HPE Reference Architecture for Veeam Availability Suite with HPE Apollo backup target

Solution overview, best practices, and recommendations

Contents

Executive summary 3
Introduction 3
  Benefits of Veeam backups to HPE Apollo storage 4
Solution overview 4
  Design features 5
Solution components 7
  Hardware 7
  Software 7
Best practices and configuration guidance for the solution 8
  Configuration of storage on HPE Apollo 4200 Gen9 server 8
  Repository configuration 10
  Backup job configuration 12
HPE Apollo and Veeam use cases 14
  Multistream throughput 15
  Capacity utilization and deduplication 16
Capacity and sizing 17
  Environment sizing 17
  HPE Apollo 4200 Gen9 server capacity sizing guide 17
  Workload 17
  Analysis and recommendations 18
Summary 18
Appendix A: Bill of materials 18
Resources and additional links 20

Executive summary

As the amount and types of data that your business owns continue to grow, and as more of your IT deployments include virtualized workloads, there is an immediate and critical need to protect this data reliably. At the same time, the threat and risk of data loss is growing in both variety and volume. Network and power outages, component failure, human error, willful malevolence, data corruption, software bugs, site failures, and even natural disasters are just a few of the sources of application downtime and data loss.

Many businesses today do not have data protection or even storage specialists on staff. This means that a reliable, automated data backup system that is manageable by existing staff without specialized expertise is critical to keeping the business running, meeting ever-evolving user expectations, and remaining competitive.

Typically, resource-constrained businesses face a three-fold data protection challenge:

1. Can I find a data protection solution that fits my budget?
2. How easy is it for my IT staff to run backups and restores?
3. How reliable is my data protection solution in the face of downtime and ransomware attacks?

Together, Hewlett Packard Enterprise and Veeam offer this Reference Architecture for a data protection solution based on HPE Apollo servers and Veeam Availability Suite. This Reference Architecture is verified jointly by Hewlett Packard Enterprise and Veeam and provides half-capacity and maximum-capacity configurations specifically built, tuned, and tested for Veeam with different performance, capacity, and features. These configurations can be used as building blocks for customers with requirements for larger systems.

The solution offers the following benefits to deliver a cost-effective data protection infrastructure for virtualized environments:

- Data recovery—instant virtual machine (VM) recovery and file-level recovery.
- Reliable protection—automatic verification of the recoverability of every backup and replica.
- Rapid restoration—for speedy restoration of critical applications and workloads.
- Scalability and flexibility—readily deployed and easily scalable, by varying the number and capacity of data disks with large and small form factor drives having capacities up to 12 TB each.

Target audience: This document is intended for presales consultants, solution architects, backup, system or storage operators, and administrators who are designing, implementing, and maintaining data protection solutions based on Veeam running on HPE storage-optimized Apollo servers.

Document purpose: This Reference Architecture highlights recognizable benefits to technical audiences and provides a blueprint for deploying a half-capacity or maximum-capacity data protection solution with HPE Apollo servers and Veeam Availability Suite.

Introduction

Data is at the core of all organizations, and protecting it is more important than ever. The risks of temporarily or permanently losing access to data are widely documented. Backup remains the key process to mitigate this risk. As server virtualization and data volumes increase, organizations are looking for ways to efficiently and affordably manage backup footprints within their data centers. Hewlett Packard Enterprise and Veeam have teamed up to offer a cost-effective data protection infrastructure for virtualized environments to run Veeam software and to store Veeam backup data. Many organizations are looking for a solution that reduces deployment complexity. Bundling Veeam with HPE Apollo storage servers provides an affordable solution that also offers the benefits of simplicity.

Hewlett Packard Enterprise offers multiple solutions and architectures specifically built, tested, and tuned for Veeam that offer different performance, capacity, and features. This solution, based on the HPE Apollo 4200 Gen9 server, provides the highest storage density and throughput to store Veeam backup data while also having HPE ProLiant-like compute resources to run the Veeam Availability Suite software and manage the Veeam Availability Suite functions. HPE Apollo storage servers have the recognized quality and leadership of HPE servers, and provide the ability to scale capacity within the box—with no need for a separate storage device. This Reference Architecture documents two capacity configurations for the HPE Apollo 4200 Gen9 storage server, which is also referred to as the HPE Apollo server throughout this document.

This solution is a fully tested and recommended configuration from Hewlett Packard Enterprise and Veeam. This document helps you minimize the risk of deployment by defining the hardware and software required to implement this solution, as well as providing configuration guidance, best practices, and data regarding expected performance.

Benefits of Veeam backups to HPE Apollo storage

There are many benefits to using the HPE Apollo 4200 Gen9 storage server in this solution. In addition to providing hundreds of terabytes of local storage capacity, it also has the compute resources required for running the Windows® operating system and Veeam Availability Suite software on the same server. This greatly simplifies the design in comparison with architectures based on compute-only, plus storage-only components.

HPE Apollo storage servers have the ability to write backup data to local storage, so backup duration is significantly reduced compared to the time required for transferring backup data to a separate storage resource using either a Fibre Channel or Ethernet-based transfer medium.

The testing done by Hewlett Packard Enterprise and Veeam, as documented in this Reference Architecture, further reduces the complexity of the solution by providing guidance on the optimal configuration of the hardware and storage resources. This achieves the most efficient use of the configuration with minimal additional tuning. Configuration guidance and best practices are provided later in this document.

As an added benefit that further minimizes the complexity of solution licensing, no capacity or other storage licenses are required to take full advantage of the storage capacity of the HPE Apollo platform. Whether the disk bays are partially or fully populated, no additional licenses are required.

Solution overview

With data centers becoming increasingly virtualized, Hewlett Packard Enterprise understands the importance of having reliable data protection infrastructure for your business-critical VMs. The Hewlett Packard Enterprise Solutions Team has designed, implemented, and performance tested this HPE Reference Architecture for integrity, reliability, and overall quality. This data protection solution consists of a highly virtualized environment including an HPE Apollo 4200 Gen9 storage server and HPE ProLiant DL380 Gen10 servers using primary storage from HPE 3PAR StoreServ 8200 and HPE 3PAR StoreServ 9450 arrays. Alternatively, an HPE Nimble Storage flash array could also have been used to achieve similar results. The solution uses VMware ESXi™ (ESXi) as the hypervisor, and can be scaled-up or scaled-out depending on the needs of the user. Veeam Availability Suite offers a robust tool compilation to quickly and effectively back up and replicate crucial business data.

HPE 3PAR and HPE Nimble Storage arrays are steadfast primary storage devices with ample integration with Veeam Availability Suite supporting the most advanced Veeam Snapshot Orchestration features. The combination of this hardware and software was performance tested to ensure dependability for enterprise applications.

This solution makes extensive use of the Resilient File System (ReFS) on the HPE Apollo storage server. ReFS is a relatively new file system built on its predecessor New Technology File System (NTFS), but with enhanced capabilities. It is designed to maximize data integrity and resistance to corruption in spite of software or hardware failures. ReFS is optimized for high scalability and performance, making it perfectly suited for environments with growing volumes of data.

Active full backups can be very resource intensive, so Veeam takes advantage of the block cloning capability in ReFS to create Virtual Synthetic Full (VSF) backups. The VSF backup does not use any network resources. Instead, the backup is created in the storage repository from backup files already existing in the repository. VSF backups are generally faster than traditional full backups, and they reduce the impact of data protection activity on the production environment and on the storage. Furthermore, the use of VSF backups in ReFS makes more efficient use of the storage capacity of the HPE Apollo 4200 Gen9 server. With the workload described for testing this solution, 33% less capacity was consumed in ReFS to store the backups compared to NTFS without the VSF backups. The reduction in capacity varies by workload, and therefore production environments might see different space savings.

Figure 1 illustrates the environment for this solution. Veeam Availability Suite is running on the HPE Apollo storage server, which also contains an ReFS on local disks for use as a primary target for Veeam backups. The hypervisors in this case are VMware ESXi servers. Data on the hypervisors is contained on virtual volumes exported over Fibre Channel from the HPE 3PAR StoreServ 8200 and HPE 3PAR StoreServ 9450 arrays. Production VMs reside on these virtual volumes. In this model, the two HPE 3PAR arrays capture HPE 3PAR storage snapshots of the virtual volumes containing the datastores for all the VMs. Veeam then backs up these snapshots to the ReFS repository on the local disk within the HPE Apollo server.

