Dell PS Series Snapshots and Clones: Best Practices and Sizing Guidelines

Dell Storage Engineering
November 2019

Dell EMC Best Practices

Revisions

Date	Description
May 2012	Initial release
December 2016	Minor updates
November 2019	vVols branding update

Table of contents

Executive summary

Today's storage administrators are tasked with ever evolving challenges in ensuring the storage resources being requested of them by their customers and their customers applications is resilient from recovery actions and highly available when needed while being aligned to the organizations RTO/RPO (Recovery Time Object/Recovery Point Object) policy.

This best practices document describes the benefits of using snapshots on Dell™ PS Series volumes for data recovery to achieve RTO/RPO policy goals as well as the advantages of cloning volumes to ensure data integrity of the backups.

1 Introduction

The lab validated best practices in this paper will help IT and SAN administrators understand and fully utilize the powerful snapshot and clone features delivered with every PS Series array.

This paper describes:

- The differences between snapshots and clones and where each can be used for data protection and recovery
- How to properly and efficiently size snapshot reserve space
- How to invoke a snapshot or clone a volume
- How to restore from a snapshot or clone of a volume
- Performance implications when using snapshots and clones
- Best practices for using snapshots and clones

The results presented in this paper will help determine specific needs for any SAN environment.

1.1 Audience

This white paper is for storage administrators who are involved in the planning, implementation, configuration, and administration of PS Series iSCSI storage area networks and the use of snapshots and clones for the purpose of data recovery and availability.

1.2 Terminology

The following terms are used throughout this document.

Base volume: A volume that a snapshot initially shares data with or a clone volume is copied.

Block: A unit of data access specified during an input/output (I/O) operation.

Clone: A full copy of a volume.

Collection: A logical grouping of 1-8 volumes that allows the same operation (such as a snapshot or clone replica) to be performed on all volumes in the collection simultaneously.

Page: A logical grouping of data segments within a PS Series volume, snapshot, clone volume or replica.

Pool: A collection of storage space comprising of one or more members. Volumes are then created from free space within an available pool.

Replica: A point-in-time copy of a volume placed on another PS Series array group using auto replication.

Snapshot: The storage feature that preserves the contents of a volume at a point in time.

Template volume: A volume that is designated as read only and then used as a source for creating thin clones

Thin clone: A space efficient copy of a template volume that only stores its unique changes.

Note: For additional details and definitions see the *Dell PS Series Configuration Guide*.

2 PS Series storage

With its unique peer storage architecture, PS Series arrays deliver high performance and availability in a flexible environment with low cost of ownership. PS Series storage solutions deliver the benefits of consolidated networked storage in a self-managing, iSCSI storage area network that is affordable and easy to use, regardless of scale. By eliminating complex tasks and enabling fast and flexible storage provisioning, these solutions dramatically reduce the costs of storage acquisition and ongoing operations.

Patented page-based volume management enables automatic movement of data while it is in use. This technology provides the foundation for online expansion, automatic configuration and load balancing, performance optimization, and advanced software functionalities — all with continuous access to data. The result is that there is no downtime required for increasing capacity, moving data between storage tiers, or load balancing storage. In addition, most management tasks are handled by the array, not the administrator. PS Series arrays make enterprise-class shared-block storage practical for all servers and applications.

Every PS Series array includes additional enterprise class features such as snapshots, clones, and replication at no additional cost. The snapshot feature enables quick recovery of files, the clone feature may be used for creating copies of data or for recovery of files or volumes, and the replication feature allows the implementation of disaster recovery initiatives.

3 Protecting data with snapshots and clones

3.1 Snapshots

Snapshots are point-in-time copies of volumes that capture the contents of a volume at a specific point in time and are often used to recover data lost by events such as human error, viruses, or database corruption. They can also be used for testing or to create a source for backup to tape or another disk. The creation of a snapshot is done without disrupting normal host access to the volume. Snapshots have some features and properties similar to the base volume as well as some unique capabilities.

When a snapshot is created, it does not consume any space, but instead is only a set of pointers to the data in the base volume. As data is modified on the base volume, disk space is allocated from the snapshot reserve to store the changes. Meanwhile, the snapshot still points to the original data pages so that the volume looks exactly like it did at the point in time when the snapshot was taken.

Like volumes, a snapshot can be assigned an iSCSI Qualified Name (IQN) and can be presented as a volume to a host. This allows a host to mount a snapshot, read or modify the data in the snapshot, or create a full volume (clone) that has dedicated space allocated to it from the free storage pool.

Figure 1 illustrates a view of a volume with multiple snapshots over a period of time.

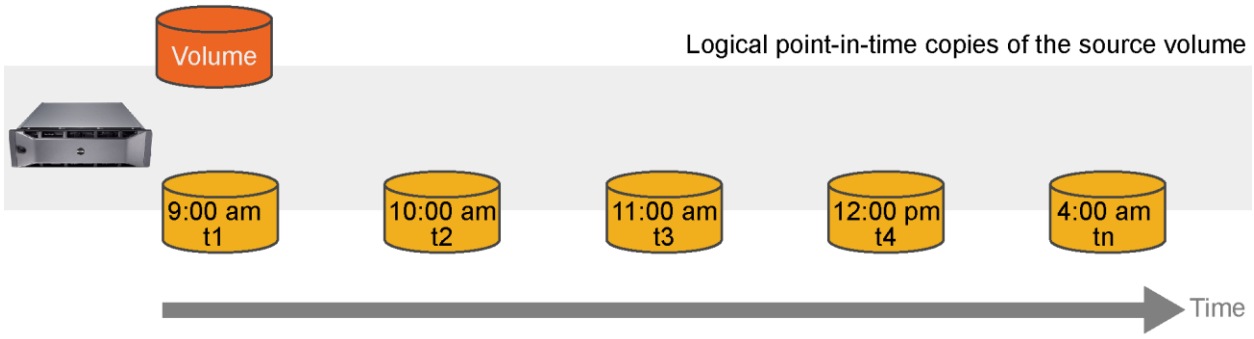

Figure 1 Multiple snapshots over time

3.2 Clone volumes

A clone volume is a full copy of an existing volume. The clone volume has the same reported size, contents, and thin-provision settings as the original volume. A clone volume can be created from a regular volume, a specific replica of a volume, or a specific snapshot of a volume. A thin clone can also be created from a template volume. Thin clones are sub-volumes of template volumes where the template is a read-only version of the volume at a specific point in time. Clone volumes are often used for data protection purposes; either to enable fast recovery of files or as a source for backup to tape or disk-to-disk backup operations. Clone volumes can also be used for testing and development purposes with less impact to the production environment.

The next figure shows a base volume with three clone volumes from different points in time. Each clone volume is a copy of the base volume, but completely independent of the base volume.

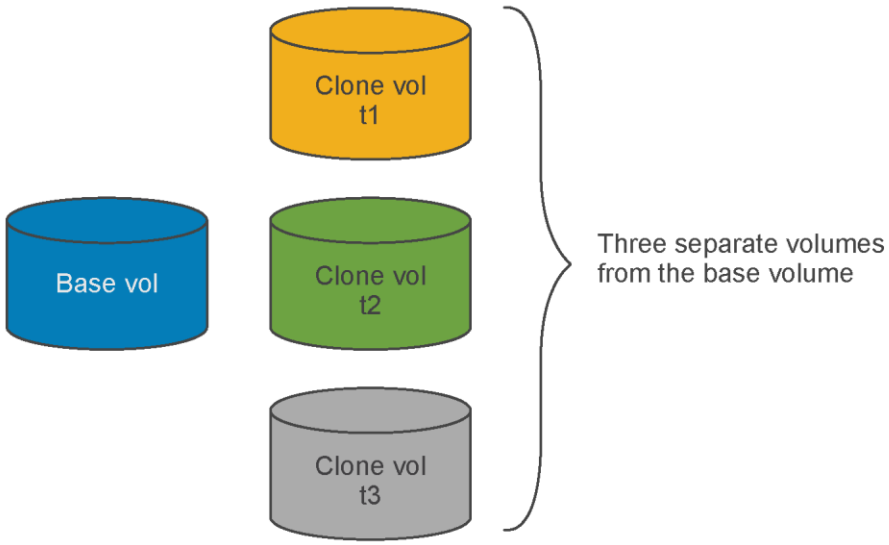

Figure 2 Independent point-in-time images of the source volume

3.3 Differences between PS Series snapshots and clone volumes

Table 1 Differentiating factors between snapshots and clones:

	Snapshot	Clone Volume
Definition	Point-in-time copy of a volume	Full copy of a volume
Location	Base volume and snapshots always reside in the same storage pool	Once a clone volume is created, it can be moved to any pool that contains adequate free space
Functionality	When a snapshot is created, it is instantly available, and no data is moved or copied	When a clone volume is created, it is instantly available to the user. All regular volume operations can be performed on the clone volume once the cloning operation is complete.
Deletion	Deleting a snapshot's parent volume deletes all associated snapshots for that volume	Deleting the source volume where a clone volume was created does not delete a clone volume.
Association with the source volume	A snapshot does not initially consume any space but instead is a set of pointers to the data in the base volume. As data is modified on the base volume, disk space is allocated from the snapshot reserve to store the modified data.	A clone volume creates a full copy of the original volume, consuming 100% of the original volume size from free space in the pool where the original volume resides. A clone volume uses unique, non-shared blocks (or pages) of data.
Use Cases	• Short-term local data protection: Rapid restores of data from multiple recovery points in case of a file corruption or accidental deletion. • Backup and Recovery: Offload backup operations to another server. • Additional copies of data: Provide other servers secure access to a copy of production data	• Test and Development: Independent full copy of the source volume can be used for testing and development purposes. • Data Reporting and Data warehousing: Clones can be used for running end of the month/quarter/year reporting. • Compliance and Legal hold: A read only copy of a clone can be used to store data history for compliance and legal hold purposes.

