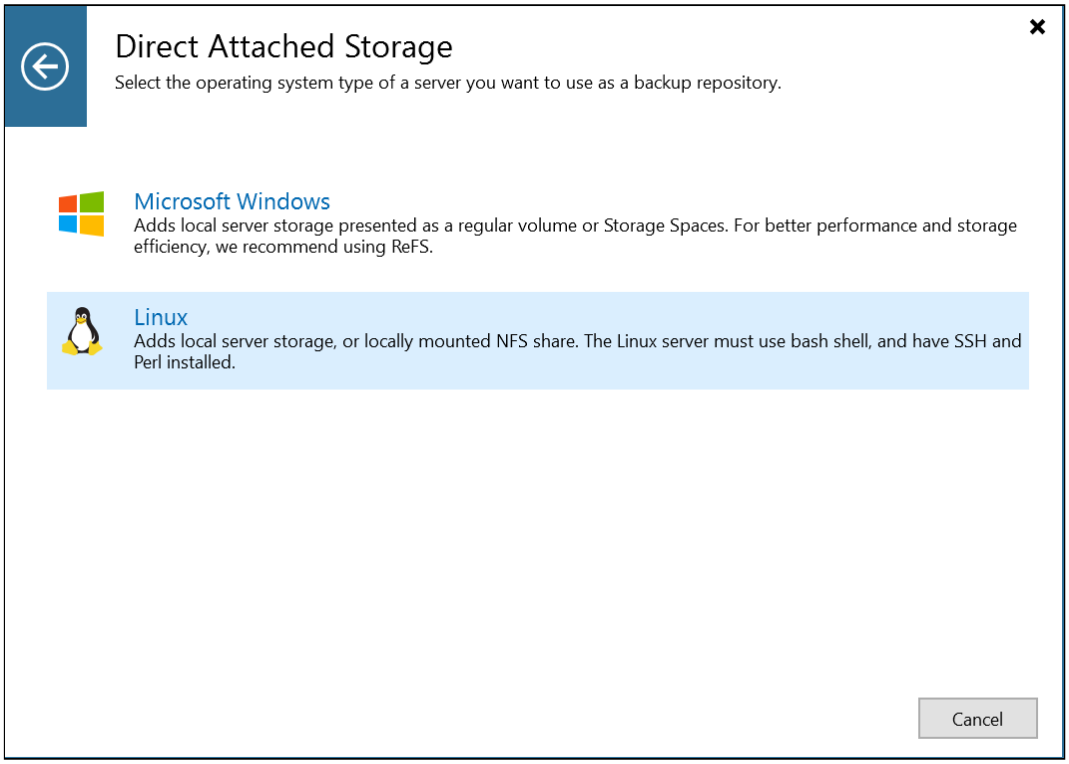
CephFS Protocol (File System)

The CephFS protocol provides a distributed POSIX compatible shired filesystem. You can use the UVS manager to deploy a CephFS, then mount the filesystem directly on a Linux server. Before setting the fstab, we have to mount the CephFS first. We need the **Ceph keyring** and the **Ceph monitor IP** to mount the CephFS. Finally, you need to edit the fstab file to mount the CephFS automatically after server boot.