All fabric connections, including the VMWare ESXi connections, the HPE 3PAR connections, and the Veeam Availability Suite server connections, use two active paths of 16 Gb Fibre Channel (FC). This is the minimum recommended connectivity to enable data processing at the maximum speed supported by this tested configuration. Weekly full backups and daily incremental backups are configured using Veeam Backup Jobs.

Figure 1 shows the solution environment layout, in which data is read directly from the two HPE 3PAR arrays and backed up to an ReFS repository locally on the HPE Apollo server.

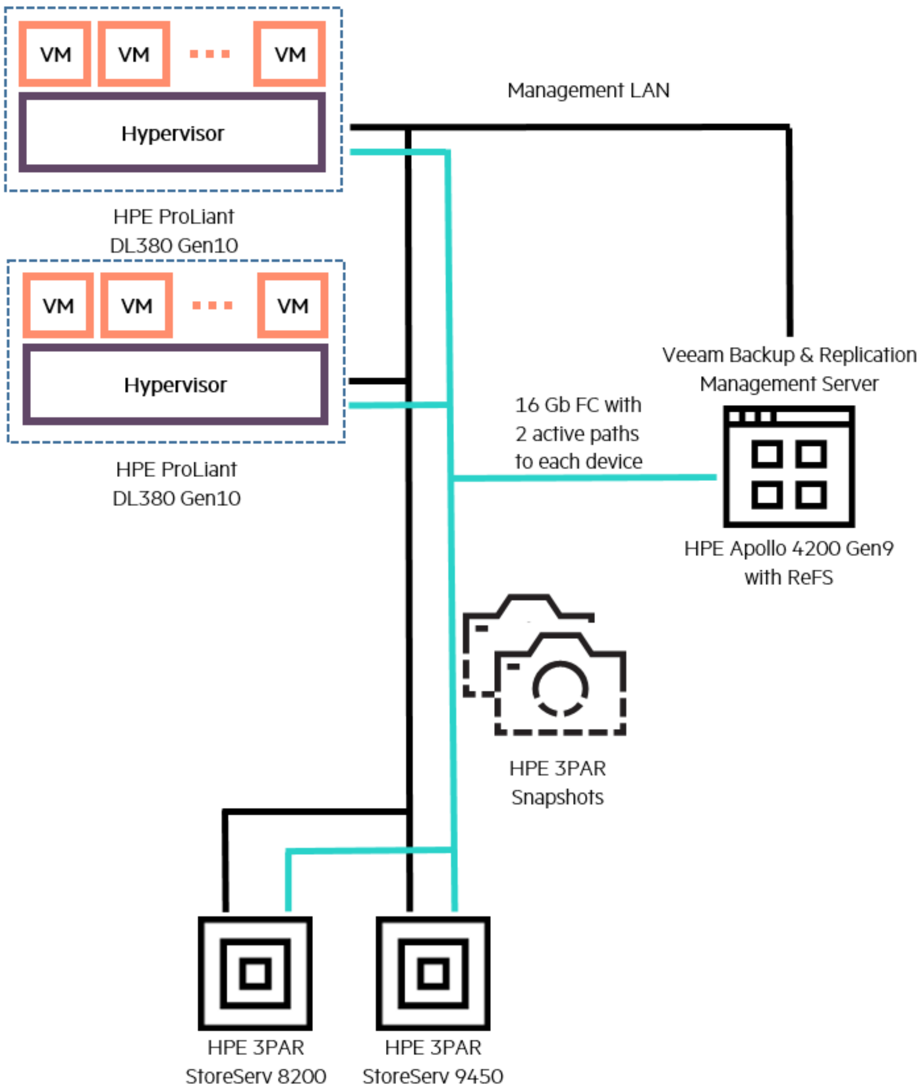

Figure 1. Solution environment

Design features

This solution highlights HPE 3PAR storage snapshot integration with Veeam Availability Suite. The same storage snapshot integration is available with HPE Nimble Storage flash arrays, so an HPE Nimble Storage array would provide the same advantages. During a traditional backup, Veeam Availability Suite defaults to using VMware vSphere® hypervisor snapshots to write data during the backup. This can be a cumbersome process, as Veeam Availability Suite continuously queries production ESXi hypervisors for backup data. After the traditional hypervisor snapshot is deleted, all changes to the VM during the lifetime of the snapshot need to be merged back into the VM before the snapshot can be deleted. This task is very I/O intensive and impacts performance of the VM. The VM could become slow to respond while changes are being merged in prior to deleting the VM hypervisor snapshot. To effectively mitigate this process, Veeam's storage snapshot integration with HPE 3PAR or HPE Nimble Storage arrays is used.

This HPE 3PAR/HPE Nimble Storage hardware snapshot integration helps to reduce the impact on production data and VM performance due to data protection activity by reading data directly from the HPE 3PAR snapshot, instead of the production ESXi hosts. The sequence below, as illustrated on the right hand side of Figure 2, shows the benefit of this integration.

1. When a backup is initiated (either full or incremental), Veeam creates a hypervisor VM-Snapshot for all VMs in a backup job.
2. Veeam then creates an HPE 3PAR storage snapshot using the HPE 3PAR storage snapshot technology. Creating the HPE 3PAR snapshot normally takes no more than 5 seconds.
3. As soon as the HPE 3PAR snapshot has been created, Veeam quickly deletes the hypervisor snapshot and allows the VM to continue with normal operation.
4. Veeam then maps the HPE 3PAR storage snapshot to the Veeam Proxy to run the actual backup.
5. After the backup has been completed, the storage snapshot is normally deleted, thus releasing capacity on the HPE 3PAR array. Optionally, Veeam can keep one or more HPE 3PAR snapshots as additional high-speed restore points.

The short lifespan of the hypervisor snapshot is a key benefit of this process, compared to the traditional process of performing the backup on the hypervisor snapshot itself, and minimizes the length of time the VM is impacted by the backup application. Another key benefit is the direct path and high-speed Fibre Channel connectivity between the HPE 3PAR array and HPE Apollo server, compared to the traditional two hops with data moving from the HPE 3PAR array to the ESXi server over Fibre Channel, and from the ESXi server to the HPE Apollo server via LAN.

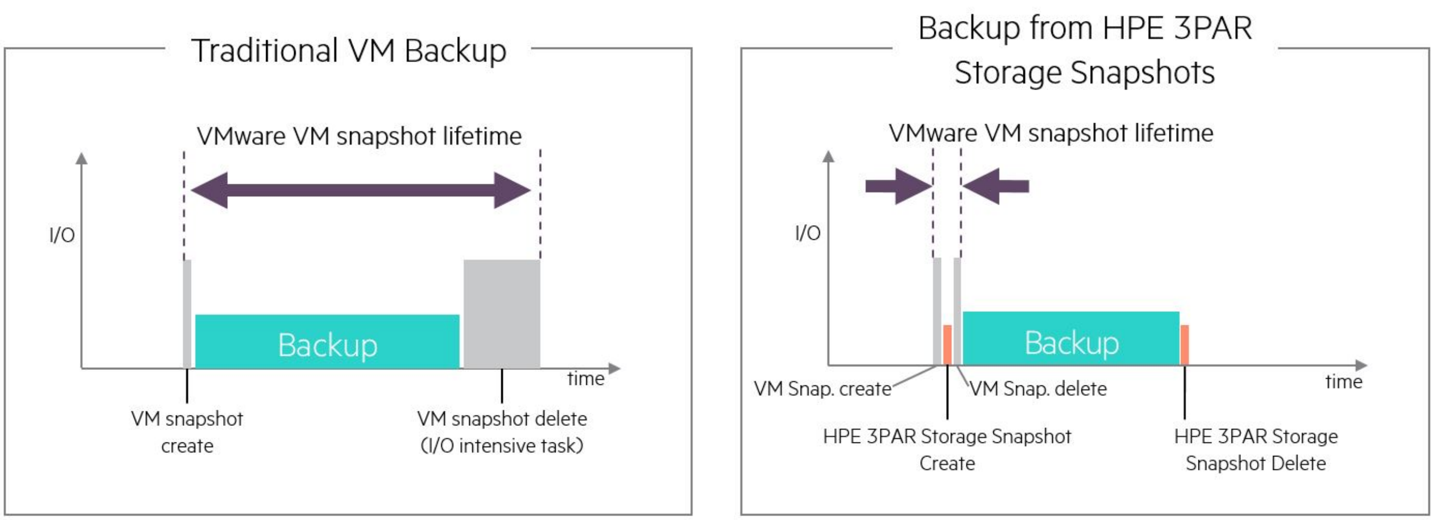

Figure 2. HPE 3PAR storage snapshot implementation

Solution components

This section outlines the hardware and software components used to create this Reference Architecture for Veeam Availability Suite with HPE Apollo Backup Target.

