How Low-Capacity Snapshots Can Save 6TB Of Storage In One Week

Minimizing capacity utilization when creating snapshots in Fibre Channel SAN environments

Today, many storage administrators utilize data snapshot technology to help them accomplish many tasks that not only enhance their levels of data protection, but also enable traditionally "disruptive" functions to be performed in a non-disruptive manner. Some of the benefits of these point-in-time snapshots include:

- Performing zero-impact, zero-window backups that occur "in the background" while production systems remain online, and without using CPU cycles on production servers
- Enabling data to be rapidly restored from disk rather than tape, minimizing the time needed for data or system recovery
- Accelerating application testing by using copies of production data for testing purposes
- Updating data warehouses using copies of production data without disturbing ongoing operations

However, many of the snapshot solutions offered in the market today consume excessive amounts of disk capacity. That's because many snapshot products create full-size copies of production volumes, and often require disk capacity equal to the original volume size to be "reserved" for use only by the snapshot application. The result is that many organizations can only afford to retain a small number of snapshots on their storage devices at any given point in time, severely limiting the potential benefits that are possible.

Fortunately, a new generation of snapshot solutions that generate space-efficient, low-capacity point-in-time copies of data have emerged that only consume a fraction of the disk capacity previously needed. In particular, this paper will focus on the use of StoreAge's multi*view* snapshot solution, and how it can benefit any organization with Storage Area Network (SAN) implementations.

Snapshots – A Definition

A snapshot is a point-in-time logical image of a volume. In other words, it represents what a particular volume looked like at the time the snapshot was created.

However, when looking at what really happens "behind the scenes" when a snapshot is created reveals significant differences in how different vendors implement this data copy technology.

A full-size snapshot does exactly what the name implies: It creates a point-in-time copy of a volume that is equal in size to the source volume.

1. Full-Size Snapshots (Clones)

A full-size snapshot does exactly what the name implies: It creates a point-in-time copy of a volume that is equal in size to the source volume. As part of this process, many full-size snapshot products require the storage administrator to set aside, in advance, a sufficient amount of space where the copy of the volume will be created. Unfortunately, many products also "reserve" this space for use only by the snapshots, preventing administrators from using this reserved space for any other data storage function, even if no snapshots currently exist.

Many full-size snapshot products "reserve" disk space for use only by the snapshots, preventing administrators from using this reserved space for any other data storage function, even if no snapshots currently exist.

In addition, some full-size snapshot products use a "triple mirror" design. In a triple-mirror configuration, a mirror of the source volume is continuously maintained (requiring disk capacity equal to the source volume size), and then full-size point-in-time snapshots

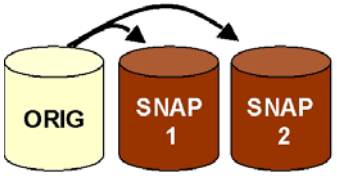
are created ("split off") from the mirrored volume upon command. The downside of is technique is that it requires disk capacity equal to three times the source volume size.

> *Example:* To create a snapshot of 1TB of data, a total of 3TB of capacity will be needed (1TB for the source volume, 1TB for the mirrored volume, and 1TB for the snapshot). And, in most cases, none of this capacity can be used by any other application.

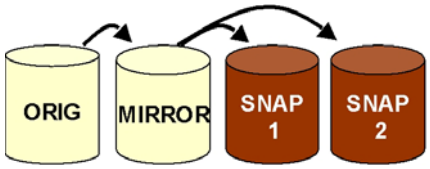
Examples of full-size snapshot solutions include:

- EMC TimeFinder (triple-mirror design)
- Hitachi ShadowImage
- IBM FlashCopy (for Shark)
- Veritas FlashSnap

The downside of the "triple mirror" design is that it requires disk capacity equal to three times the source volume size just to create the first snapshot.



Full-size Snapshots (clones) – *Each snapshot is the same size as the source volume.*



Triple-Mirror Design – *Maintains a mirror of the source volume, and each snapshot is the same size as the source/mirror volume.*

2. Low-Capacity Snapshots

Low-capacity snapshots are space-efficient alternatives to full-size snapshots, because they only keep track of the changes that are made to a volume after the snapshot is created. In other words, if 5% of the data in a volume has changed since the snapshot was created, the snapshot will be approximately 5% of the source volume size.

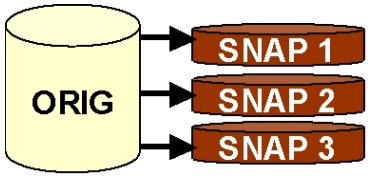
Low-capacity snapshots are space-efficient alternatives to full-size snapshots, because they only keep track of the changes that are made to a volume after the snapshot is created.