3.4 Storage requirements and configuration limits

In a PS Series group, snapshots can be used to protect data against human error, viruses, or database corruption. PS Series storage administrators can create a point-in-time copy or multiple copies of the base volume, which can be retained for data protection or made accessible to another host. Creating snapshots can be performed while the base volume remains online, therefore, users and applications are not disrupted.

Before creating a snapshot of a volume, the administrator must allocate snapshot reserve to hold the snapshot data. Snapshot reserve space is always consumed from the same storage pool as the volume. By having the administrator decide how much space to reserve, the storage keeps track of the commitment. This allows the administrator to avoid having to manually estimate the free space that may be quickly consumed. The PS Series Group snapshot reserve space, allocated by default, is equal to 100% of the host volume allocated space. This ensures that a 100% data change in the volume can be protected by a single snapshot. This value can, and often is, easily set to a lower value based on the application data change rate, recoverability requirements, or role that the snapshot will be used for.

The amount of required snapshot reserve space depends on:

- The number of writes that occur to the base volume during the life of the snapshot. In general, initial changes to the base volume will result in the utilization of more snapshot reserve space, while subsequent writes to logical blocks that have already been modified do not require additional space.
- The range of logical blocks where the data change occurs. If the same blocks of data are consistently modified, then snapshot reserve is consumed only when the first change is made.
- The number of simultaneous snapshots in use. More snapshots tend to use more space since multiple points in time must be preserved.
- The life span of the snapshots. Longer-lived snapshots tend to use more space since more data is likely to change over a longer time period.

Even though snapshot reserve space is not allocated until needed, it is immediately subtracted (or reserved) from the pool and reflected in the available free space. This guarantees that the snapshot reserve space will be available to hold any new updates to the volume as well as preserve the point-in-time view of the volume. In addition, the system will automatically delete older snapshots as necessary to stay within the reserve usage limit set by the administrator. This method reflects a more accurate representation of space that is available for user data. An administrator can adjust or change a volume snapshot reserve at any time through the GUI or CLI management utilities.

The following table shows the supported configuration limits for arrays running the 9.0.x firmware release. Configuration limits for groups of PS4XXX class arrays are shown separately. Refer to the release notes of the current firmware for the most current limits and other important information.

Table 2 Supported configuration limits

Configuration	PS4XXX groups	All other groups
Snapshots + replicas + VMware® vSphere® Virtual Volumes™ (vVols) per group[1]	2048	10,000
Snapshots per volume	128	512
Snapshot schedules per volume or collection	64	64

[1] This value includes vVols, vVol snapshots, and vVol-linked clones.

3.5 Creating a snapshot

In a PS Series storage group, the base volume and all of the snapshots for that volume reside in the same storage pool. Internally, the PS Series group organizes physical storage in logical segments or pages of data.

When a user creates a volume, an internal table is created to track pointers to each page of data that makes up the base volume. When a snapshot of a volume is created, a copy of that table is made for the snapshot volume. At this time, data is not moved or copied and no additional disk space is consumed. Additional disk space (snapshot reserve) is only consumed when there is a modification to the base volume (this can be to user data or metadata). If the base volume is never changed, hundreds of snapshots (up to a maximum of 512) could be created for the volume without requiring additional space.

Figure 3 illustrates how the base volume table and the snapshot table at the time of snapshot creation are pointing to the same data pages.

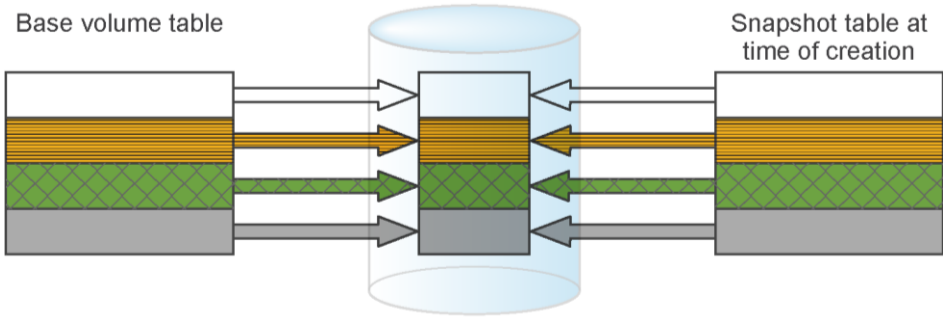

Figure 3 Snapshot table at creation

As the base volume is modified, snapshot reserve space is utilized. The first time a change is made to the logical segment of the base volume that is different from the time the snapshot was created, a new segment is allocated out of the snapshot reserve space. The group then updates the volume page table pointer to point to the new logical segment. However, because a PS Series volume is created from virtual segments of data, it is not necessary to copy or move the original data before it is changed. Instead, a new segment is allocated to hold the changed data and only pointers are updated. This is more efficient than the Copy on First Write (COFW) technique that requires additional disk operations to move the old data out of the way before flushing

new data from cache to the physical disks. Reducing unnecessary workload on the disks reduces latency and results in more available resources for the system to meet an application's requirement.

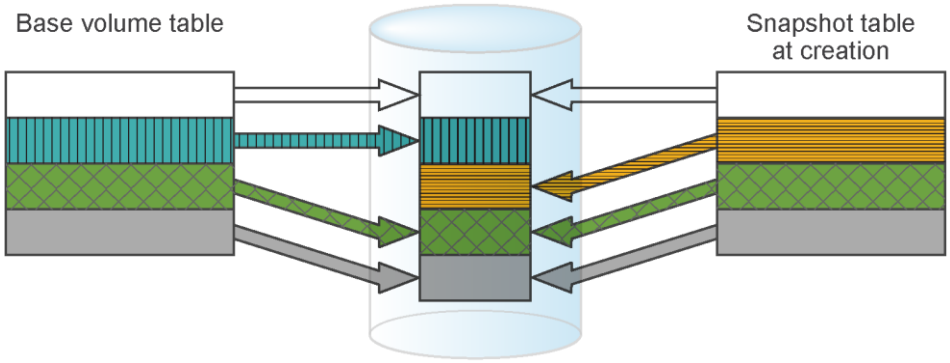

Figure 4 Snapshot changes

Snapshot reserve can never grow beyond 100% of the base volume for a single snapshot. However, if there are multiple snapshots, the reserve usage may be greater than the size of the base volume. This usage depends on the actual change rate and I/O pattern occurring on the base volume and how many snapshots are retained.

The following example explains sharing pages of data between multiple snapshots of a base volume:

1. Assume that the base volume occupies four pages of data. At time 0, the user takes the first snapshot (snapshot 0). Since there are no changes to the base volume, the first snapshot and the base volume point to the same logical page of data (refer to Figure 3).
2. When data changes, for example an application writes new data, the base volume pointers are updated to the new pages and the first snapshot continues to point to the original data (refer to Figure 4).
3. A second snapshot is created (snapshot 1) capturing the changes made after snapshot 0. However, changes made after snapshot 1 will be captured on a subsequent snapshot (Figure 5).