Hardware

HPE Apollo 4200 Gen9 storage server

The HPE Apollo 4200 Gen9 server offers one of the highest storage capacities in any 2U server. In addition to this high-density storage, the HPE Apollo 4200 Gen9 chassis also contains HPE ProLiant-like compute capabilities and HPE integrated Lights Out (iLO) management. No special racks are required as it fits easily into standard racks with a depth of 32-inches per server. The configuration can easily be customized using a different number of disks, as well as selecting disks with different capacity and performance.

For this testing, the maximum storage configuration available at time of writing was used. There were 24 large form factor (LFF) SATA data disks plus two hot spares. The capacity of each disk is 12 TB. The data drives were configured in RAID 60 for a total usable capacity of 240 TB (218 TiB). The HPE Apollo server also contained two 1.6 TB SSDs for the OS and for Veeam vPowerNFS cache.

HPE 3PAR StoreServ 8200 and HPE 3PAR StoreServ 9450 storage arrays

Both HPE 3PAR storage arrays used in this testing provided primary storage to the ESXi hosts. The HPE 3PAR StoreServ 8200 storage array, using 6 x 1.9 TB SSDs and running OS version 3.3.1 MU2, was connected by two paths to a 16 Gb Fibre Channel backend fabric. Storage from this array was used to contain the datastores for half of the VMs being backed up with Veeam.

The HPE 3PAR StoreServ 9450 storage array, using 24 x 1.2 TB SSDs and also running OS version 3.3.1 MU2, was connected to the same 16 Gb Fibre Channel backend fabric using two paths. Similar to the HPE 3PAR StoreServ 8200 array described above, this array was used to contain datastores for the remaining VMs that were being backed up with Veeam.

Veeam Availability Suite Management Server

Veeam Availability Suite was hosted on the HPE Apollo server described above, which was running Windows Server® 2016. The Backup & Replication server had the roles of Veeam server, proxy server, and repository server, and was also running the SQL server free edition installed with Backup & Replication.

VMware ESXi Servers

An HPE ProLiant Apollo 2000 Gen9 16-core (2 x Intel® Xeon® CPU E5-2640 @ 2.60GHz) server was configured with 4 TB of storage and 128 GB of physical memory. ESXi™ Server 6.7.0 was installed on this server.

An HPE ProLiant DL380p Gen10 20-core (2 x Intel Xeon Gold 5115 CPU @ 2.40GHz) server was configured with 10 TB of storage and 256 GB of physical memory. ESXi Server 6.7.0 was installed on this server.

Note

All VMs hosted on these servers utilized HPE 3PAR storage on either the HPE 3PAR StoreServ 8200 or HPE 3PAR StoreServ 9450 storage array described above.

Software

Veeam Availability Suite 9.5

Veeam Availability Suite enables the recovery of any IT service, related applications, and data within seconds and minutes. Turn your enterprise into a hyper-available enterprise with Veeam Availability Suite for any app, any data, and any cloud.

Best practices and configuration guidance for the solution

Configuration of storage on HPE Apollo 4200 Gen9 server

The HPE Apollo 4200 Gen9 server provides ample flexibility for a variety of different configurations. For this test the system was configured with the highest available capacity; therefore, 25 LFF 12 TB SATA disk drives were used. Even though these disks are optimized for capacity rather than for performance, this solution demonstrated excellent ingest performance during this testing.

RAID Level

To reduce the risk of data loss caused by hardware failures, two independent RAID-6 groups were created with 12 disks each, and one hot spare. These were combined via striping into a single RAID-60 volume. Both layers (RAID-6 and striping) are entirely managed by the SmartArray controller card in the HPE Apollo 4200 Gen9 server and are thus offloaded from the OS. RAID-6 is designed to survive two concurrent disk failures without losing data. When a disk failure occurs, the automated reconstruction begins using the spare disk, but can take several hours. RAID-6 protects from the risk of a second disk failure that might occur before the reconstruction is completed. To further reduce the risk of data loss and reduce the time required for the reconstruction process, two RAID-6 volumes with 12 disks each were striped to form a single RAID-60 volume. This configuration produces better data protection overall, compared to a single RAID-6 volume with 24 disks.

Strip Size

The choice of strip size in a RAID-60 configuration has a direct influence on read/write performance. This impact is significant, as overall system performance can double simply by changing strip size from the worst value to the best one for a given configuration. Because the optimal strip size depends on a number of factors, including number of disks, disk types, raid configuration, and the predominant read/write block size and pattern in an application, often the only way to determine the best setting for a given environment is through experimentation. The SmartArray controller supports six different strip sizes between 32 KiB and 1024 KiB. To determine the optimal setting for this configuration, each possible configuration was tested with Veeam using the same workload. The best performance was obtained with a strip size of 128 KiB.

File System and block size

As described above, ReFS was used because of the advantages it provides over the more traditional NTFS. With the release of Veeam Availability Suite 9.5, Veeam recommends using ReFS due to the advanced integration and performance enhancements[1]. In testing, it was determined that an allocation unit size of 64K produced the best results.

In summary, the storage of the HPE Apollo 4200 Gen9 server was configured as described above using the Smart Storage Administrator utility, as shown in Figure 3:

- RAID-60 (12+12 drives) plus 1 hot spare using 12 TB, 7200 RPM SATA disks
- Strip size = 128 KiB
- Assign all the controller cache to Write

[1] Advanced ReFS Integration with Veeam Overview, https://www.veeam.com/wp-availability-suite-advanced-refs-integration.html

Note

The SmartArray model used in this solution supports write-back cache. In case of power loss, the super capacitor on the SmartArray adapter allows the cache data to be written to a flash module until power is restored and data can be flushed to disk to maintain data integrity.

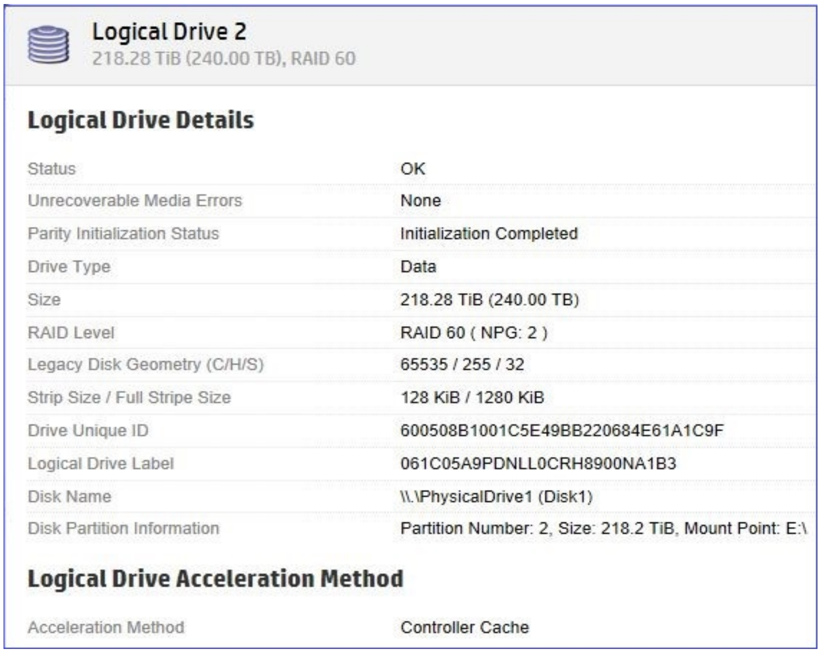

Figure 3. Logical volume properties

After the storage on the HPE Apollo server has been configured as described above and the new volume made accessible to the OS, create a new disk using the maximum capacity available, a file system of ReFS, and an allocation unit size of 64K, as shown in Figure 4.

