



Synchronizing Data

*Joel Lovell
Senior Storage Consultant
Storage Engine, Inc.*

One of the major technical challenges in maintaining a backup data center is keeping the data at the backup data center synchronized with the data in the primary data center.

While many database applications feature some built-in self-management and data protection capabilities, total enterprise data protection—comprehensive and efficient backup, restore, and disaster recovery—in such a critically important and dense environment can be daunting and ineffective. Even though they design and implement highly available server environments, server availability is typically going to be restricted not by the application server's performance or capabilities, but by the backup, restore, and recovery windows, which are driven by process- and task-time. As a result, the key to the most effective and highly performing enterprise database system is the optimization of data protection capabilities, and this means effectively synchronizing data from primary to backup.

In any server environment, the heightening availability requirements and burgeoning user traffic— impacting the rate of data change and corresponding resource requirements—as well as the associated volume of data growth command constant attention from any database or IT administrator. Frequent data protection challenges and concerns related to keeping production data in synch with backup or recovery data include:

- Shrinking backup windows but larger backup requirements (making more traditional methods of backup less realistic)
- Hard-to-achieve recovery point and time objectives (putting enterprise data and business continuity at risk in the event of unplanned downtime due to a disaster)
- A constant need to reconfigure disk resources and add storage capacity (a result of growth in data volumes and application servers, producing an inability to dynamically manage storage and storage capacity growth across multiple servers).

Slow and Inadequate Backup, Restore, and Recovery

In an enterprise server environment, administrators often spend too much time trying to meet backup needs, correct problem databases, and perform archive retrievals, emergency restores, or system recoveries. In the end, despite built in application capabilities, users do not experience the necessary application uptime because administrators cannot achieve the system performance required. Once again, backup, restore, and recovery optimization in a database server environment is the key.

Difficult-to-Manage Infrastructure and Processes

In the typical enterprise server environment, the variety and placement of the infrastructure alone can make storage management difficult. Enterprise-wide systems can depend on a combination of storage types (DAS, NAS, and SAN), adding complexity and duplicity to data

protection tasks, to say nothing of overall system management. Due to the complexity of the typical enterprise environment, database administrators on average spend a significant amount of time carrying out ongoing management tasks that include provisioning, configuring, and mapping for performance tuning, space management, and system resource tuning.

The Solution To Data Synchronization

By combining a standard open architecture storage virtualization solution you can ensure that enterprise server environments are reliably protected, highly available, and easily managed. *By moving storage functionality and management—specific to the application—to the layer between the host and the storage device, into the storage network - data protection is improved without changing the existing storage paradigm (unless desired).* The storage devices appear to their respective servers as the same devices that they were using before the solution was implemented. Once the solution has been placed in the “data path” between the host (application server) and the storage device, allows the introduction of an advanced data protection scheme for backup, restore, and disaster recovery which can be achieved immediately.

Administrators gain an instant, granular, point-in-time rollback capability (via disk-based journaling) to protect data from soft errors such as accidental deletions, corruptions, and viruses. At any given point in time, an administrator can roll back to a database image from minutes, hours, days, weeks, or months prior. Even individual database files or table spaces, deleted outside of an applications supplemental data protection features, can be retrieved quickly and fully intact.

If your solution permits the creation of continual, consistent point-in-time images of any managed volume (including of both data and logs) by using copy-on-first-write technology, the snapshots behave much like physical copies of data, but, unlike a physical mirror, they do not require a significant amount of time or storage capacity to create. Snapshots require mere moments and only a fraction of the disk storage space demanded by the original application. Consequently, a greater volume of data can be protected using snapshot technology, and backups and restores can be accomplished from the snap-shot instead of production data. Since the snap-shots can be accomplished in even very small intervals (minutes), (and in some cases even offers continuous journaling of a production system's write-command in the background for a more efficient asynchronous mirroring over WAN), achieving established Recovery Point Objectives (RPO) becomes very straightforward.

Virtualization with snap-shot & delta-based replication are available from a number of sources. For the most part, the leading solutions offer the following advantages:

- Replicates only the final version of changed blocks
- Dirty-block addresses are accumulated (not the data, just the addresses)
- When replication starts, a snapshot is taken
- Dirty-blocks are read from the stable snapshot and sent to DR site
- Only latest version of block is sent (best efficiency)
- Microscan eliminates “same-content-writes”, insures efficient synchronization of data in both directions
- OS or App writes in units of “clusters”
- Can result in writing of same contents back to the block
- Block-level hashing can identify same-content-writes and send only the blocks that are truly changed
- 30% to 90% reduction of WAN traffic

- Some solutions offer protocols optimized for storage traffic over high-latency WAN connections.
- Designed specifically to avoid excessive round-trip acknowledgement
- Maximize utilization of available WAN bandwidth
- Works with any IP-based WAN, with or without VPN/Compression
- Optional Compression/Encryption available
- Key-based authentication ensures maximum security
- Supports multiple WAN links to aggregate bandwidth

Get right back to business after a disaster

Most discussions about disaster recovery end at the notion of having systems available during a disaster. However, those who have ever been through a disaster know that while having systems up and running during a disaster is already a major accomplishment, the real challenge is actually the long and tedious process of restoring all the systems at the primary site—or new primary site— back to normal afterwards. Solutions that incorporated a Re-synchronize function as part of their replication services help IT administrators transfer data back to the primary site for normal operations. Replication options provide enterprises with a high-performance, easy-to-use, cost effective, customizable, flexible, secure solution to minimize system downtime during a disaster, as well as during the recovery process.

visit us on the web for more information

www.storageengine.com

Storage Engine, Inc.
One Sheila Drive Tinton Falls, NJ 07724
Phone: 732.747.6995
Fax: 732.747.6542