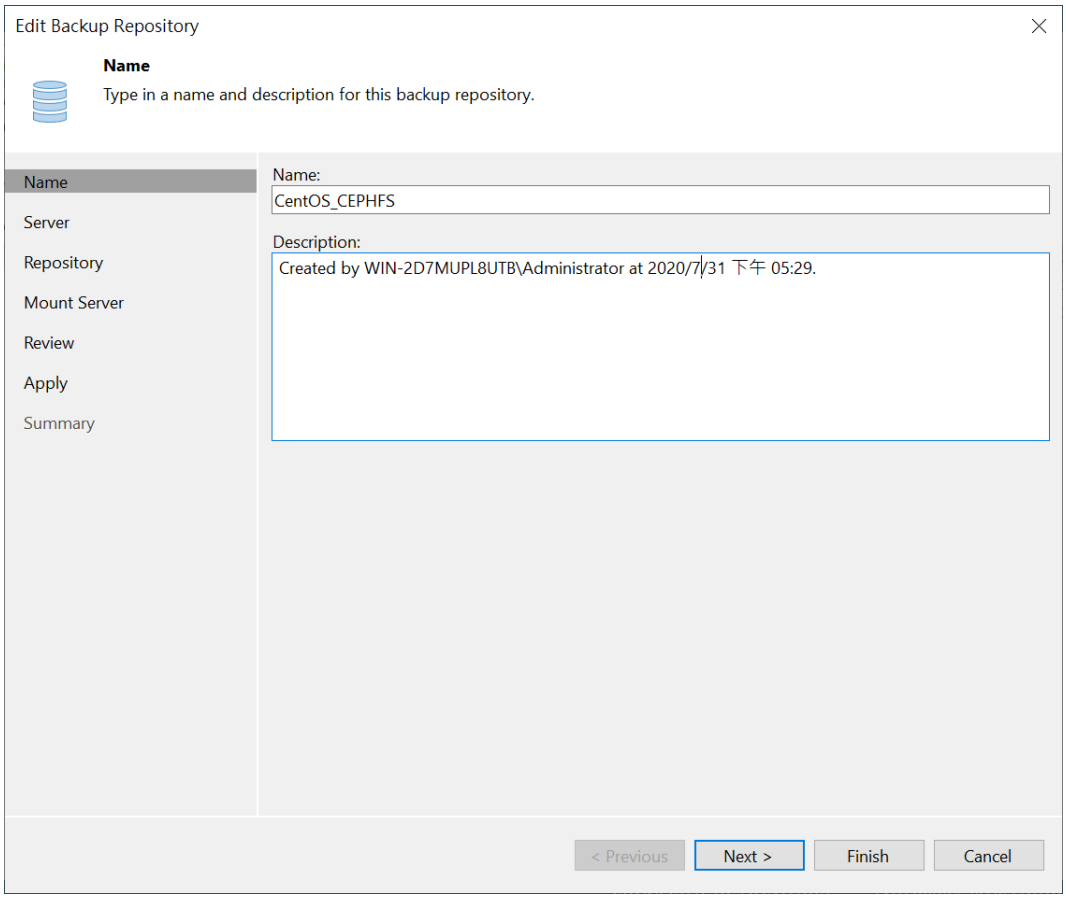


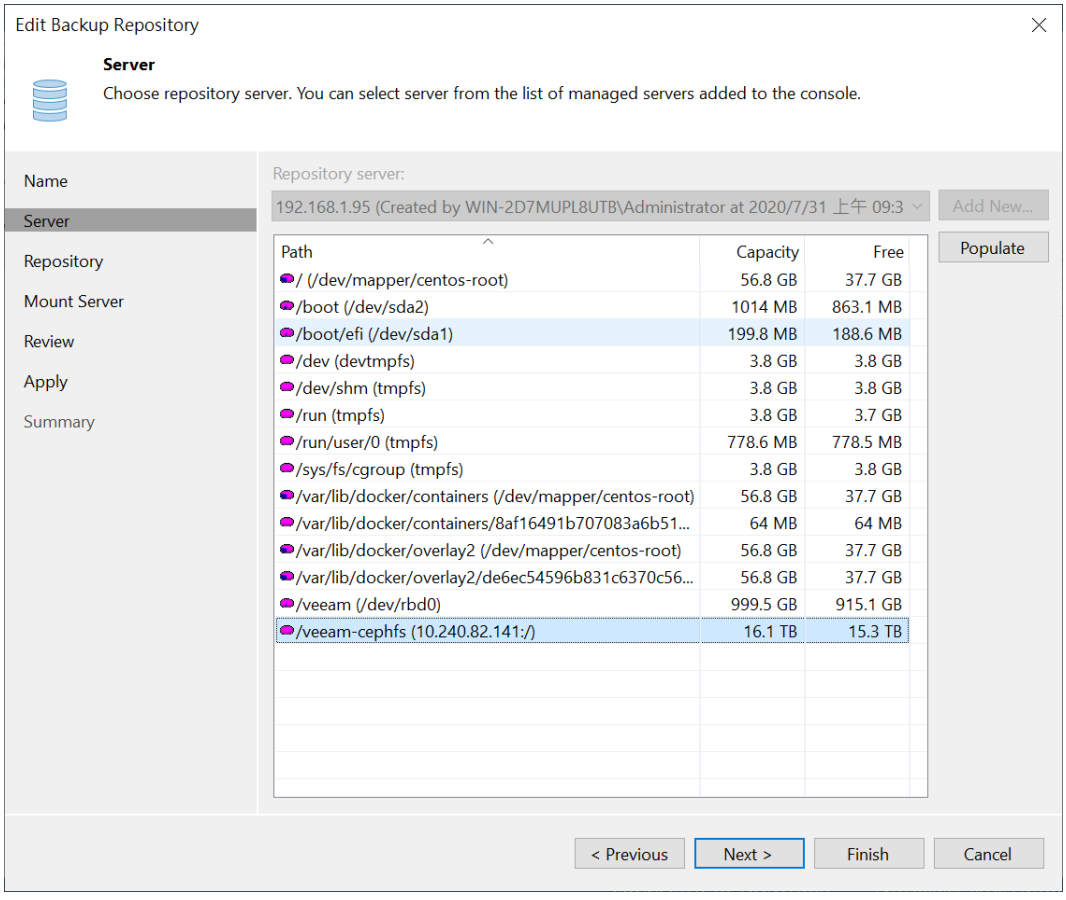
Add Ceph File System as the Backup Repository