In addition, once they are created, low-capacity snapshots consume disk capacity in a granular "as needed" fashion, further enhancing their space-efficient characteristics. Typically, these products will automatically expand the size of the snapshot volume in small increments as changes to the source volume occur.

Examples of low-capacity snapshot solutions include:

- StoreAge multi*view*
- EMC Clariion SnapView
- IBM FlashCopy (for FAStT)

Low-capacity snapshots consume disk capacity in a granular "as needed" fashion, automatically expanding the size of the snapshot volume in small increments as changes to the source volume occur



Low-Capacity Snapshots – *Each snapshot only contains changes to data in the source volume, and is a fraction of the size of the source volume.*

3. "Promoting" a Snapshot Into a Volume
From a system perspective, snapshots are created "in the background" and are unseen and inaccessible by server file systems until the storage administrator "promotes" the snapshot to become a volume. At that point, the snapshot appears as a normal volume that can be mounted and accessed by any appropriate server.

Compatibility & Management Issues

Many snapshot solutions have compatibility and manageability issues that should be fully understood prior to deployment.

With a few notable exceptions, snapshot solutions often have glaring compatibility and manageability issues that should be understood prior to deployment. For example, many solutions are compatible only with specific storage devices or operating systems, and may also lack the ability to be centrally managed. Limited compatibility may force you to buy and learn multiple snapshot products, impose an inefficient management scheme, or effectively "lock you in" to certain vendors in the future.

Currently, snapshot products on the market today use three different deployment architectures when used in SAN environments. These are:

- *Server-based snapshot products* – Licensed separately for each server, and can create snapshots only of data belonging to that server.
- *Storage-based snapshot products* – Licensed separately for each storage device, and can create snapshots only of data residing on that storage device
- *Network-based snapshot products* – Licensed per SAN, and can create snapshots of any data on any storage device in the SAN.

To further examine the pros and cons of each approach, the following chart compares the compatibility and management features of many of the most popular snapshot solutions:

Several factors will influence which snapshot product should be used, including disk capacity consumption, licensing fees, manageability and compatibility.

	Snapshot Size	Compatible With	Management
EMC SnapView	Low-capacity	EMC Clariion	Per storage device
EMC Timefinder	Full-size	EMC Symmetrix	Per storage device
Hitachi ShadowImage	Full-size	Hitachi 9xxx	Per storage device
Veritas FlashSnap	Full-size	All storage devices	Per server
IBM FlashCopy (Shark)*	Full-size	IBM Shark	Per storage device
IBM FlashCopy (FAStT)	Low-capacity	IBM FAStT	Per storage device
StoreAge multi*view*	***Low-capacity***	***All storage devices***	***Per SAN***

** FlashCopy snapshots equal source volume size even when using the NOCOPY option.*

Performance Issues

Full-size snapshots will suffer some level of performance degradation, as the storage device must create a duplicate copy of the source volume.

From a performance standpoint, full-size snapshots will incur some level of performance degradation, as the storage device must create a duplicate copy of the source volume.

For example, suppose you have a storage array with 1TB of data. To create a full-size snapshot of 1TB of data, the storage device must physically copy all of that data to the snapshot while servicing other I/O requests from other applications.

Low-capacity snapshots, on the other hand, do not impact performance at the time they are created, but do utilize some level of storage device processing power to update the

snapshot as data in the source volume changes. The amount of processing power needed is determined primarily by two factors: How frequently source volume data changes, and how many snapshots are retained at any given point in time. On average, industry analysts estimate that 5% or less of data changes on a daily basis, so in most cases, only a modest amount of data movement needs to occur to keep each snapshot updated.

Since the overall cost of managing storage will greatly exceed the initial purchase costs, snapshot solutions that can be centrally managed and uniformly applied across all storage in a SAN, regardless of the number or type of servers and storage devices in use, will result in a lower total ownership cost compared to other management approaches.