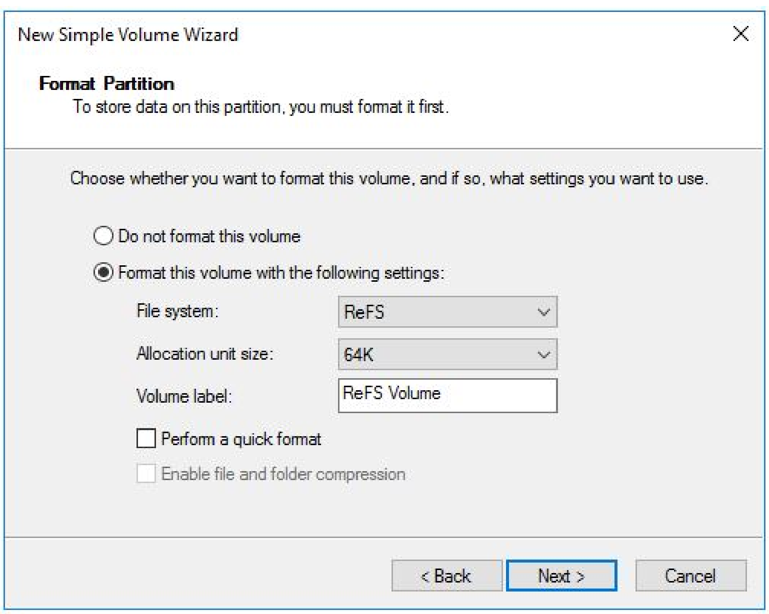

Figure 4. Creating the ReFS file system

Testing was performed with one physical processor and again with two physical processors. In this application, best performance was obtained when the HPE Apollo server was populated with only one physical processor having 18–22 cores. Memory should be a minimum of 128 GB, although 192 GB would be optimal.

When configuring the Fibre Channel fabric, use at least two 16 Gb ports on each device connected to the fabric, and enable and properly configure Windows Multipath I/O (MPIO) on the Veeam server. During backup tests with the solution configured this way, both 16 Gb Fibre Channel ports were fully saturated. Having only a single 16 Gb Fibre Channel path, or one or more paths in an infrastructure with a slower speed such as 8 Gb, will impair the performance of this solution.

Repository configuration

When creating the backup repositories in Veeam, use the following settings.

In the **New Backup Repository** wizard, specify the type of server and other parameters based on where the repository will be located. Since the repository will be ReFS in local storage on the HPE Apollo 4200 Gen9 server, select the type as **Microsoft Windows Server**, select the HPE Apollo server as the **Server**, and specify the appropriate path to the storage volume on the server.

On the **Repository** page of the wizard, deselect the default setting of **Limit maximum concurrent tasks to:**, as shown in Figure 5. This defaults to a value of 4, but no limit is needed. During testing of this solution, the highest backup performance was measured with all repository types when the number of concurrent tasks was only limited by the available resources of the Veeam server.

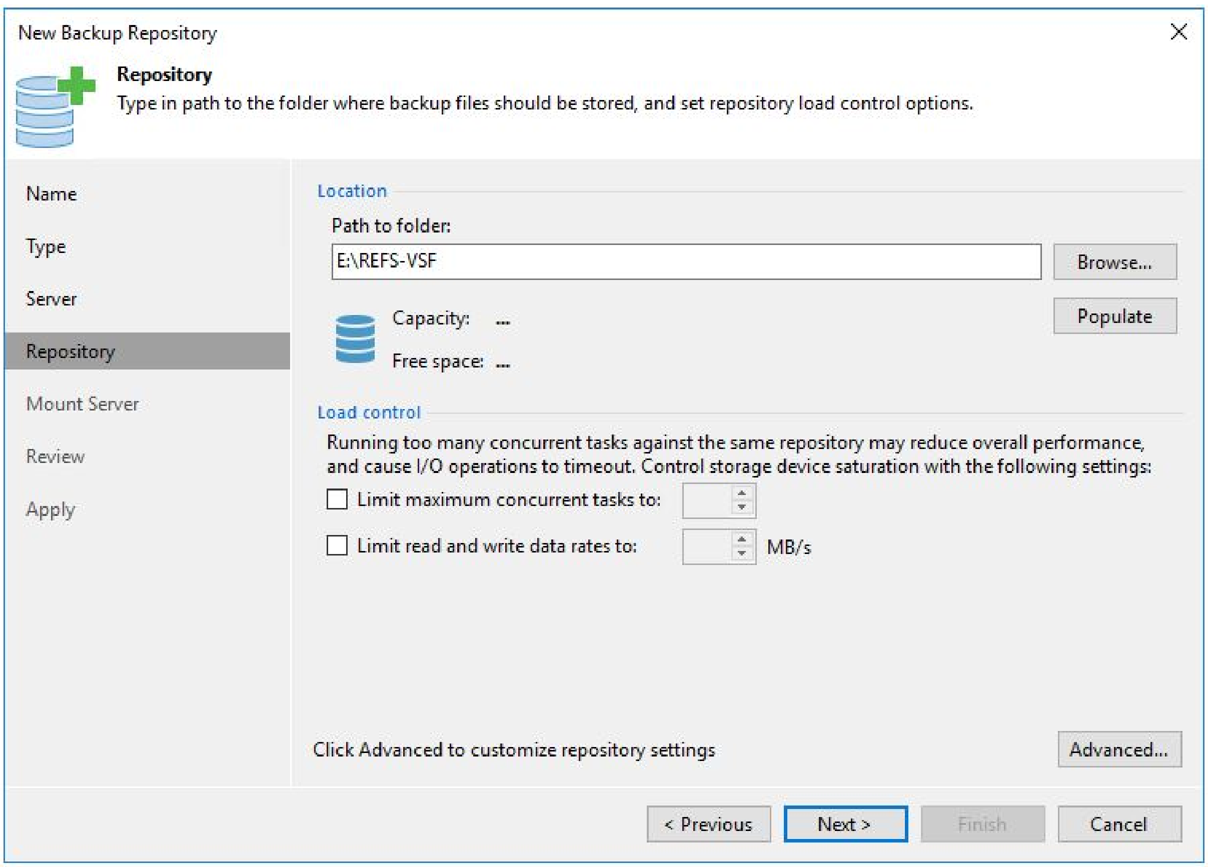

Figure 5. Veeam repository concurrent task setting

On the **Repository** page of the **New Backup Repository** wizard, click **Advanced…** and then select **Use per-VM backup files** (see Figure 6). This allows Veeam to process each VM as an independent thread. In our testing, this configuration produced higher performance than testing with the checkbox disabled.

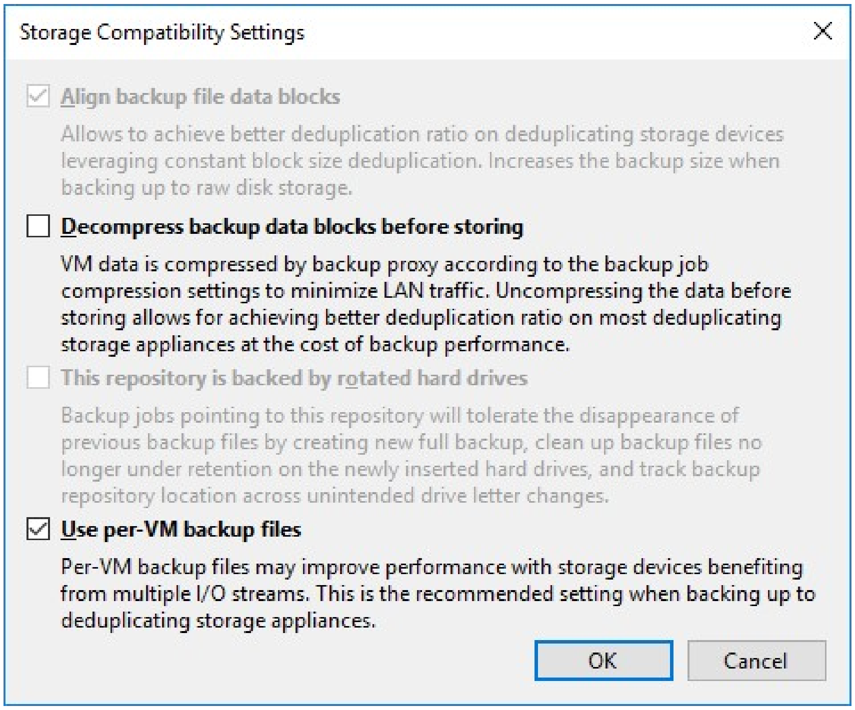

Figure 6. Veeam Repository per-VM backup files

Backup job configuration

When creating a new backup job, launch the **New Backup Job** wizard and follow the prompts to select the devices to be backed up. On the **Storage** page of the wizard, click **Advanced.** In the **Advanced Settings**, click the **Backup** tab, and then ensure the default settings are selected for both **Incremental** backup mode and **Create synthetic full backups periodically**, as shown in Figure 7. This allows Veeam to use block cloning for synthetic full backups as described above.