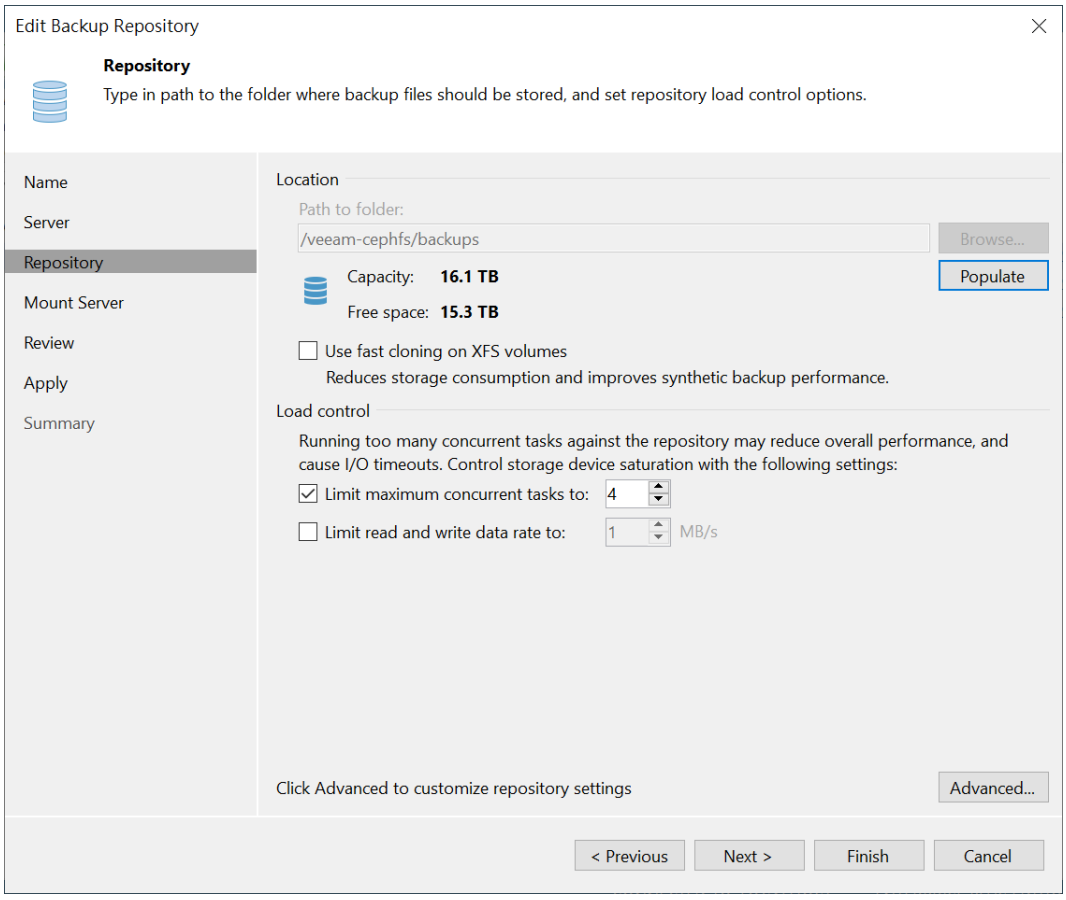


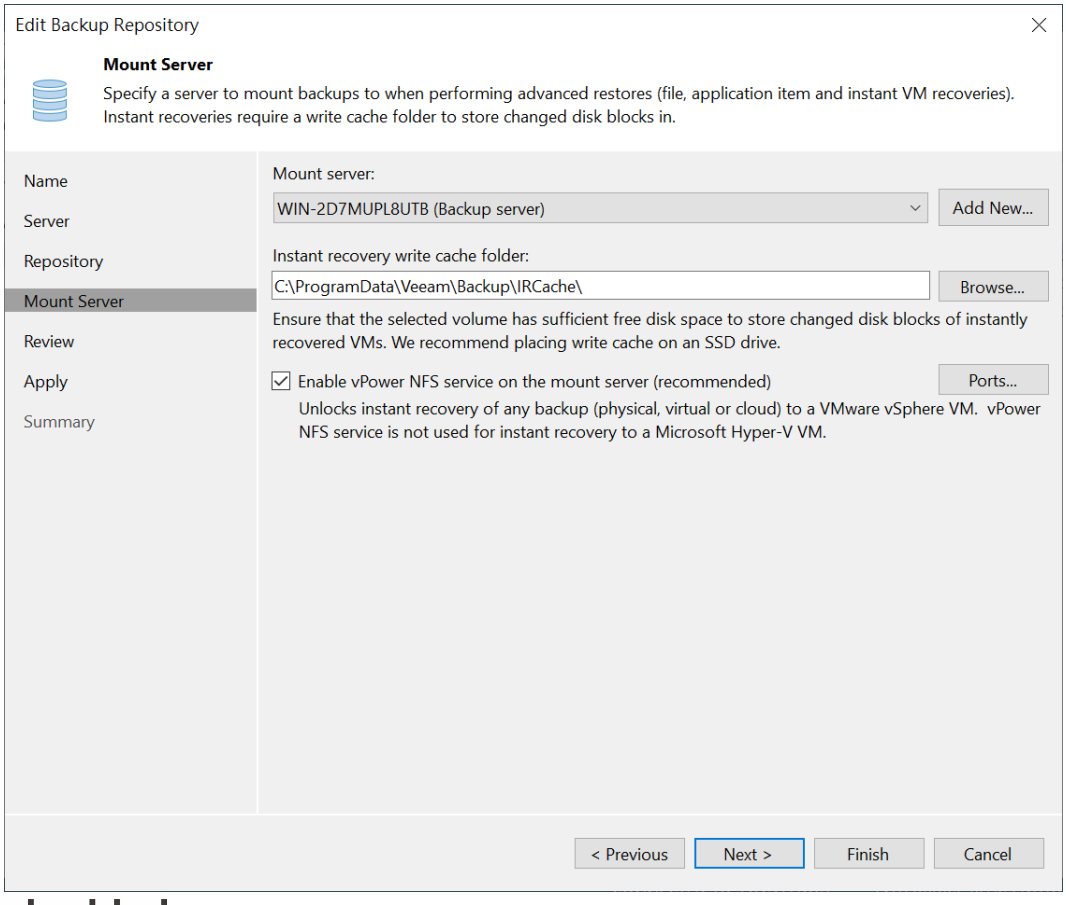


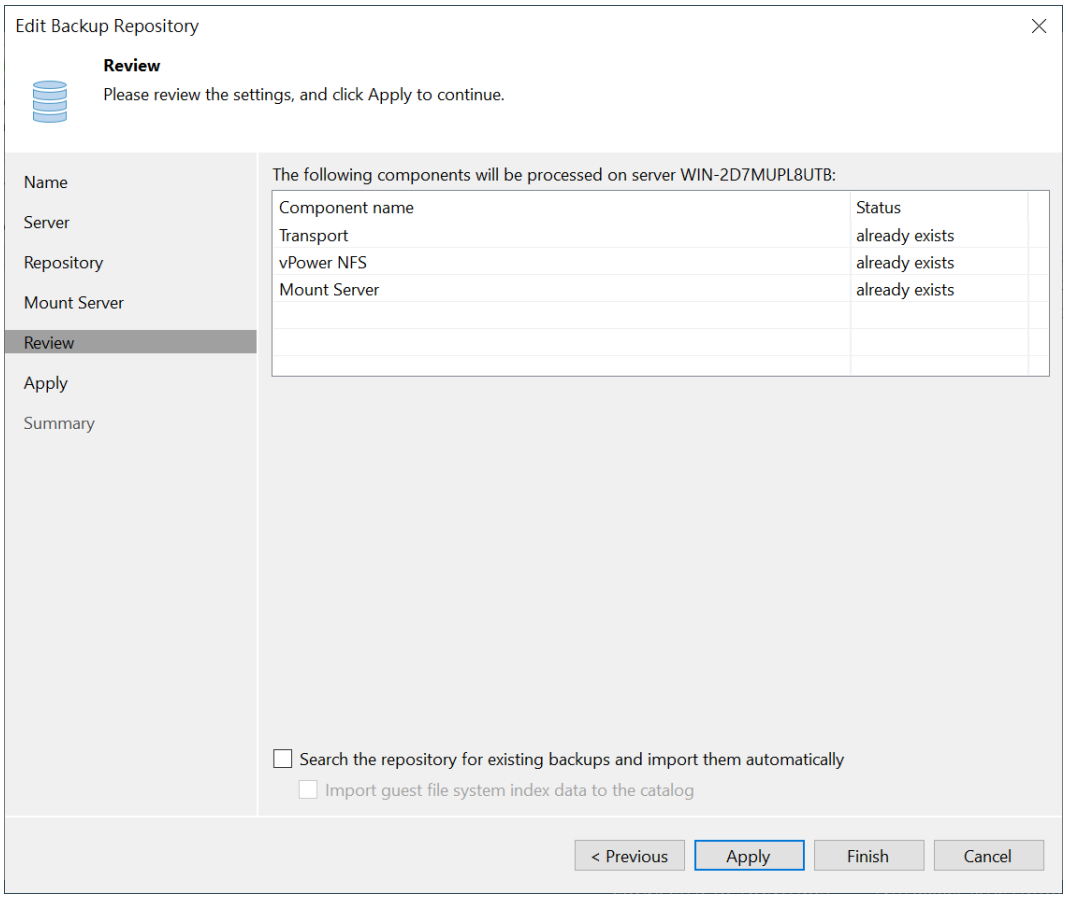
Edit Backup Repository


Edit Backup Repository
Name
Type in a name and description for this backup repository.
Name
Server
Repository
Mount Server
Review
Apply
Summary
Name:
CentOS_CEPHFS
Description:
Created by WIN-2D7MUPL8UTB\Administrator at 2020/7/31 下午 05:29.
< Previous
Next >
Finish
Cancel


Edit Backup Repository
Server
Choose repository server. You can select server from the list of managed servers added to the console.
Name
Server
Repository
Mount Server
Review
Apply
Summary
Repository server:
192.168.1.95 (Created by WIN-2D7MUPL8UTB\Administrator at 2020/7/31 上午 09:3
Add New...
Populate