Cost Issues

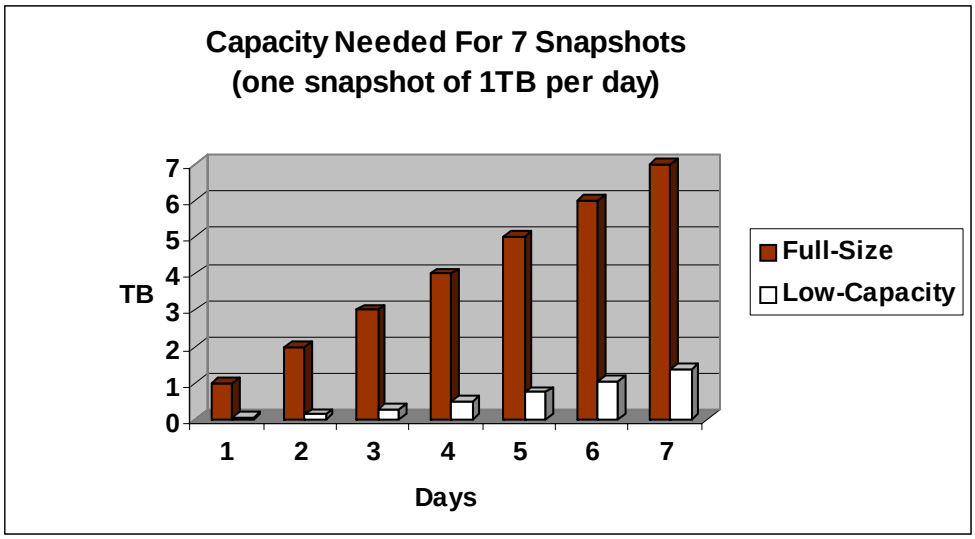
From a cost standpoint, there are several factors that should be considered before selecting a snapshot solution.

1. Disk Capacity

As discussed above, full-size snapshot products will consume significantly greater amounts of disk capacity compared to low-capacity snapshot products. This is especially true if multiple snapshots will be retained online for several days, which is the preferred option for many storage administrators.

For example, the following chart illustrates the disk capacity required to create and retain one snapshot of 1TB of data per day, for seven days.

Low-capacity snapshot products can make it economically feasible to create and retain multiple snapshots each day.



Assumptions: Data in source volumes changes 5% per day.

In this example, the low-capacity snapshots save 6TB of disk capacity that can be used for other purposes. At a cost of $0.08 per MB, this represents a savings of $480,000 in disk capacity alone.

2. Licensing

As discussed above, different snapshot products use different licensing schemes. In general, license fees for network-based snapshot solutions (usually licensed per SAN) will tend to be lower than products licensed either per server or per storage device.

3. Management

As is frequently cited in the storage world, the overall cost of managing storage will greatly exceed the initial purchase costs of storage. Therefore, snapshot solutions that can be centrally managed and uniformly applied across all storage in a SAN, regardless of the number or type of servers and storage devices in use, will result in a lower total ownership cost compared to other management approaches.

4. Compatibility
Since many snapshot products have limited compatibility with different servers and storage devices, a heterogeneous SAN environment may require multiple snapshot products to cover all the servers and storage devices in use. However, this is not the case with network-based snapshot solutions, which can create snapshots of any data on any storage device in the SAN, and are often the most cost-effective solution.

Creating multiple snapshots of critical data each day is a highly effective method of minimizing risk and reducing data loss. Frequent snapshots not only allow data to be rapidly restored from disk, but also minimize the amount of data that is "at risk" since the last backup.

5. Data Risk & Data Loss
Creating multiple snapshots of critical data each day is a highly effective method of minimizing risk and reducing data loss. Frequent snapshots not only allow data to be rapidly restored from disk, but also minimize the amount of data that is "at risk" since the last backup. For example, restoring data from a snapshot that is one hour old is far preferable to restoring data from tape that is 24 hours old. Low-capacity snapshot products can make it economically feasible to create and retain multiple snapshots each day.

Summary

StoreAge multi*view* low-capacity, point-in-time snapshots combine the space-efficiency, compatibility and centralized management features that significantly lower the total cost of implementing snapshot technology in SAN environments.

Compared to other snapshot solutions, the key benefits of implementing multi*view* include:

- Consumes significantly less disk capacity, enabling more frequent creation of snapshots for backup, restore, application testing and other purposes
- Compatible with virtually all operating systems and storage devices
- Reduces acquisition, training and maintenance costs by providing a centralized and uniform method of creating snapshots across all servers and storage devices
- Eliminates "vendor lock-in," allowing future storage purchases to be based on price/performance criteria rather than on vendor compatibility
- Reduces the amount of data "at risk" at any given point in time

U.S. Headquarters
One Technology Drive, Bldg. C-515
Irvine, CA 92618
T: (949) 754-0640
F: (949) 754-0645

International Headquarters
63 Bar Yehuda Road
Nesher 36651 Israel
T: +972 4 8203454
F: +972 4 8203464

For more information:
www.StoreAge.com
info@StoreAge.com

07/30/04