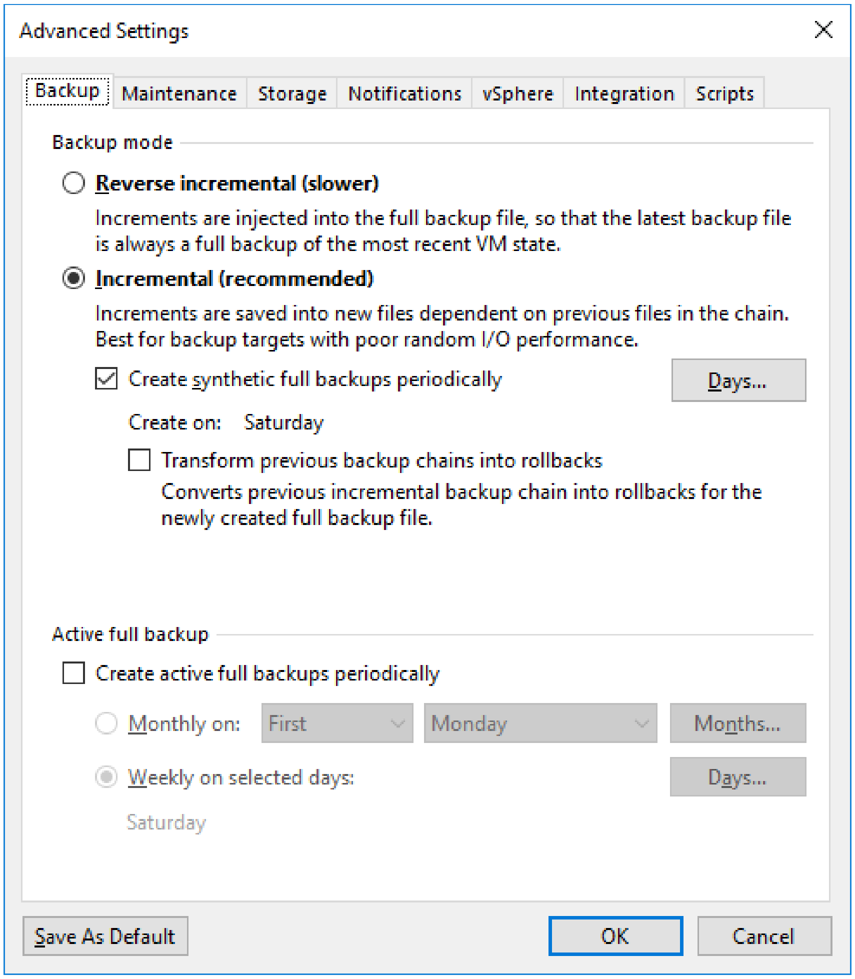

Figure 7. Veeam Backup Job "Advanced Settings"

On the **Integration** tab of the **Advanced Settings** window, select the default setting **Enable backup from storage snapshots** under **Primary storage integration**, as shown in Figure 8. This takes full advantage of Veeam's integration with HPE 3PAR storage snapshots as described above.

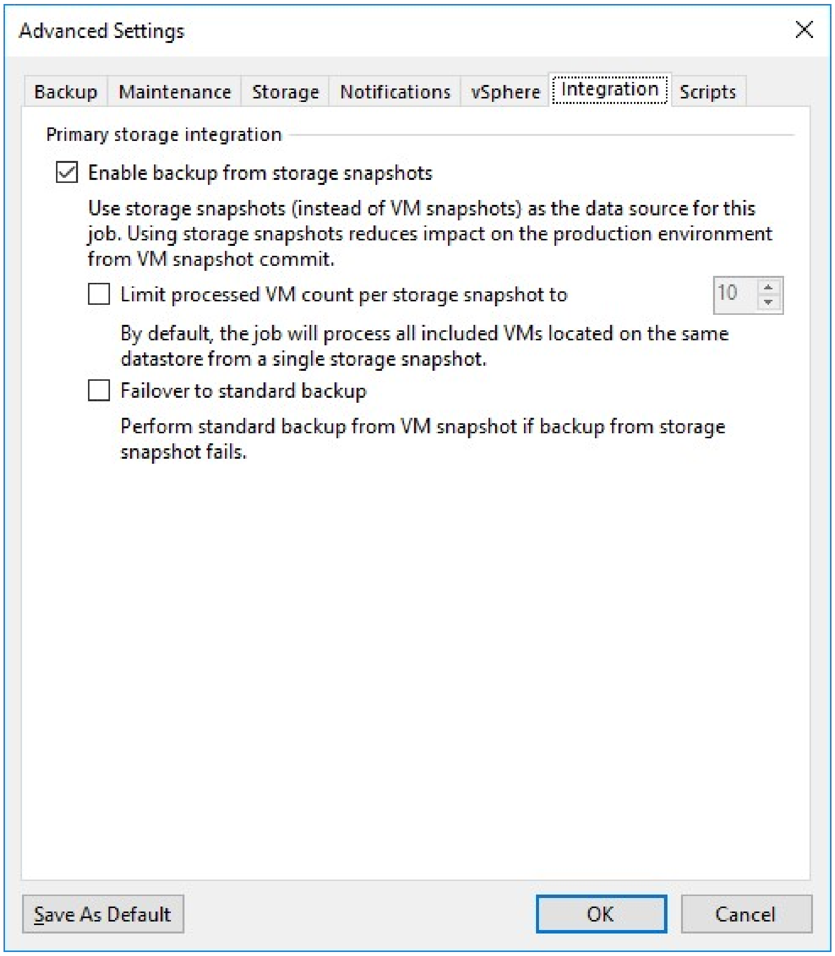

Figure 8. Veeam Backup Job "Integration" settings

If the job being created is for backups to a target repository of ReFS on the HPE Apollo server, click the **Storage** tab of the **Advanced Settings** window, and select the default **Enable inline data deduplication** and **Compression level: Optimal** settings, as shown in Figure 9.

For all jobs regardless of backup target, **Storage optimization** on the **Storage** tab (see Figure 9), should be set to **Local target (16TB+ backup files)**. With this setting, Veeam uses a block size of 4 MB when writing to the target repository. In testing, it was determined that selecting this option produced the best backup performance regardless of the capacity or characteristics of the target storage. Although this setting increases backup throughput, it can slightly reduce the random I/O performance of Instant VM Recovery (IVMR).