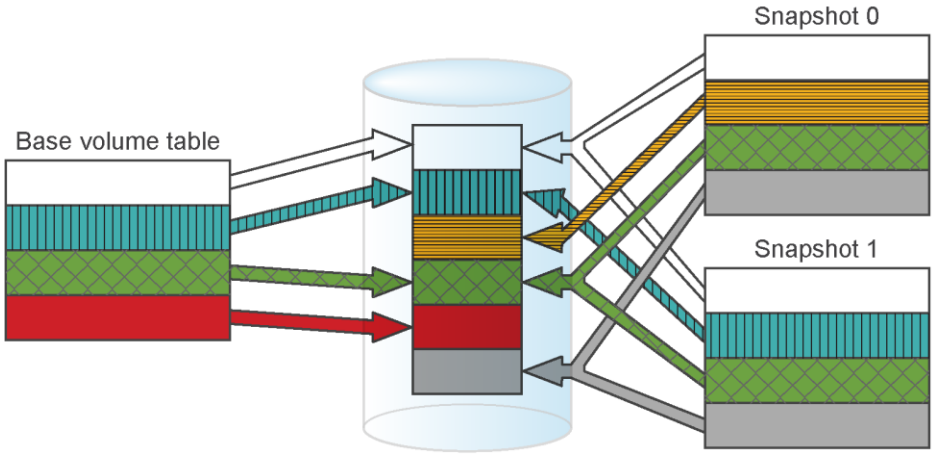

Figure 5 Base volume and snapshot 1

Depending on the number of snapshots and the amount of changes made to the base volume, the snapshot reserve space can exceed the total allocated space. If the snapshot reserve is entirely consumed, the default

action is to delete the oldest snapshot, providing the necessary free space to create a new snapshot. However, the user can set a policy for recovering snapshot space with the following options:

- Delete oldest snapshot (default).
- Set the base volume and its snapshots offline.

If a snapshot has active iSCSI connections, they will be terminated before the snapshot is deleted or before the volume and its snapshots are set offline. Administrators can also monitor snapshot usage and prevent the loss of snapshot data due to automatic snapshot deletion by cloning the snapshot or increasing the snapshot reserve space. While deleting a snapshot does not result in deleting any data from the base volume, it does remove the ability to restore the base volume to that specific point in time.

3.6 Creating a clone volume

When a clone volume is created, it is a complete copy of the base volume. Unlike the snapshot volume, it no longer shares any data with the base volume. Since it is an independent copy of that base volume it does not use snapshot reserve. Initially, a clone must be created in the same pool as the base volume. However, it can subsequently be moved to any pool with adequate free space. If the base volume is a thin provisioned volume, a clone of that volume will also default to be thin provisioned and will initially report the same size and space utilization as the base volume.

3.7 Creating a template volume and thin clones

Any volume can be designated as a template volume. When a volume is marked as a template volume, it is set offline and is read only. A template volume is not intended to be mounted by a host, but instead can be used to quickly create copies (thin clone volumes) with the same starting data set. For example, in a Virtual Desktop Infrastructure (VDI) environment a template volume may be used to create a golden boot image for virtual desktops.

Thin clones can be created from a template volume to use pool space more efficiently. Each thin clone shares common data in the template volume and only uses additional space to store changes in the volume. Initially, a thin clone will report the same size and space utilization as the template volume and its contents will be identical to the template volume. If a host modifies a thin clone, additional space will be allocated up to the maximum setting for the thin clone volume.

3.8 Volume collections

Dell EqualLogic Group Manager allows for the creation of a collection or a grouping of volumes to allow the same action to be applied to more than one volume at a time. For example, an administrator may want to group multiple volumes that contain database and log files for a database server and snapshot them all at the exact same point in time to maintain consistency across the volumes. In a single operation, simultaneous snapshots will be taken of all of the volumes that are part of the collection.

Collections may also be used to group volumes for creating replicas. A collection may contain up to eight volumes from one or more pools, however template volumes are not supported as part of a collection.

3.9 Snapshot schedules

Using Dell Storage Manager, Dell EqualLogic Group manager or Auto-Snapshot manager, an administrator can create a schedule for taking snapshots (or replicas) of a volume or a volume collection. The schedule can execute on an hourly or daily basis and the frequency of snapshots (or replicas) can be set to occur once or at regular intervals (from five minutes to 12 hours apart). The schedule can also set how many snapshots (or replicas) should be retained (assuming there is sufficient reserve space allocated). For example, an administrator may schedule a snapshot to occur every two hours between 7 a.m. and 7 p.m. each day with a maximum of ten snapshots to be retained.

3.10 Dell EqualLogic Auto-Snapshot Manager

Dell EqualLogic Auto-Snapshot Manager/Microsoft® Edition (ASM/ME) is also available for hosts running Windows Server®. ASM/ME is installed as part of the Dell EqualLogic Host Integration Tools (HIT) for Microsoft® and can be used to automate point-in-time, application consistent snapshots, clones, and replicas; also known as Smart Copies. ASM/ME uses the Windows Volume Shadow Copy Service (VSS) to ensure that any I/O is quiesced and any in-flight data is committed to disk before the snapshot or clone volume is created. Without VSS, the only way to guarantee that the data on a volume is consistent, is to dismount the volume (i.e., bring it offline) or shut down the host it is connected to, to ensure all cached data is flushed to the disk.

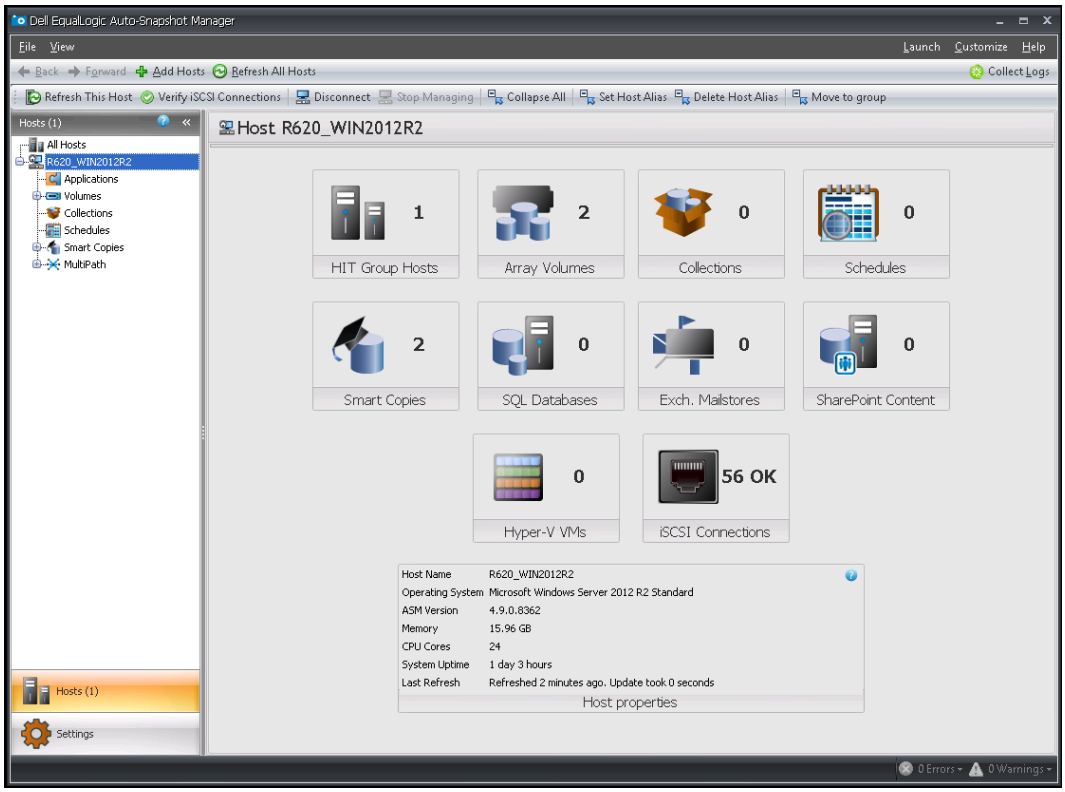

Figure 6 Dell EqualLogic Auto-Snapshot Manager

ASM/ME provides both a graphical user interface and a command line interface for manual or scripted operations. ASM/ME also supports popular Microsoft applications such as SQL Server®, Exchange®, SharePoint®, and Hyper-V virtual environments.

Besides single volume operations, ASM also supports collections. Snapshot, clone, and replica operations can also be scheduled using ASM.

3.11 Supported versions of the Windows operating system

The below tables list the supported versions of the Windows operating system and identifies the host integration tools components that do not support specific operating system versions. Host integration tools do not support any evaluation version of Windows or any version not listed in the below tables.

