A WestWorld Productions, Inc. Publication
OCTOBER 2004 VOLUME XXIII Number 10
www.wwpi.com

Computer Technology Review®

THE STANDARD OF EXCELLENCE FOR STORAGE TECHNOLOGY INFORMATION

Better Backup and Recovery with Snapshots

By Peter Hunter

Go back fifteen years and poll IT managers on their number one problem and they'd say "backup and recovery." Ask them the same question today, and the answer is the same. This reigning champ of IT problems has a new challenger that is promising to upset traditional thinking and finally put the issue of backup and recovery to rest. The challenger is snapshots. On their own, snapshots offer an efficient quick-restore capability, but when combined with networked storage and the best practices in today's backup systems, the result is a revolution of elegance.

What is a Snapshot?

Snapshot is the user-friendly word for point-in-time copy (the terms are interchangeable). A point-in-time copy represents an exact duplicate of a data volume at the moment the copy was created. Subsequent changes to the data volume do not affect the point-in-time copy. At any future date, one can restore the data from the point-in-time copy. The clever part is that these duplicates don't actually take up any additional hard-drive space. Only the changes that occur after you take the snapshot are stored. Data that remains unchanged is only stored in one place and is referenced by both the baseline volume and the point-in-time copy. The granularity of the stored changes is at the block level (the smallest chunk of data that a disk system can work with), so this results in a very space-efficient method for creating restorable copies of data.

The principal application of snapshots is to provide a quick backup-and-restore option. Any given volume can have multiple snapshots representing different time slices in the history of the data volume, and the taking of snapshots is a nearly instantaneous procedure. If a problem occurs that requires a recovery operation, the system can be restored to a previous snapshot. The restoration procedure is also very quick, as the volume is brought offline and the snapshot brought online—a process often referred to as rollback.

Unlike traditional tape or disk-to-disk backup restorations, snapshot rollback does not require that the data be "unpacked" from a special backup format by being run through a backup server. The snapshot image is the original disk image—what the operating system expects to see. This greatly improves recovery times.

Snapshots are co-located with baseline volumes, meaning they reside in the same storage system and are protected with the same data-protection features such as RAID and redundant controllers. However, the snapshots also are vulnerable if the primary storage systems should somehow fail. As they stand alone, snapshots cannot replace traditional tape backup and recovery; however, they can vastly improve them.

Backup Windows

The "backup window" problem results from the traditional approach to backup. Application servers traditionally stream data to a tape device. For many applications, this backup operation can take many hours. During the backup, the application might not be able to provide normal service. In cases where the application is still providing service, its performance will be degraded, as the backup involves moving a large amount of data to the tape device. This period during the backup is referred to as the backup window.

Most businesses have constraints on how long the backup time can last. This varies from a few hours in the early morning to only a few minutes. In addition, this small window of opportunity reduces the frequency of backups, which in turn reduces the recovery options.

How Snapshots and Networked Storage Work Together

Networked storage decouples the storage devices from the servers and interconnects them over a storage area network (SAN). SANs allow the application servers and backup servers to see the same disk storage systems. When coupled with snapshots, this visibility into the storage enables a new form of backup and recovery. Here's how it works:

The application server takes snapshots of its volumes. The backup server then mounts the snapshots and copies them over the SAN to tape devices. The tapes can then be moved offsite to a secure location.

In the end, the application server operates uninterrupted during the backup process. The application server is completely unhindered during the backup operation. In addition, the backup server and tape libraries can operate all of the time, eliminating the backup window problem.

Understanding Snapshots and Application Interactions

Applications play with their data like children play with their toys. The application starts and opens up its data files and then throws the data all over the place in memory and on disk. When the application is finished (unlike children) it puts the data away in a nice clean format, so that it can quickly resume work once restarted.

With snapshots, the data is captured in a disordered state. Application recoveries from this state are not always successful. This concept is called "snapshot consistency", and a snapshot that is taken in this way is considered an inconsistent snapshot.The solution is to make the application aware of the snapshot process. In this case, the application briefly flushes dirty buffers of data to disk before the snapshot is taken. Some coordination occurs between the application and the storage device, such that the snapshot is taken at the right moment. The application remains online and continues its normal services while this process is underway. All of this happens in a matter of a few seconds and then the snapshot is taken.

The result of this coordination with the application is an application-consistent snapshot. This snapshot image is as good as gold and can be copied to tape for safe keeping.

A New Paradigm

Today's operating systems, applications and storage arrays are integrating more and more snapshot coordination capabilities, simplifying the deployment and management of snapshots with backup and recovery. The promise of snapshot technology and better backup and recovery is here, as it finds its way into every storage management toolkit and becomes established as a standard data center function. CTR

Peter Hunter *is product marketing manager at EqualLogic, Inc. (Nashua, NH)*

www.equallogic.com

Reprinted by permission from the publisher of the October 2004 edition of ***Computer Technology Review***®.