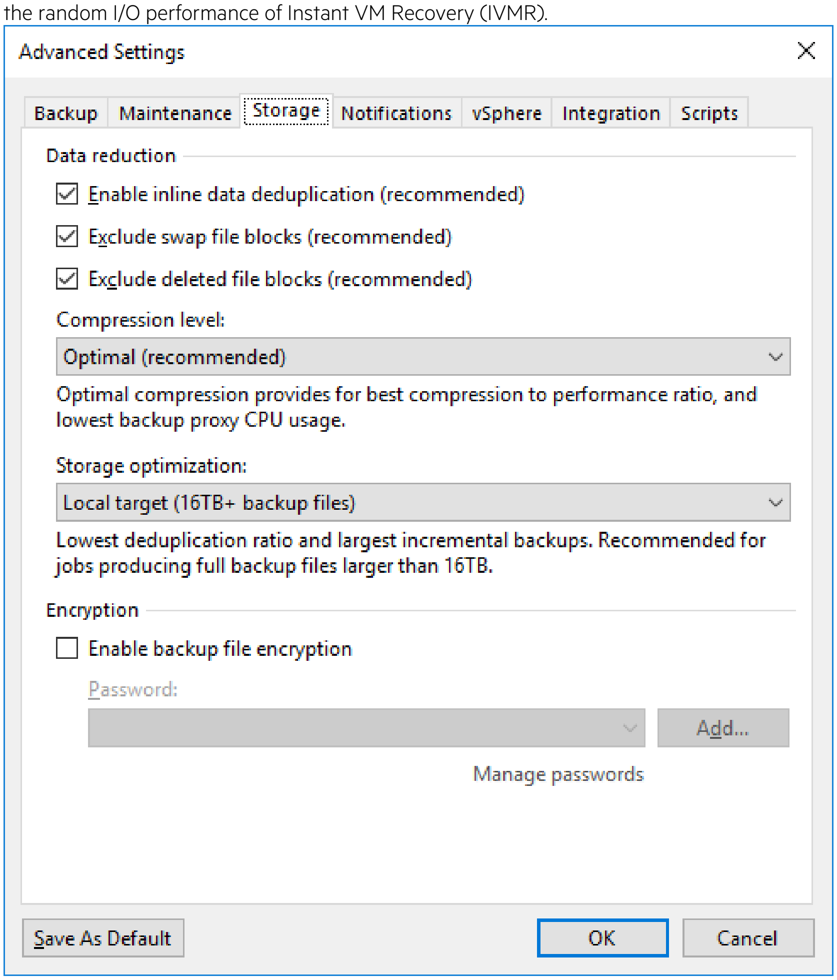

Figure 9. Veeam Backup Job "Storage" settings for HPE Apollo-based repositories

HPE Apollo and Veeam use cases

For multistream backup and restore, ReFS on the HPE Apollo server performs well as a backup target repository. ReFS on the HPE Apollo server excels as a primary target when used for single-stream backups and restores. This backup target also yields outstanding performance when restoring a VM with multiple VMDKs, or when doing single or multiple Instant VM Recovery (IVMR).

Veeam includes advanced features such as Data Labs and IVMR that require performance characteristics different from what is typical in traditional sequential backup and restore workloads. In particular, these features produce large volumes of random I/O over extended periods of time. In many cases, Data Lab and IVMR workloads can be directed to primary storage hardware snapshots. An ReFS repository on the HPE Apollo server is easily able to handle the extended random I/O associated with these workloads.

Multistream throughput

A total of 18 VMs were backed up using multistream and single-stream backup jobs. When a multistream backup was used, the processing rate reported by Veeam was 3 GB/s, as shown in Figure 10. It is important to note that the configuration tested used SATA disks in the HPE Apollo server, which have slower access times than is typical in SAS disks. Even with these “slower” disk drives, the configuration was able to sustain a steady Veeam backup throughput higher than 3 GB/s.

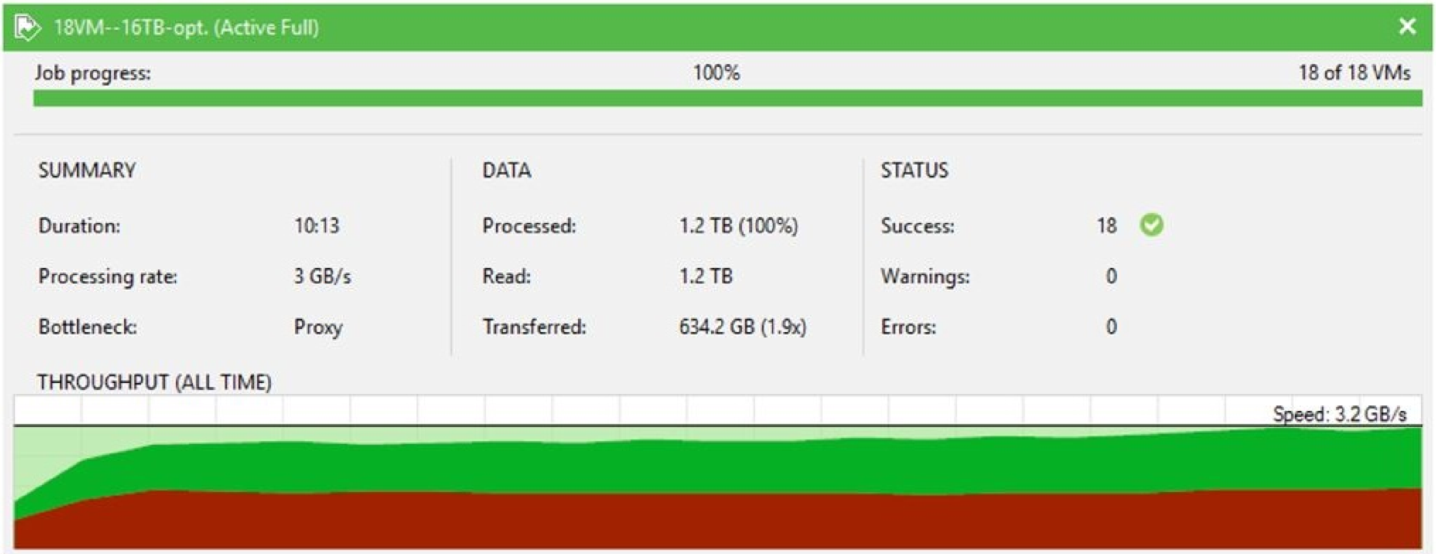

Figure 10. Multistream backup throughput

Single-stream throughput

A full backup was repeated for only one of the 18 VMs. This resulted in the Veeam job having only one VM to process, and thus a single write stream to the Veeam backup repository. For the single-stream backups, the processing rate reported by Veeam was 2 GB/s for the single-stream backup written to ReFS on local storage, as shown in Figure 11.

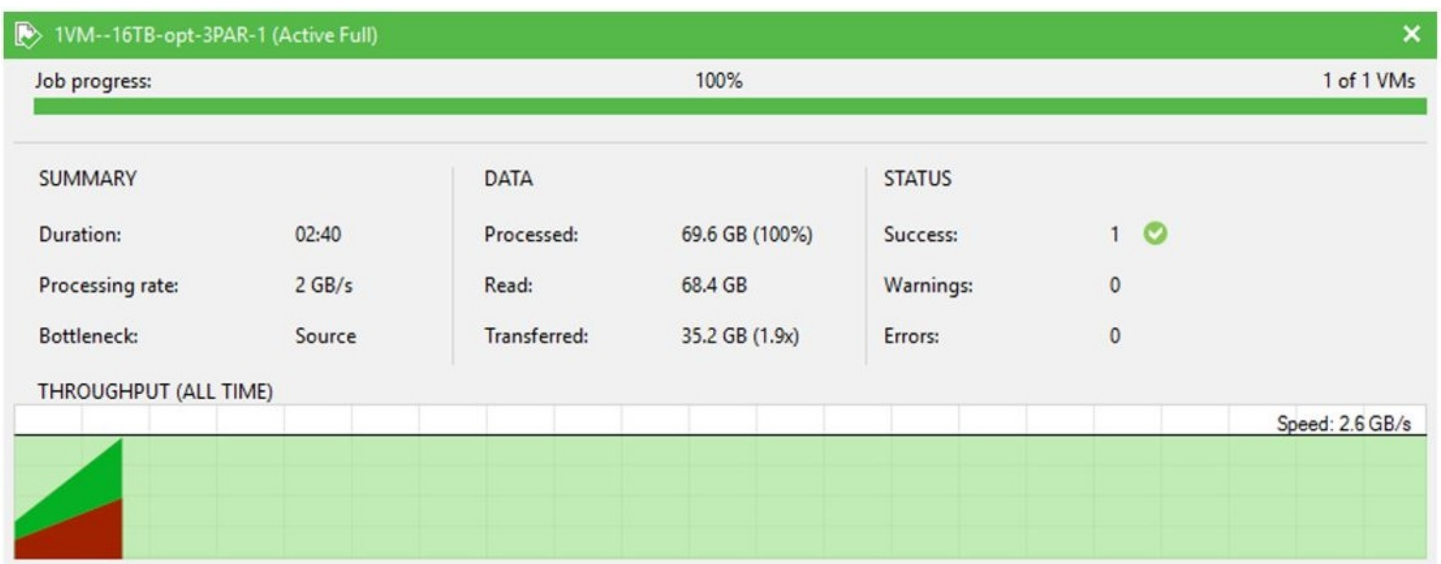

Figure 11. Single-stream backup throughput

Capacity utilization and deduplication

The configuration used in testing this solution contained 18 VMs with the same OS. This produced the potential for significant deduplication across the total backup data set. A similar advantage in deduplication is obtained in any environment where multiple VMs of the same OS are backed up in the same Veeam backup job. The data contained on the test VMs shared common characteristics, although each data set was unique. After every full or incremental backup, 2% of the data set on each VM was changed.