Table 3 Windows Desktop Operating System Support

Component	Windows 7 SP1	Windows 8.0 and 8.1	Windows 10
RSW	Yes	Yes	Yes
Remote Setup CLI	Yes	Yes	Yes
MPIO/DSM	No	No	No
ASM/ME	No	No	No
VSS Provider	No	No	No
VDS Provider	No	No	No
PowerShell Tools	Yes	Yes	Yes
Storage Management Provider (SMP)	No	Yes	Yes
Rethinning driver	No	No	No

Table 4 Windows Server and Server Core Support

Component	Server 2012 and 2012 R2	Server Core 2012 and 2012 R2
RSW	Yes	No
Remote Setup CLI	Yes	Yes
MPIO/DSM	Yes	Yes
ASM/ME	Yes	No
VSS Provider	Yes	Yes
VDS Provider	Yes	Yes
PowerShell Tools	Yes	Yes
Storage Management Provider (SMP)	Yes	Yes
Rethinning driver	No	Yes

4 Backup and recovery operations

4.1 Restoring from a snapshot

There are times when data in the base volume needs to be restored to a particular point in time. PS Series snapshots provide several quick restore options to maximize the availability of critical data.

- A full volume can be quickly restored by setting a snapshot online (the host server must be powered off or the volume must be set offline before doing this).
- A snapshot volume can be mounted to the same or another host server to allow for manual recovery of an entire volume, a single file, or simply for verification or test and development use.

The restored volume will contain all of the data that existed in the volume at the time the snapshot was created and have the original volume name and iSCSI target name. During the restore operation, the base volume and the snapshot must be dismounted from the host. For example, on a Windows 2012 R2 host, the Disk Management utility can be used to mark the volume offline or a host may be shut down or powered off while the snapshot is restored. When a volume is restored from snapshot, Dell EqualLogic Group Manager sets the base volume and snapshot offline for the array and, by default, returns the volume back online when completed. Another snapshot is automatically created to preserve the contents of the volume prior to when it was restored.

All array members that contain data from the base volume or the snapshot must be online to restore the volume from the snapshot.

4.2 Restoring from a clone volume

A clone volume can also be used for restoring an entire volume or a single file. The clone volume can be attached in place of the original volume (for example, if it becomes corrupt, or infected with a virus) or it can be mounted on another host. Once mounted on another host, an administrator can then manually restore as many files as needed to the original volume while the original volume remains online.

4.3 Deleting volumes, snapshots, and clones

If a volume is deleted, the snapshots associated with the base volume are also deleted. Users can delete the snapshots manually if they no longer need the snapshot data or wish to reduce the number of snapshots for the volume. In addition to this, the PS Series group can delete snapshots automatically in the following scenarios:

- If creating another snapshot would exceed the snapshot reserve space, then the group will delete the oldest snapshot in order to create a new snapshot.
- If a scheduled snapshot reaches the maximum number of snapshots, as specified by the schedule keep count, the oldest snapshot created by the schedule will automatically be deleted before a new snapshot is created.

Deleting a clone volume does not affect the original base volume. Each clone volume is independent of the originating base volume.

A thin clone volume that was created from a template volume is dependent on the associated template volume. Deleting a thin clone does not delete the template volume it is associated with. Before deleting a template volume, all thin clones that depend on it must first be deleted or converted to regular volumes. If an administrator wants to preserve one of the thin clones, a full clone volume (essentially a copy) can be created from that thin clone before it is deleted.

4.4 Reducing backup time with snapshots and clones

Snapshots can greatly simplify and increase the performance of backup and recovery operations by providing the ability to offload the backup copy operation to a different server than where an application is running. In addition, they offer improvements to backup operations with regard to open file handling.

While snapshots provide a fast and efficient manner to create copies of SAN volumes, the snapshots are still stored with the SAN volumes and may share data with the base volume. This means that both primary application data and its snapshots are vulnerable to catastrophic loss scenarios such as fire, flood, and earthquake. Administrative mistakes, such as deleting the volume, open the possibility to lose primary data and associated snapshots. Therefore, snapshots should not be considered a substitute for traditional backup technologies, but a supplement to them.

Snapshots are inherently temporary. An administrator can configure the system to keep many snapshots for days or weeks, rather than months or years, as is typical of backup archives. Clone volumes contain a full copy of the base volume, and can even be moved to a separate storage pool. This means they are not dependent on the base volume, however, if they are stored in the same SAN will be vulnerable to the same catastrophic site disasters. Using snapshots and clone volumes in conjunction with PS Series Auto-Replication is an efficient way of creating offsite copies of volumes that can be used in the case of a disaster at the primary site.

Well-designed backup environments ensure copies of data are regularly created and stored away from the primary volume data. This backup data is typically stored in a secure location, and kept for months to years depending on the retention policies of the organization. Using a combination of PS Series snapshots, clone volumes, auto-replication, and external backups (to tape or other disk), backup data can be used both for small data recovery operations (such as a user accidentally deleting a file) and full recovery from a catastrophic failure in the data center.

Note: Dell strongly recommends customers run a comprehensive backup environment, and consider utilizing snapshots or clone volumes as part of this environment to improve backup and restore operations.

5 Snapshot and clone testing

Several tests were performed to evaluate the behavior of snapshots and clones. In particular, the utilization of snapshot reserve space was monitored under varying load conditions and data access patterns, such as sequential versus random access. The impact on the source volume when using snapshots for back-up was also tested.

The following sections explain the architecture, topology, and test scenarios in addition to how the tests were actually performed. See A for a list of software and firmware revisions used in the tests.

5.1 Test topology and architecture

For the test scenarios, two Dell EMC PowerEdge™ Servers running Windows 2008 R2 were used for hosts. Both servers were attached via 10 Gb Ethernet to a pair of Dell PowerConnect™ 8024F switches for the SAN and to a PowerConnect 6248 for LAN or client connectivity. Three PS6010XV and one PS6010E storage arrays were also connected to the SAN switches and their out-of-band management ports were connected to the PowerConnect 6248.

For management and monitoring, a PowerEdge R710 running Windows 2008 R2 was also connected to the LAN switch. This system was used to manage the storage systems via Dell EqualLogic Group Manager and SAN Headquarters (SAN HQ).

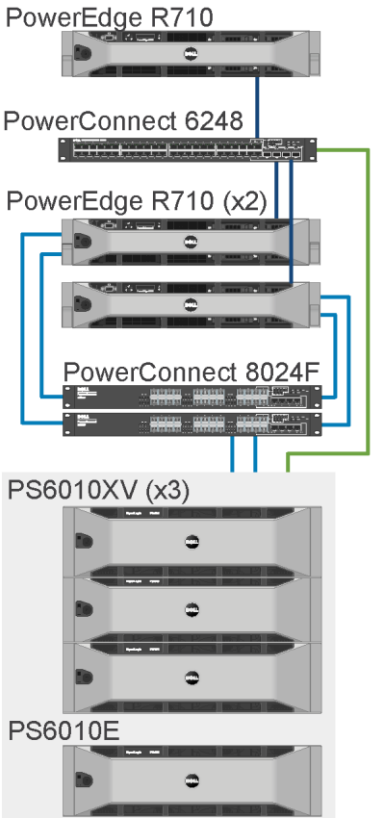

Figure 7 Test architecture

Both of the SAN attached hosts had the Dell EqualLogic Host Integration Tools (HIT) for Microsoft Edition loaded and utilized the Microsoft software iSCSI initiator and Multi-Path I/O (MPIO). The HIT Kit installed the EqualLogic Device Specific Module (DSM) for Windows which automatically created the optimal number of iSCSI connections for the volume at the default settings.

5.2 Test methodology

The three PS6010XV arrays were configured in one RAID 10 pool, and the PS6010E was configured in a separate RAID 10 pool. The PS6010E was only used in test cases involving backup scenarios as a backup target. It was attached to a host that was storing backup data. All background initialization functions were allowed to complete before running any test cases.

For test cases where a workload was involved, the open source utility IOmeter was used. The completion time reported by the backup application was used for backup scenarios. For most scenarios, each test case was executed at least twice and the average values were used.

6 Test results and analysis

This section details the various test scenarios performed and explains the results of each test.

6.1 Effect of block size and random vs. sequential I/O pattern

Data warehouse and business intelligence applications typically have a higher percentage of sequential I/O operations as well as a higher percentage of disk reads. Online transaction processing (OLTP) workloads are typically more random in nature and disk writes may be 30-50% of the total I/O operations. Although each application may have a unique I/O pattern, the effects of purely random and sequential I/O access patterns on the snapshot reserve space were tested to demonstrate their effects on snapshot reserve.

In order to simulate such an environment, three PS6010XV arrays were configured in a single RAID 10 pool. A 100 GB volume was created and attached to a Windows 2008 R2 host. After a snapshot of the test volume was created, IOmeter was used to run a random I/O workload to the base volume (as the snapshot volume remained offline) while the snapshot reserve usage was monitored at regular intervals. After collecting the results, the snapshot was deleted and the test repeated with a sequential workload.