Path
Capacity
Free
/ (/dev/mapper/centos-root) 56.8 GB 37.7 GB
/boot (/dev/sda2) 1014 MB 863.1 MB
/boot/efi (/dev/sda1) 199.8 MB 188.6 MB
/dev (devtmpfs) 3.8 GB 3.8 GB
/dev/shm (tmpfs) 3.8 GB 3.8 GB
/run (tmpfs) 3.8 GB 3.7 GB
/run/user/0 (tmpfs) 778.6 MB 778.5 MB
/sys/fs/cgroup (tmpfs) 3.8 GB 3.8 GB
/var/lib/docker/containers (/dev/mapper/centos-root) 56.8 GB 37.7 GB
/var/lib/docker/containers/8af16491b707083a6b51... 64 MB 64 MB
/var/lib/docker/overlay2 (/dev/mapper/centos-root) 56.8 GB 37.7 GB
/var/lib/docker/overlay2/de6ec54596b831c6370c56... 56.8 GB 37.7 GB
/veeam (/dev/rbd0) 999.5 GB 915.1 GB
/veeam-cephfs (10.240.82.141:/) 16.1 TB 15.3 TB
< Previous
Next >
Finish
Cancel


Edit Backup Repository
Repository
Type in path to the folder where backup files should be stored, and set repository load control options.
Name
Server
Repository
Mount Server
Review
Apply
Summary
Location
Path to folder:
/veeam-cephfs/backups
Browse...
Capacity: 16.1 TB
Free space: 15.3 TB
Populate
Use fast cloning on XFS volumes
Reduces storage consumption and improves synthetic backup performance.
Load control
Running too many concurrent tasks against the repository may reduce overall performance, and cause I/O timeouts. Control storage device saturation with the following settings:
Limit maximum concurrent tasks to: 4
Limit read and write data rate to: 1 MB/s
Click Advanced to customize repository settings
Advanced...
< Previous
Next >
Finish
Cancel