Figure 12 summarizes the different storage utilization of the same backup data written across the different deduplication and storage reduction technologies tested. As can be seen by the turquoise line, the total user data backed up at the end of the 56-day period was just over 25 TB. The purple line shows the storage utilization using Veeam's built-in deduplication and compression, but without the benefit of the Virtual Synthetic Full technology implemented in ReFS. This is the capacity required when the backups are written to an NTFS repository. The orange line indicates the storage utilization when Veeam's deduplication and compression are used with the additional space-saving features of ReFS.

The actual capacity consumed by eight weeks of backup data in the ReFS repository on the HPE Apollo server was slightly more than 8 TB. Without the benefit of Virtual Synthetic Full backups made possible by the block cloning features of ReFS, the capacity used on the HPE Apollo server grew to almost 12 TB.

As highlighted in these test results, Veeam backups to an ReFS repository using Virtual Synthetic Full backups requires approximately one third less capacity than writing the same backups to an NTFS repository. In addition to this important space savings, the virtual synthetic full backup process requires less time and fewer resources than a traditional full backup or other backup policies supported by Veeam, such as Reverse Incremental or Incremental Forever.

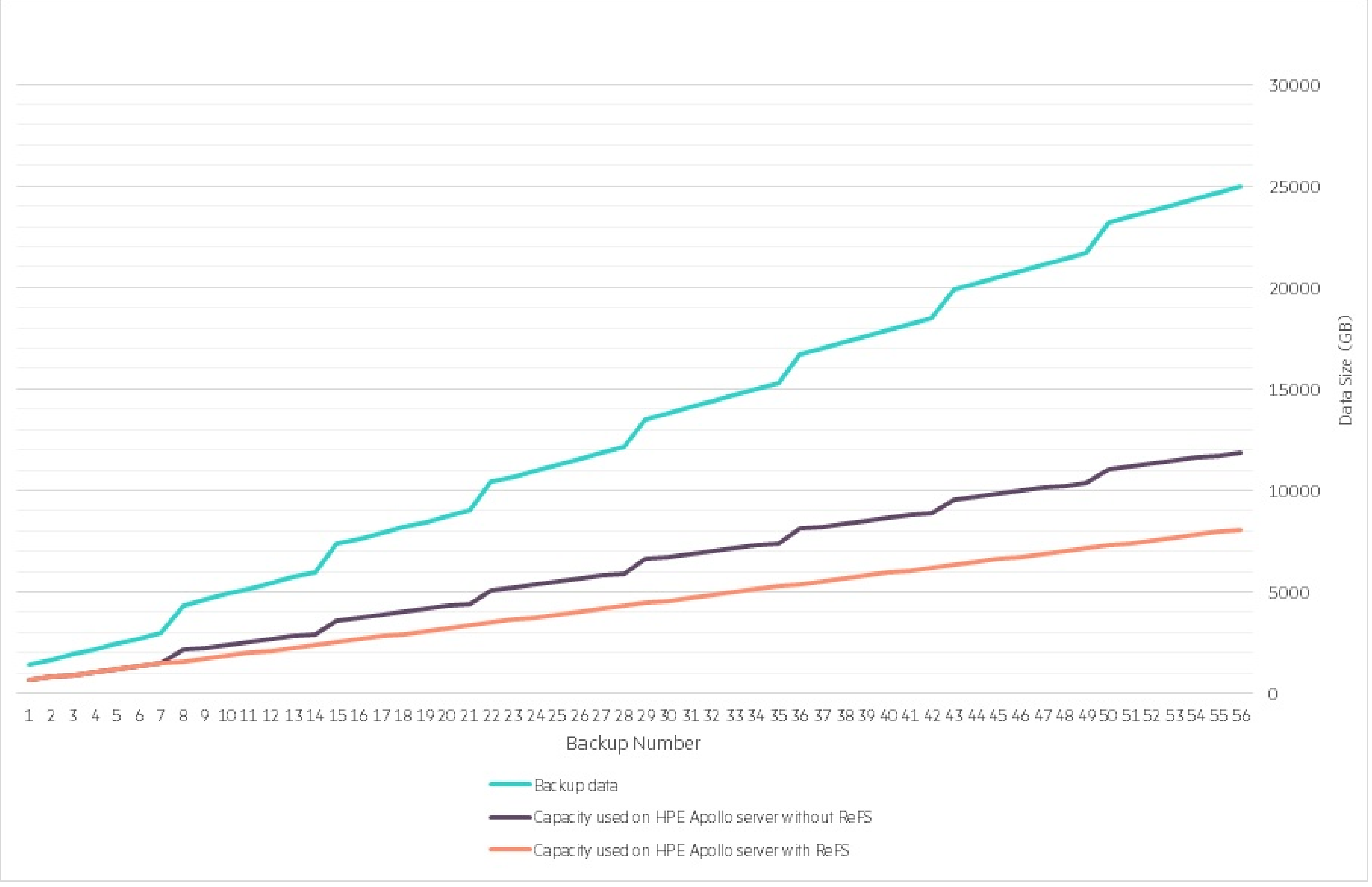

Figure 12. Size of data stored on the HPE Apollo server

Capacity and sizing

Environment sizing

The number of VMs backed up in this testing was chosen to match the number of logical processors in the CPU of the HPE Apollo server. Optimum performance is obtained when each processor core processes a single thread, with each VM making up one thread for processing. Furthermore, the fastest Veeam backup performance is achieved with only one physical processor installed. The size of the VMs, data set, and rate of change were chosen to represent a typical customer-like environment.

In a production environment with hundreds of VMs, Veeam will process one VM per logical processor concurrently, with the rest being queued. As soon as the backup is completed for one VM, a new VM is moved from the queue to the available processor, until all VMs have completed the backup process.

HPE Apollo 4200 Gen9 server capacity sizing guide

Two configurations are presented in Table 1, based on varying capacity. The system can scale up data-in-place by adding more disks of the same type. The RAID set and file system can then be expanded accordingly. Other configurations can be built as well, leveraging the best practices documented elsewhere in this paper and in Table 1.

Table 1. HPE Apollo 4200 Gen9 server configurations

HPE Apollo 4200 Gen9 server	Half-capacity	Maximum-capacity
Description	Half-capacity	Full-capacity
Number of SSDs for OS	2	2
Number of 12 TB HDDs	12	24
Usable capacity	120 TB (109 TiB)	240 TB (218 TiB)
CPU	1 socket, 18–22 cores.	1 socket, 18–22 cores
Memory	128 GB	192 GB
Networking	10 GbE	100 GbE

The solution described in this document is based on Large Form Factor (LFF) 12 TB SATA HDDs and is adequate for most workloads. However, some Veeam features such as SureBackup, Virtual Lab, and Instant VM Recovery (IVMR) produce a very high random I/O workload on the HPE Apollo server. If you plan to make extensive use of these features, it might be necessary to use different storage configurations based on a larger number of HDDs with smaller capacities, or in extreme cases, all SSDs. To increase the number of installed disks it is possible to configure the HPE Apollo server with Small Form Factor (SFF) disks. This doubles the maximum number of disks per system from 28 LFF to 56 SFF disks.

To lower the initial cost of the solution, the HPE Apollo server can be configured with the smaller number of disks described in Table 1 (12 disks). As your need for capacity grows, the solution can easily be scaled up, preserving existing data, by adding more physical disks, new RAID groups, and new ReFS volumes. It is also possible to expand the existing RAID set and ReFS volume, but that operation could take a considerable amount of time. If a data-in-place upgrade is not required, it is possible to re-create the single existing volume and ReFS file system to take advantage of the additional capacity.

Workload

The workload used in this testing consisted of 18 VMs, each 100 GB in size and set up similarly. The OS on each VM was Windows Server 2016, and consumed approximately 23 GB. The data set on each VM was 50 GB in size, with the remaining 27 GB of space on each VM free. The 50 GB data set was unique on each VM, though they all shared common characteristics. Each data set contained approximately 112,200 files of multiple sizes with many small files and fewer large ones. Each file contained strings of randomized binary data, with strings repeated twice in each file to allow for a 2:1 compression ratio, which is typical of customer data. After each full or incremental backup, 2% of files on each VM were modified.