For both workloads, the read/write mix was 67% reads and 33% writes with one worker and two outstanding I/O's accessing 100% of the volume. The intent was only to provide a steady state workload – not to test any maximum capability of the system.

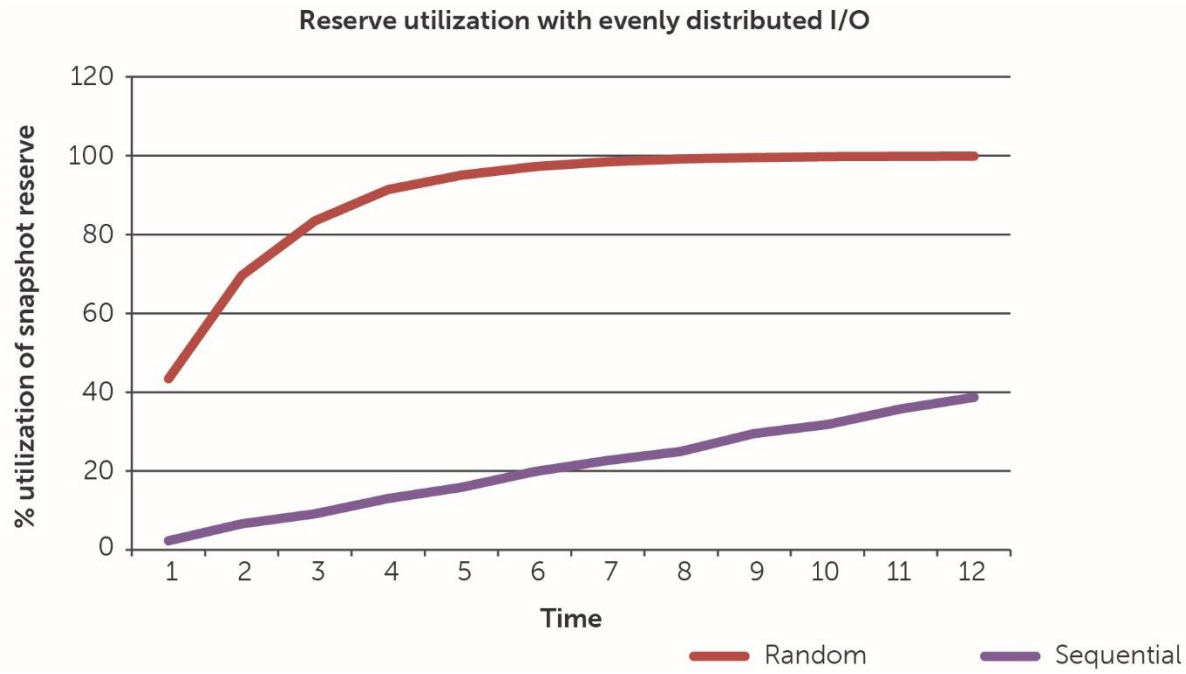

Figure 8 Reserve utilization with evenly distributed I/O

The results show that I/O access patterns will make a difference in how quickly snapshot reserve is consumed by an application. A PS Series array will allocate virtual pages of storage to a volume when it is created or expanded. The same is true for the snapshot reserve and the diagram illustrates, a purely sequential access to a volume takes longer, because it requires more I/O operations, to fill up a page before

another must be allocated to the snapshot volume (or reserve). However, when a random I/O pattern is used, additional pages are allocated for reserve more quickly as I/O's are not always adjacent and may modify a different page each time. As each subsequent I/O occurs to pages that have already been allocated, the rate of consumption slows and ultimately levels off.

It should be noted that not all applications access data in purely random or purely sequential fashion and many do not access 100% of the entire volume space. In fact, many applications access specific areas of data more often than others. The results of this test only demonstrate the behavior of simulated access patterns and do not represent actual customer results.

6.2 Reserve usage after cumulative snapshots

Many companies use snapshot technology to improve recovery time when compared to daily tape backups. Snapshots allow users to take point-in-time copies of a volume. Users can create schedules to take multiple snapshots per day of a single volume which provides the ability to recover a corrupted volume or a lost file more quickly than from tape. However, having multiple point-in-time copies can impact the amount of storage space utilized by the snapshot reserve. Moreover, as seen in section 6.1, having random or sequential I/O access patterns makes a difference in how quickly snapshot reserve is consumed by an application.

Using the same pool configuration as in previous tests, multiple snapshots of a 10 GB volume were taken over time while observing the snapshot reserve usage. However, in this test case the amount of the base volume accessed was varied. This was done because different applications may access large or small data sets and within those data sets, may access all of the data or just portions of it. For example, a user has an application running and has a schedule to take hourly snapshots of the volume. In between these hourly snapshots, the application is making changes to a percentage of the base volume. This test shows the impact of this behavior on the snapshot reserve space. By varying the number of sectors in the Maximum Disk Size value for IOmeter, only 10% of the base volume was initially accessed. Then, the access was increased in steps of 10% until 50% of the volume was accessed. Each time the access size was increased; there was another snapshot of the volume and the I/O workload was run. These tests were repeated for both random and sequential access patterns.

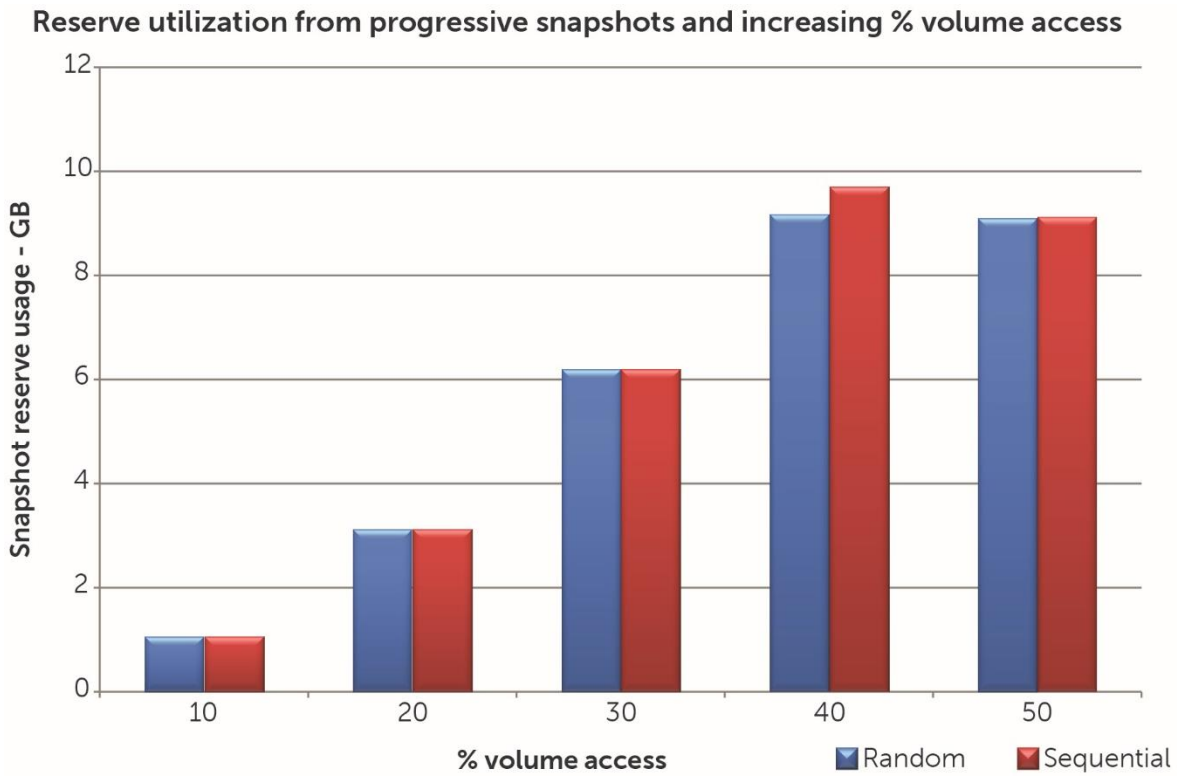

Figure 9 Snapshot reserve utilization from progressive snapshots and increasing % volume access

After the first snapshot, the snapshot reserve usage was about 1 GB (10% of the 10 GB base volume). Since only 10% of the volume was accessed, all the pages making up that 10% were affected, resulting in the same amount to be consumed by snapshot reserve space. In other words, the data in 10% of the base volume was completely changed.