Edit Backup Repository
Mount Server
Specify a server to mount backups to when performing advanced restores (file, application item and instant VM recoveries). Instant recoveries require a write cache folder to store changed disk blocks in.
Name
Server
Repository
Mount Server
Review
Apply
Summary
Mount server:
WIN-2D7MUPL8UTB (Backup server)
Add New...
Instant recovery write cache folder:
C:\ProgramData\Veeam\Backup\IRCache\
Browse...
Ensure that the selected volume has sufficient free disk space to store changed disk blocks of instantly recovered VMs. We recommend placing write cache on an SSD drive.
Enable vPower NFS service on the mount server (recommended)
Ports...
Unlocks instant recovery of any backup (physical, virtual or cloud) to a VMware vSphere VM. vPower NFS service is not used for instant recovery to a Microsoft Hyper-V VM.
< Previous
Next >
Finish
Cancel



Benchmark on Various Setups

To benchmark the backup performance of various backup repositories, we set up tests with different backup repositories and three backup sources.

Backup sources we use for the tests are a SATA SSD based volume on a server, a Windows VM of Hyper-V, and a CentOS 7 VM and a Windows VM of VMWare.

There are several report counters recorded after each test.
Veeam backup report counters:
Duration: The duration is the total time used for the backup job.
Load: The resource component usages in percentage. They represent the amount of time that component is busy during the backup job.
Processing Time: This is the ratio between total amount of data read from source and the job duration.
Average Data Write Rate(MB/s): To understand the load of the backup repositories, we calculate this rate with the amount of data written to the repository divided by the duration of the job. This represents the load that the backup job gives to the Ceph cluster.

(1) Backup a volume on a SSD Drive

In this benchmark, we run five backup jobs to backup a volume on a SATA SSD inside the Veeam backup server to the backup repositories backed by five kinds of repository backend.
The test results are listed in Table 1.
All of their backup source is a volume based on an SATA SSD.

1. The first test uses a Linux VM as the repository server and mount a virtual disk backed by a Ceph RBD with replica 3 pool.
2. The second test uses a Linux VM as the repository server and mount a CephFS volume backed with replica 3 pool.
3. The third test uses a Linux VM as the repository server and mount a virtual disk backed by a Ceph RBD with erasure code (K=4, M=2) pool.
4. The fourth test uses a Linux VM as the repository server and mount a CephFS volume backed by erasure code (K=4, M=2) pool.
5. The fifth test uses a physical Linux server as the repository server and mount a virtual disk backed by a Ceph RBD with erasure code (K=4, M=2) pool.

The backup job read 200.1 GB data from the source. After Veeam B&R applies deduplication and compression, only 69.7GB of data is transferred to the backup repository.

All of the backup jobs with different repositories have similar procession rates. The loads to the target backed by CephFS on a virtual machine and RBD on a bare metal server are only 4% and 3%. Target backed by RBD on a virtual machine has a higher load. All three back jobs have the bottleneck on the backup source which is the SATA SSD.

The processing rate around 560MB/s is about the maximum throughput of the backup source SATA SSD.

Using a replica pool or erasure code pool as the backend of a repository does not make a difference on a single job. The benchmark of replica vs. erasure code pool needs to do a test with more concurrent backup jobs.

Table 1. Backup a volume from a server with a SATA SSD.

Disk Size (Data processed)	237.9 GB
Data Read from the source	200.1 GB
Data Transfered to Ceph after Deduplication and Compression	69.7 GB
Deduplication	1.3X
Compression	2.7X

Table 2.

Backup Repository	Duration (sec)	Source (%)	Proxy (%)	Network (%)	Target (%)	Processing Rate (MB/s)	Average Data Write Rate (MB/s)
Linux VM, RBD-replica 3	646	83	33	84	21	554	110
Linux VM, CephFS-replica 3	521	97	25	31	5	564	137
Linux VM, RBD, EC	645	82	34	83	24	554	111
Linux VM, CephFS, EC	536	97	26	27	4	564	133
Linux Server, RBD, EC	526	97	21	16	3	561	136

Note: The Average Data Write Rates are calculated by Data Transferred divided by Duration. These rates represent the workloads of the Ceph cluster in these backup jobs.