The backup cycle implemented was a weekly full backup, followed by six daily incremental backups. The cycle was then repeated for a simulated eight weeks of backups. After the initial full backup, all remaining weekly full backups were implemented as Virtual Synthetic Full backups.

Analysis and recommendations

The HPE Apollo 4200 Gen9 storage server has HPE ProLiant-like compute and high-density storage in the same chassis, making it an ideal server to host all the Veeam components on one platform. The Veeam Server and SQL Server database can be installed on the HPE Apollo server, as well as the Veeam Proxy, Backup Repositories, and even a Veeam Tape Server if that is required. All of these components could be installed on an HPE ProLiant server as well, but because HPE ProLiant servers are not as storage-dense as the HPE Apollo servers, an HPE ProLiant-based solution would take up much more physical space than an HPE Apollo-based solution of comparable capacity. For simplicity and efficient space utilization, Hewlett Packard Enterprise recommends using the HPE Apollo 4200 Gen9 server.

Impact to the production VMs is minimized due to the close integration of HPE 3PAR and HPE Nimble Storage snapshots with Veeam Availability Suite. The HPE storage snapshot is used for the backup, thus reducing the impact of data protection on the production VMs.

Because Veeam makes extensive use of array-based hardware snapshots, HPE recommends grouping all VMs in a volume in the same backup job in Veeam. In this way, each job requires only one hardware snapshot instead of many. This streamlines the resource utilization on the storage array, allowing data protection activities to run as efficiently as possible with minimal impact on the VMs.

Summary

As the volume of data owned and managed by businesses continues to grow, so does the demand to keep that data protected. The ever-present possibility of data loss and external threats have made the need for fast and reliable data recovery essential. The HPE Apollo 4200 Gen9 storage server paired with Veeam Availability Suite offers a robust solution for this data-driven world. The combination of HPE ProLiant-like compute ability and high-density storage in the same chassis makes the HPE Apollo storage server the clear choice for implementing a simple and affordable, yet powerful, backup solution for virtualized environments. In addition, Veeam Availability Suite can now leverage HPE 3PAR and HPE Nimble Storage snapshots to significantly reduce the impact of data protection on production environments.

The solution highlighted in this Reference Architecture is an efficient, affordable, reliable, and secure backup and recovery infrastructure for data protection in virtualized environments. Using Hewlett Packard Enterprise award-winning servers, storage, network infrastructure, and support services will keep your mission-critical applications functioning at their best, while protecting them from the worst.

Appendix A: Bill of materials

Two different BOMs are presented below. The first is for a half-capacity configuration, containing the minimum recommended memory, networking, and disk capacity. The second BOM is for a maximum-capacity configuration, containing more memory, faster networking, and higher overall capacity. Compute power is the same for each configuration. The main functional difference is the amount of data each configuration is able to back up. A summary of each configuration is shown in Table 1.

The following BOMs contain electronic license to use (E-LTU) parts. Electronic software license delivery is now available in most countries. Hewlett Packard Enterprise recommends purchasing electronic products over physical products (when available) for faster delivery and for the convenience of not tracking and managing confidential paper licenses. For more information, please contact your reseller or a Hewlett Packard Enterprise representative.

Note

Part numbers are at time of publication and subject to change. The bill of materials does not include complete support options or other rack and power requirements. If you have questions regarding ordering, please consult with your Hewlett Packard Enterprise Reseller or Hewlett Packard Enterprise Sales Representative for more details. hpe.com/us/en/services/consulting.html

---

Table 2a. Bill of materials – half-capacity configuration

Qty	Part number	Description
		HPE Apollo 4200 storage server – half-capacity configuration
1	808027-B21	HPE Apollo 4200 Gen9 CTO Svr
1	806563-B21	HPE Apollo 4200 Gen9 Rear HDD Cage Kit
1	830754-B21	HPE Apollo 4200 Gen9 Intel Xeon E5-2699v4 FIO Processor Kit
4	728629-B21	HPE 32GB (1x32GB) Dual Rank x4 DDR4-2133 CAS-15-15-15 Registered Memory Kit
1	665249-B21	HPE Ethernet 10Gb 2-port 560SFP+ Adapter
2	JD096C	HPE X240 10G SFP+ to SFP+ 1.2m DAC Cable
1	813546-B21	HPE SAS Controller Mode for Rear Storage
2	P9D94A	HPE StoreFabric 16Gb Host Bus Adapter
2	846788-B21	HPE 1.6TB 6Gb SATA MU-2 LFF LPC SSD
12	881787-B21	HPE 12TB 6G SATA 7.2K LFF 512e LP HDD
1	806565-B21	HPE Apollo 4200 Gen9 IM Card Kit
1	806562-B21	HPE Apollo 4200 Gen9 Redundant Fan Kit
2	720479-B21	HPE 800W FS Plat Ht Plg Pwr Supply Kit
1	822731-B21	HPE Apollo 4200 Gen9 Hardware Rail Kit

Table 2b. Bill of materials – maximum-capacity configuration

Qty	Part number	Description
		HPE Apollo 4200 storage server – maximum-capacity configuration
1	808027-B21	HPE Apollo 4200 Gen9 CTO Svr
1	806563-B21	HPE Apollo 4200 Gen9 Rear HDD Cage Kit
1	821792-L21	HPE Apollo 4200 Gen9 Intel Xeon E5-2699v3 FIO Processor Kit
6	728629-B21	HPE 32GB (1x32GB) Dual Rank x4 DDR4-2133 CAS-15-15-15 Registered Memory Kit
1	825111-B21	HPE InfiniBand EDR/Ethernet 100Gb 2-port 840QSFP28 Adapter
2	JD096C	HPE X240 10G SFP+ to SFP+ 1.2m DAC Cable
1	813546-B21	HPE SAS Controller Mode for Rear Storage
2	P9D94A	HPE StoreFabric 16Gb Host Bus Adapter
2	846788-B21	HPE 1.6TB 6Gb SATA MU-2 LFF LPC SSD
24	881787-B21	HPE 12TB 6G SATA 7.2K LFF 512e LP HDD
1	806565-B21	HPE Apollo 4200 Gen9 IM Card Kit
1	806562-B21	HPE Apollo 4200 Gen9 Redundant Fan Kit
2	720479-B21	HPE 800W FS Plat Ht Plg Pwr Supply Kit
1	822731-B21	HPE Apollo 4200 Gen9 Hardware Rail Kit

Resources and additional links

HPE Apollo 4000 systems, hpe.com/us/en/servers/hpc-apollo-4000.html

Veeam Backup & Replication 9.5 User Guide for VMware vSphere, https://helpcenter.veeam.com/docs/backup/vsphere/overview.html?ver=95

- Creating Backup Jobs, https://helpcenter.veeam.com/docs/backup/vsphere/backup_job.html?ver=95
- Backup Copy, https://helpcenter.veeam.com/docs/backup/vsphere/backup_copy.html?ver=95
- Backup Proxy, https://helpcenter.veeam.com/docs/backup/vsphere/backup_proxy.html?ver=95
- Data Recovery, https://helpcenter.veeam.com/docs/backup/vsphere/data_recovery.html?ver=95
- Instant VM Recovery, https://helpcenter.veeam.com/docs/backup/vsphere/instant_recovery.html?ver=95
- GFS Retention Policy, https://helpcenter.veeam.com/docs/backup/vsphere/backup_copy_gfs.html?ver=95

Veeam Backup & Replication 9.5 PowerShell Reference: Getting Started, https://helpcenter.veeam.com/docs/backup/powershell/getting_started.html?ver=95

Veeam Backup & Replication 9.5 User Guide for Microsoft Hyper-V: Data Compression and Deduplication, https://helpcenter.veeam.com/docs/backup/hyperv/compression_deduplication.html?ver=95

Backup Copy – Animation of Retention, https://www.veeam.com/kb1931

Hewlett Packard Enterprise Technology Consulting Services, hpe.com/us/en/services/consulting.html

To help us improve our documents, please provide feedback at hpe.com/contact/feedback.

Learn more at hpe.com/storage

a00064646enw, April 2019, Rev 1