After the second snapshot, about 3 GB is consumed in snapshot reserve space. In this case, 20% of the disk was accessed. If all pages were affected, 2GB of snapshot reserve space should be consumed. However, one snapshot consuming 1 GB of space was already retained. Adding these two snapshots together reflects the current usage at about 3 GB.

Adding the third snapshot results in the same cumulative effect (1GB + 2GB +3GB) returning a total of 6 GB. After the forth snapshot, there is a similar result causing the total to approach 10 GB, or 100% of the snapshot reserve.

At the fifth snapshot, the results begin to look different. Adding 10% + 20% + 30% + 40% + 50% (or 1GB + 2GB +3GB + 4GB +5GB) is clearly above the complete (100%) snapshot reserve which is 10 GB. The 100% consumption of the snapshot reserve forces the default policy to delete the oldest snapshot and make room for new snapshots. In the end, only 9 GB of the snapshot reserve is used. With this scenario, there are only two retained snapshots (from the 40% and 50% access test cases). To retain additional snapshots, the size of the snapshot reserve would have to be increased to more than the default setting of 100%.

6.3 Backup from a snapshot and clone volume

Users often use snapshots to create backups of their data. During this process there is often an application workload running on the source volumes. In this test, the performance impact on the source volume was analyzed while snapshots and clones were being backed up. During this time, the source volume was under load.

To evaluate the difference between snapshot and clone volume backup, the Windows Server Backup feature included with Windows Server 2008 R2 was used. The base (or source) volume was mounted on one Windows host and the snapshot or clone on another. The base volume resided in the pool that contained the three PS6010XV members configured for RAID 10. The second host was also attached to a 1TB volume on a PS6010E that was configured independently, in a second RAID 10 pool. A volume created from this pool acted as the backup target (where the backup files were stored).

A workload (8K random, 67% read) was applied with a variable number of worker threads to the source volume during the backup. The time it took to complete the backup was measured, and after running each backup job twice, the average was calculated.

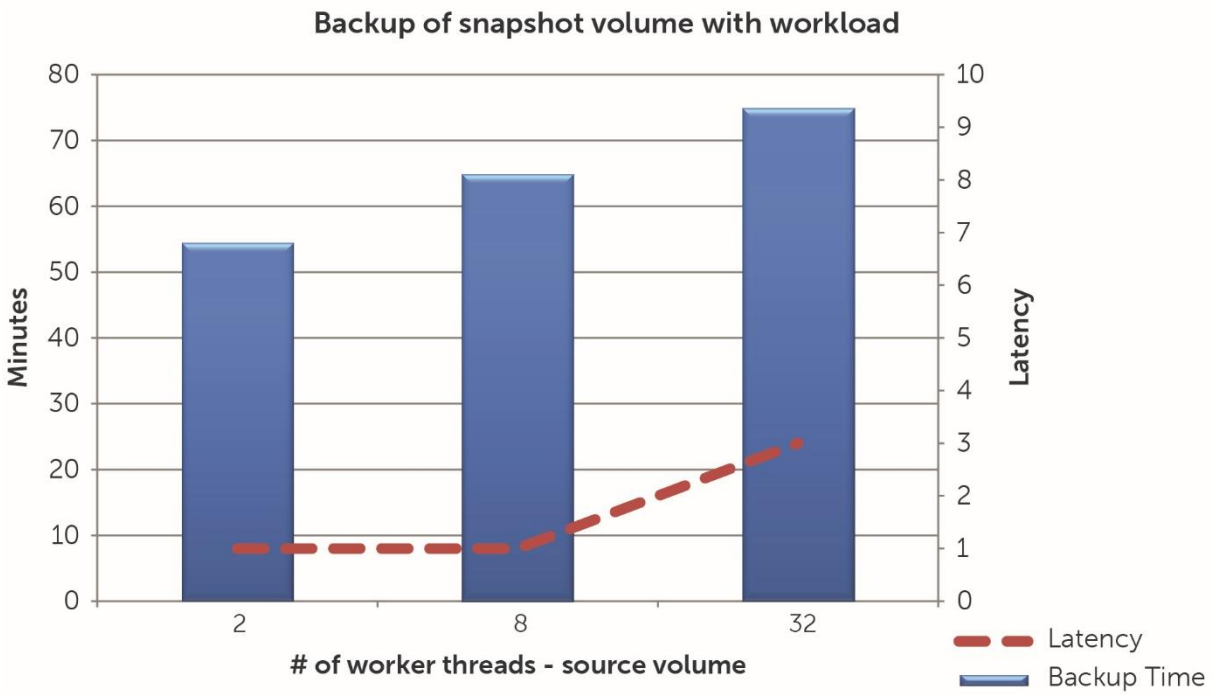

Figure 10 Backup of snapshot volume with workload

As seen in Figure 10, as the workload on the source volume increased (the number of threads was increased from two to eight, and then to 32), the time it took to complete the backup increased. The latency measured on the source volume remained very low (1ms) until the worker threads were increased to 32, at which point a slight rise (to about 3ms) in the latency was observed.

When running the same tests from a clone volume (see Figure 11), a slight improvement in backup completion time was seen. This can be attributed to the clone volume being an independent volume that does not share portions of data with the base volume (or the volume it was cloned from). Once again, latency was

very low until the worker threads were increased to 32. In both cases, this was the effect of increasing the overall workload on the storage system.

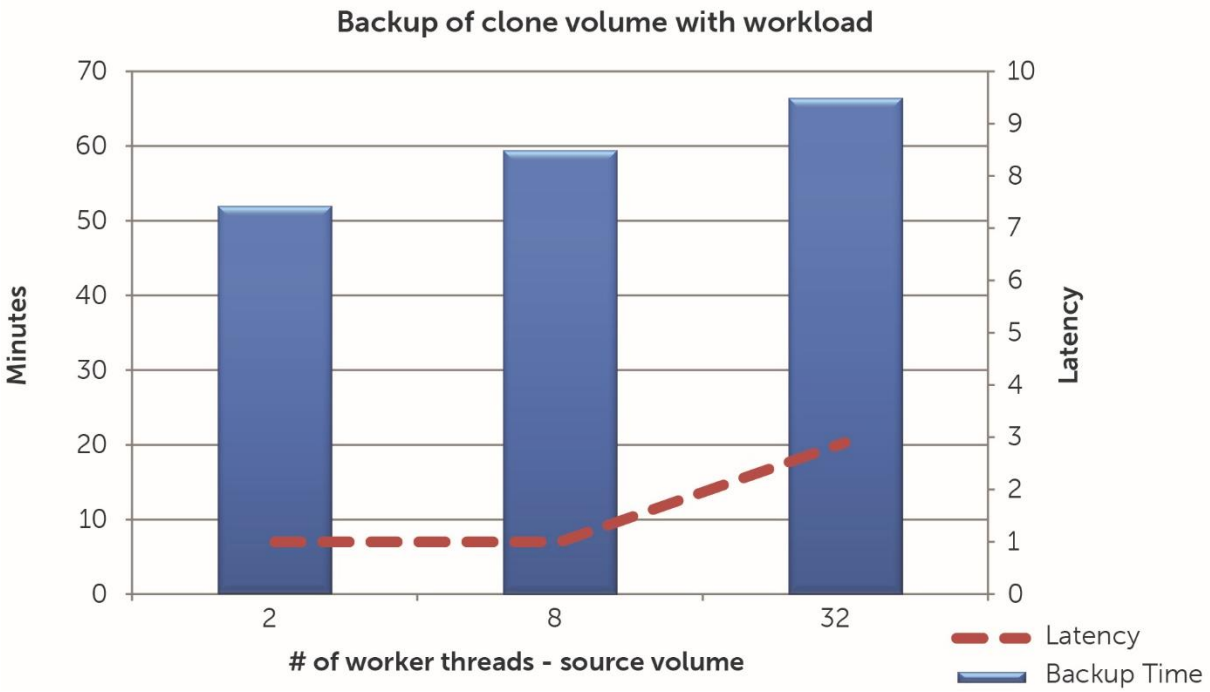

Figure 11 Backup of clone volume with workload

6.4 Backup from multiple snapshot volumes

In the previous section, it was observed that the performance impact (latency) on the source volume was very low while performing a backup from a single snapshot volume. In this test, we wanted to measure the performance on the source volume as we scaled the number of snapshot volumes that were used for backup.

While performing backups using Windows Server Backup, it was observed that the average block size on reads and writes was 128K. This is similar to many other backup applications which typically read and write data in block sizes ranging from 64K to 256K.

To determine the effect of backing up multiple volumes simultaneously, IOmeter was used to simulate the workload of a backup application. The three PS6010XV arrays in a RAID 10 pool were used to run this test.