(2) Backup a Windows 10 VM on Hyper-V

In this benchmark, we backup a Hyper-V instance that is stored on a SATA hard drive. The processing rates of these jobs reach the upper limit of HDD bandwidth. We can also find the bottleneck is on the source because their loads are busy during 99% of the job duration. Ceph cluster, the target, workload from the Veeam backup jobs is light. Ceph cluster is only busy at 6% to 1% of the working time.
Compared to the previous benchmark, the processing rate of the VM backup is much lower than the SSD backup. This is mainly because the VM data is stored in a hard drive.

Table 3.

Diak Size (HDD)	127 GB
Data Read from source	37.9 GB
Data Transfered to Ceph after Deduplication and Compression	21.4 GB
Deduplication	3.3X
Compression	1.8X

Table 4. Backup a virtual machine image on SATA3 HDD

Backup Repository	Duration (sec)	Source (%)	Proxy (%)	Network (%)	Target (%)	Processing Rate (MB/s)	Average Data Write Rate (MB/s)

Linux VM, RBD volume, EC	363	99	7	3	6	145	60
Linux VM, CephFS volume, EC	377	99	7	2	1	142	58.1
Linux Server, RBD volume, EC	375	99	6	2	2	140	58.4

Note: The Average Data Write Rates are calculated by Data Transferred divided by Duration. These rates represent the workloads of the Ceph cluster in these backup jobs.

(3) Backup Virtual Machines on ESXi

This test backs up a CentOS 7 and a Windows 10 Virtual machines running on a HDD of VMWare ESXi 6.5 host to a repository backed by a Ceph RBD with 4+2 erasure code protection.

Table 5.

Source	CentOS VM	Windows 10 VM
Disk Size (HDD)	40 GB	32 GB
Data Read from source	1.8 GB	12.9 GB
Data Transfered to Ceph after Deduplication and Compression	966 MB	7.7 GB
Deduplication	22.1X	2.5X
Compression	1.9X	1.7X

Table 6.

Backup Source	Duration (sec)	Source (%)	Proxy (%)	Network (%)	Target (%)	Processing Rate (MB/s)	Average Data Write Rate (MB/s)
CentOS 7	122	99	10	5	0	88	8
Windows 10	244	99	11	5	1	93	32

Note: The Average Data Write Rates are calculated by Data Transferred divided by Duration. These rates represent the workloads of the Ceph cluster in these backup jobs.

Conclusions

According to the test reports, Ceph RBD and CephFS have similar performance. This meets our experience regarding the benchmark of RBD and CephFS performance. Comparing the characteristics of CephFS and RBD, they have their advantages and disadvantages. If you need to deploy multiple repository servers, you have to create an RBD image for each backup repository server as you can only mount Ceph RBD on one host. Compared to CephFS, using RBD is simpler as it does not need the metadata servers. We have to assign the RBD capacity size when created, so you have to resize its capacity when you need more space.

If you use CephFS as the repository, you have to deploy at least one metadata server (MDS) in the Ceph cluster. We also need a standby metadata server for high availability. Compared to the Ceph RBD, you don't need to give the file system a quota. So, you can treat the CephFS as an unlimited storage pool.

In this article, our tests back up only one VM in each backup job. According to the above test reports, we know the average data writing rate is related to the processing rate and data deduplication and compression efficiency. A faster source disk reduces the backup job duration and results in a faster processing rate. Depending on your infrastructure, you can deploy several concurrent jobs to back up different objects simultaneously. Ceph is very good at supporting multiple concurrent jobs. A 20x HDD OSD Ceph cluster powered by 3x Ambedded Mars 400 can offer up to 700MB/s aggregated writing throughput to the 4+2 erasure code pool. Deploying multiple current backup jobs gets the benefit of reducing the overall backup duration. The maximum performance of a Ceph cluster is almost linearly proportional to the total number of disk drives in the cluster.

We didn't test using S3 object storage as the backup repository in this article. S3 object storage can be used as the capacity tier in the Veeam Scale-Out backup repository and target archive repository for NAS backup. You can easily set up a RADOS gateway and create object storage users easily using the Ambedded UVS manager.