

Getting in front of backup

http://searchStorage.techtarget.com/magazineFeature/0,296894,sid5_gci1518767,00.html

by: [Lauren Whitehouse](#)

Issue: [Sep 2010](#)

Learn about a handful of key technologies that can help storage managers meet their backup recovery time objectives by making the first steps -- data capture and transfer -- simpler and more efficient.

The focus on backup modernization during the last few years has been squarely on the backup target device: tapes and disks. That's where the majority of users have made the most changes. But now that so many users and IT shops have become disk friendly, there's a new focus on the front end of the backup process: the capture and transfer phase.

In 2004, nearly 60% of Enterprise Storage Group (ESG) survey respondents reported backing up directly to tape. By 2010, only 20% were using tape exclusively. These days, approximately 80% of IT organizations tell ESG they're augmenting backup processes with disk, which helps them meet backup windows and recovery time objectives (RTOs). Still, exponential data growth means greater backup demands and a need for new backup processes. As a result, technologies such as continuous data protection (CDP), replication, source-side deduplication and snapshot are being implemented more frequently. ESG research found a significant uptake in several of these technologies: while the use of snapshots grew only 2% between 2008 and 2010, replication use increased 34%, CDP expanded by 58% and deduplication use improved 66% in the same two-year period.

Snapshot and image-level backup

What if you could eliminate your backup window, accelerate system recovery and facilitate efficient disaster recovery (DR)? Effectively, that's what snapshot- and image-based backup can deliver. A snapshot is a copy of a volume or file system created at a specific point in time. Taking advantage of snapshot functionality for backup can dramatically reduce the impact on apps by eliminating the backup window, providing RTOs of seconds to minutes and enabling better recovery point objectives (RPOs) by enabling more frequent copies per day.

Image backup uses snapshot technology to create a point-in-time image of a system (such as hardware configuration, OS, applications and data), storing it in a single portable file. Because the recovery point is captured "hot," critical applications don't have to be shut down during backup. This approach eliminates the backup window and enables rapid whole-system recovery to any system (virtual or physical), including to dissimilar hardware. Both of these methods are efficient in the capture, transfer and storage of data. After the initial base copy is made, only incremental blocks are captured and stored.

CDP

CDP technology continuously captures changes to data at a file, block or application level, supporting very granular data capture and recovery options. It time stamps each write and mirrors it to a continuous data protection retention log. When a recovery is needed, the CDP engine creates an image of the volume for the point in time requested without disrupting the production application.

Block-level CDP operates at the logical volume level and records every write. This type of continuous data protection stands out at transparent data capture and presentation of views at different points in time. Typically

running on the same server as the application it's protecting, file-level CDP operates at the file-system level and records any changes to the file system. Application-aware CDP tracks critical application process points within the CDP data stream that can greatly simplify recovery, such as transaction-consistent database checkpoints or application-consistent points within email applications.

Continuous data protection completely eliminates discrete backups, replacing them with a transparent, continuous data capture process that puts very low overhead on production servers. Because it captures data as it's created, that data is immediately recoverable. This allows CDP-based solutions to deliver near-zero RPOs.

Replication

Replication is the bedrock of these strategies and it's increasingly being used for data protection as a standalone process to provide operational and disaster recovery for applications with tight RPOs or RTOs; as a method of consolidating distributed data for centralized file-level backup; or in conjunction with snapshot or CDP to maintain an off-site copy and facilitate disaster recovery. Replication provides an exact mirror copy of data on a local or remote primary system that can be mounted to rapidly recover from a failure. Storage capacity and bandwidth are optimized with block-level updates and network compression after the initial copy is made.

Replication is available on host systems, storage arrays or in network-based products. Typically, array- and network-based products replicate at the block level and host-based offerings replicate at the file-system level. Host-based replication operates asynchronously, while array- and network-based replication are configurable for synchronous or asynchronous modes. Synchronous replication occurs in real-time as data is written to primary storage; then it's replicated on secondary storage. Asynchronous replication occurs in near real-time. Once data has been completely written to primary storage, the written data is replicated on secondary storage.

Source-side deduplication

Deduplication identifies and eliminates redundancy, storing only unique data and shortcuts to unique data for duplicates. Data deduplication's role in optimizing backup processes is fairly well documented; however, the focus has mostly been on target-side deduplication solutions. Source-side deduplication ensures that only changed segments are backed up after the initial full copy. That means significantly less data is captured, transferred and stored on disk. This reduces the time needed to perform backups. Because the backup window requirements are minimal, it's possible to back up more frequently, which increases the number of recovery points on disk storage to meet RPO and RTO requirements.

A wholesale replacement of file-level backup is likely for many organizations today, according to ESG research. For example, 55% of IT organizations surveyed by ESG plan to replace existing file-level backup with snapshot and/or CDP solutions. That said, the integration of snapshot, replication, CDP and deduplication into existing backup platforms to augment file-level approaches seems to be a strong trend. That's why several backup vendors have made recent strides to match capture techniques to recovery objective policies, simplifying implementations and optimizing the front end of backup processes.

BIO: Lauren Whitehouse is an analyst focusing on backup and recovery software and replication solutions at Enterprise Strategy Group, Milford, Mass.