The base volumes were mounted on one Windows host, and the snapshot volumes on the other. A workload was simultaneously run against the base volumes while running the backup workload on the snapshot volumes to simulate an application. The test started with two volumes and then incremented the volumes and associated snapshots by two until there were 12 total volumes and the latency approached 20 ms.

The IOmeter workloads used to simulate the backup application and application workload are shown in the following table.

Table 5 IOmeter workload types

Workload type	# workers	I/O type	Read/Write Mix	Block Size	Volume	Volume Access
Backup	1	Sequential	100/0	128K	Snapshot	100%
Application	8	Random	67/33	8K	Base	10%

The read/write mix indicates the ratio of read to write on the source volume when referring to the application workload, and on the snapshot volume when referring to the backup workload. Volume access is the amount of the base volume or snapshot volume the workload was allowed to access.

As shown in Figure 12, when increasing the number of volumes, and therefore the workload against both sets of volumes, the latency of the base (source) volume increased as expected. The workload was actively modifying portions of the base volume after the snapshot was created and therefore consuming snapshot reserve. However, because snapshots share data with the base volume, portions of the data were being accessed by both workloads.

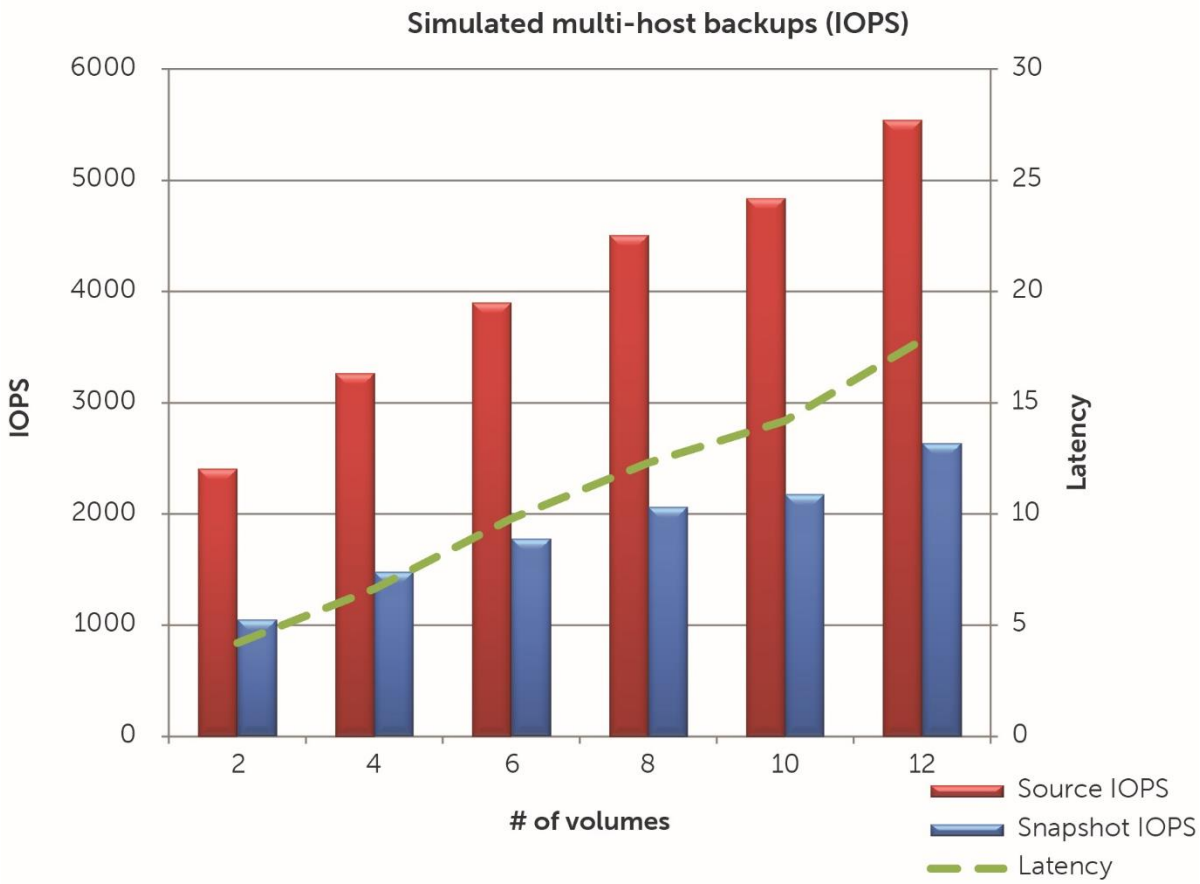

Figure 12 Simulated multi-host backup (IOPS)

Using snapshot volumes as the source for backups may affect the performance of the source volumes. In this case, the latency increased with the application simulation workload and the number of volumes being backed up (i.e., running the backup simulation workload). IOPS continued to scale, indicating the I/O processing abilities of the storage system had not yet been reached. However, had the increase to both workloads continued, eventually a point would be reached where the system would no longer deliver more IOPS and the latency would sharply increase.

7 Planning and design best practices

7.1 Use ASM for Windows and VSM for VMware

To ensure that consistent snapshot and clones are created for Windows, ASM/ME must be used so that VSS is invoked before the snapshot is performed by the underlying hardware (the storage arrays). When a clone volume is created, ASM will call the VSS writer to freeze I/O and flush any unwritten data prior to creating the clone volume, resulting in an application consistent clone.

To ensure that hypervisor aware snapshots and clones are created in VMware environments, Dell Virtual Storage Manager (VSM) for VMware® should be used. This allows integration with the VMware client utilities and Virtual Center so that snapshot and clone operations are passed to the hardware (the PS Series group) to perform where applicable.

7.2 Snapshot reserve

Snapshot reserve defaults to 100% of the base volume capacity. This is to ensure that a single snapshot can be held even if 100% of the base volume is modified, but it can be changed at any time without taking the volume offline. Snapshot reserve can be increased to values greater than 100% (up to 10,000%) or it can be completely disabled by setting the value to 0%. The global defaults may also be adjusted by going to the Defaults tab under Group Configuration (Figure 13). To conserve free pool space, snapshot reserve can always be disabled for a volume if there are no plans to use snapshots, or if another method (such as an application-level snapshot or replication) is being used instead.

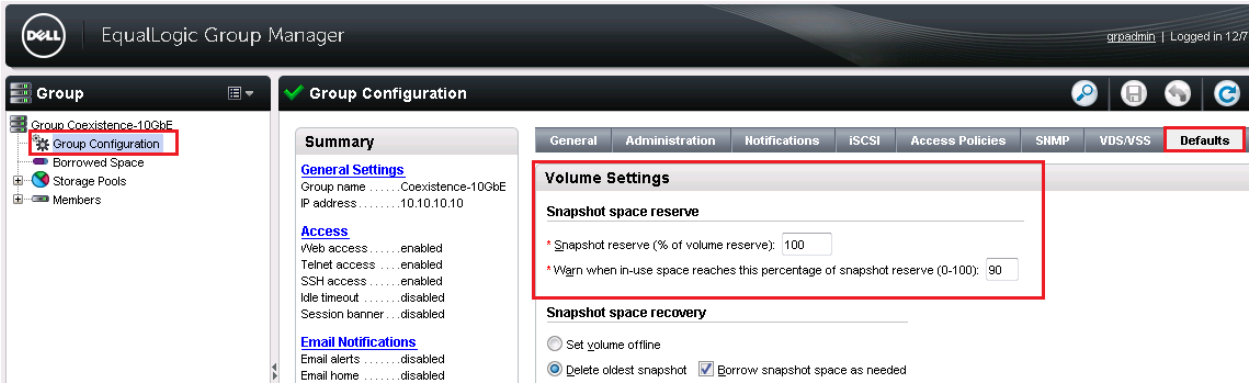

Figure 13 Default volume settings

Administrators should use the Dell EqualLogic Group Manager or SAN HQ to monitor the usage of snapshot reserve space to ensure that there is adequate space to retain the desired number of snapshots. The number of snapshots retained should be driven by the organizations RTO/RPO (Recovery Time Object/Recovery Point Object). For example, if a business determines that an acceptable loss of data is one hour, then administrators may schedule a snapshot at least once every hour during business hours.

The rate that data changes on the base volume will determine how much snapshot reserve is consumed between snapshots. If multiple snapshots are retained, then multiply the estimated amount of snapshot reserve needed by the number of snapshots retained.

Total Snapshot Reserve = (Δ snapshot 1) + (Δ snapshot 2) + (Δ snapshot 3) + … (Δ snapshot *n*)

For example, if 1 GB of data changes between snapshots of a volume and the administrator wishes to retain 5 snapshots, then at least 5 GB of snapshot reserve is needed. Because the data change rate may fluctuate slightly. Allocate slightly more than 5 GB to the snapshot reserve to ensure that all five are retained.

7.3 Clones versus snapshots

Because snapshots share data with a base volume, snapshots are not a replacement for backups. If the base volume is lost, all associated snapshots would be lost too. Snapshots can be used to facilitate more efficient backups (to tape or disk backup targets), especially when used with application-aware frameworks like Microsoft's VSS, which allow the volume and application to remain online during a backup.

PS Series snapshots and clones can enhance the recovery process by allowing an administrator to quickly rollback a volume to a specific time, or even access a single file on a snapshot or clone volume for fine grain recovery. A typical use case for snapshots is to allow system administrators to respond to and recover from storage loss and data corruption situations overnight. Clones can be used for longer term recovery or for test and development purposes since they are a full independent copy of the source volume.

7.4 Potential causes of unexpected snapshot growth

When a volume is defragmented, data is moved in the file system to ensure that large files are contiguous. This prevents the file system from having to search the disk to access a complete file. The process of defragmenting typically results in a large amount of the volume being modified; not the actual data contained in the files, but simply the way the files are stored. When a snapshot is performed before defragmenting, the PS series snapshot reserve may rapidly grow to as high as 100% depending on how severely the disk is defragmented.

Some applications also defragment or reorganize database files on a schedule. These operations can lead to a similar effect as defragmenting the file system. Large amounts of data may be modified or deleted leading to a higher amount of snapshot reserve being required to retain the number of desired snapshots.

Similarly, using a backup application that resets the archive bit on files also causes the file system to be modified resulting in a greater consumption of snapshot reserve space. Many modern backup applications use a database (or other method) to track which files have already been backed up and modified files that need to be backed up, preventing the application from needing to modify the archive bit.

Windows 2003 systems (or systems that were upgraded from Windows 2003) may have the NtfsDisableLastAccessUpdate feature enabled (its value in the Windows Registry set to 0). When this feature is enabled, the NTFS file system will update the timestamp on the file when it is opened, even if it is not modified. This results in the file system, and therefore the page of the PS Series volume, to be modified, forcing snapshot reserve to retain the old version. Setting this value to 1 disables the feature (default for Windows 2008 and later).

See the following Microsoft Knowledgebase article for more information:

http://technet.microsoft.com/en-us/library/cc758569(WS.10).aspx

7.5 Monitoring snapshot reserve

As mentioned in the section titled, “Snapshot reserve”, the default setting allocates snapshot reserve equal to 100% of the volume size (reserve). The default settings also include a warning threshold of 90%. Both of these defaults can be changed to affect new volumes at any time, either globally, or per volume. Existing volumes will not be affected.

When the amount of snapshot reserve used reaches the threshold, Dell EqualLogic Group Manager displays a warning condition as well as a message in the event log. Optionally notifications can be configured to be sent by email, logged to a syslog server, and collected by SAN HQ.

SAN HQ can also be used to monitor snapshots and reserve usage. Detailed information ranges from overall utilization of an entire group (or groups), to pools, or a specific volume. Because SAN HQ tracks usage over time, administrators can use it to monitor trends and plan for future storage needs. The following diagram shows an example of volume and snapshot reserve usage over a period of time as viewed in SAN HQ.

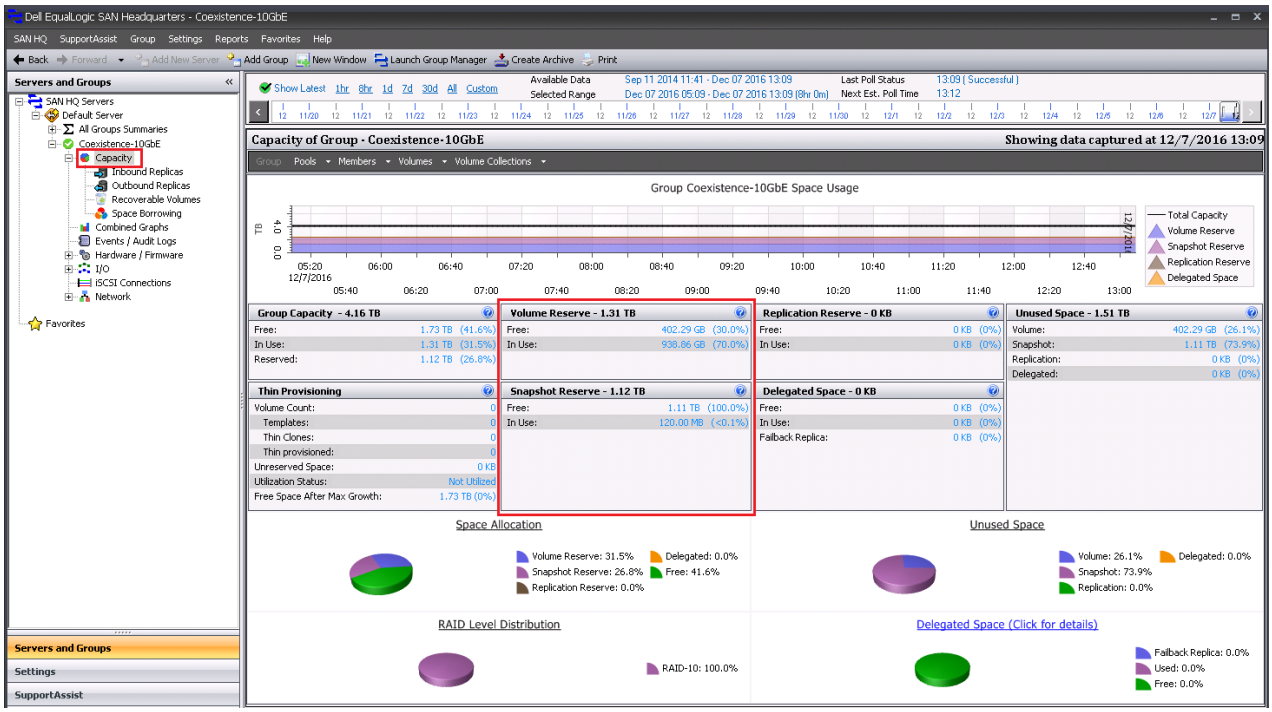

Figure 14 PS Series SAN HQ – Volume & Snapshot Reserve

8 Summary

The snapshot and clone features of PS Series storage arrays provide a set of powerful data recovery features delivered with every PS Series storage array. The information presented in this whitepaper can be used to assist administrators in achieving the most from these features.

- As the results from the lab testing show, PS Series storage systems use efficient methods of creating snapshots and clones that allow for instantaneous access as well as low latency and low disk utilization during their creation.
- Snapshots and clone volumes can be used to enhance backup solutions by allowing applications and volumes to remain online while a snapshot is created.
- Snapshots and clone volumes can be used to facilitate fast recovery of volumes and individual files.
- Snapshot reserve can be adjusted at any time to accommodate for changes in data access patterns, or to make more efficient use of available free space in the pool.
- Application access patterns, whether they access data randomly or sequentially, may affect the amount of snapshot reserve required.
- Auto-Snapshot Manager is included at no additional charge with every PS Series array and can be used to automate and simplify the creation or scheduling of snapshots and clone volumes. ASM creates application aware (or hypervisor aware) consistent snapshots and clones and aids in the automated recovery of the data.
- SAN HQ, also included at no additional charge with PS Series arrays, allows an administrator to monitor snapshot reserve usage and alerts from one or more PS Series groups.

A Solution infrastructure hardware and software versions

PowerEdge R710	
Windows 2008 R2	SP1
HIT kit	3.5.1
BIOS	6.0.7
LOM firmware	6.4.4
Broadcom BCM5709C Driver	6.2.9.0
iDRAC firmware	1.70
Lifecycle Controller firmware	1.50.671
Intel X520 Driver	2.5.52.2
Intel X520 firmware	12.5.2
PS6010XV	
Controller firmware	5.1.2
450GB 15K SAS drive firmware	ERHB
PS6010E	
Controller firmware	5.1.2
1TB 7.2K SATA drive firmware	KD03
PowerConnect 8024F	
Switch firmware	4.1.0.19
Additional Applications	
IOmeter	2006.07.27
SAN Headquarters	2.2.0.5924

B Additional resources

Dell.com/support is focused on meeting customer needs with proven services and support.

Storage technical documents and videos provide expertise that helps to ensure customer success on Dell EMC storage platforms.