

# Examining data replication and data mirroring technologies as part of a disaster recovery plan

By Ravi Chalaka

Implementing a remote data replication policy is the first step toward a comprehensive disaster recovery plan – which is no longer an option, but a necessity for many IT environments. In addition to the basic replication process, data mirroring can also be employed as part of the DR solution. The process of mirroring involves the use of a "shadow" disk that is updated in parallel with the primary disk, providing a real-time or near real-time copy of the primary disk. Local mirroring provides the first level of data protection with a mirror disk attached to the host machine or an appliance located at the primary site. In the event of data loss on the primary disk, the data is retrieved seamlessly from the mirror disk.

There are two types of mirroring techniques that may be used over a remote IP to establish a copy of the primary site data - synchronous or asynchronous. Synchronous mirroring solutions' performance depends on the link bandwidth (speed) and distance spanned by the remote link. Each write transaction to the disk is sent to the remote mirror disk and the application cannot continue until that transmission is acknowledged from the remote location. Synchronous mirroring solutions provide maximum data protection at the expense of degraded primary site performance and reduced network throughput on the link to the disaster recovery centre.

An optimum alternative that has minimal impact on performance, while minimising data loss is asynchronous mirroring. With this technique, multiple writes are transmitted without waiting for individual acknowledgments. The use of asynchronous mirroring offers the additional benefit of near real-time availability of data; online standby of data is only a few writes behind the primary site.

## **Data replication made easy**

Data replication is the basis of all disaster recovery solutions and involves periodically copying of a volume's data onto a secondary storage device, which can be located any distance from the original, preferably far away. If the main storage device should fail, data on the secondary storage device can immediately be promoted to primary status and brought online. Replication is a continuous process that begins by establishing a complete copy of data at risk at the disaster recovery (secondary) site. With that copy as a baseline, the replication process continues, recording any changes to data and forwarding those changes to update the secondary site based on watermarks with a user-specified policy.

## **Fast recovery**

For true protection against major disasters, disaster recovery centres require remote sites to be located hundreds of miles away from the business, raising issues of data loss and synchronisation of data between the production and DR site. The ideal disaster recovery solution provides quick time-to-recovery (TTR), assuring continuance of operations through near-uninterrupted access to data. Additionally, the solution must minimise data loss for a graceful recovery, keeping primary and DR centre data sets synchronised while minimising human intervention to reduce errors during the recovery process.

For robust data protection, the remote replication solution must also work in conjunction with options for multiple point-in-time snapshots to provide full, incremental, or differential automated instant backup to disk. This allows administrators not to have to spend needless hours trying to recover accidentally deleted records or virus-infected files from tape, even in the event of a disaster.

Snapshot agents minimise recovery time. When data replication is used in conjunction with snapshot agents, the data has full transactional integrity in addition to point-in-time consistency. This means the replica can be immediately put into active use without going through a 'consistency check' process that can be very time-consuming for large databases. Simply put, disaster recovery and replication solutions require flexibility, ease of deployment, scalability and rapid recovery.

### **Reducing cost**

With the IP-based remote replication, data is replicated from one site to another over any existing LAN, MAN or WAN network infrastructure. Data centres use Fibre Channel for their storage networks and a simplified DR solution must provide interface to MAN and WAN routers without the need for extra FC-to-IP converter boxes. These solutions also need to be independent of the type of storage subsystems, application servers and operating system platforms rather than work only in a homogeneous environment. By eliminating the need to deploy matching disk arrays, file servers or application servers at the secondary disaster recovery site (except for hot standby scenarios), such systems can offer unprecedented flexibility in creating disaster recovery environments and allowing for low-cost DR planning by using low-cost JBOD or ATA based RAID arrays at the DR centre.

### **Accelerating pre- and post-disaster operation**

During initial set up of pre-disaster operation, synchronisation of data between the two sites can be done using tape backups or local mirroring, shipped to the remote site, followed by a delta-sync process to facilitate minimal transfer of data over expensive WAN links. During an emergency, an IP link may be deployed to allow emergency access of the data at the disaster recovery centre over WAN by any server located anywhere. This IP connectivity provides significant advantages, offering more ways for temporary offices to access data during an emergency. A reverse delta-sync process facilitates fast recovery of the primary site when the emergency is over.

### **Flexibility in preparing for a catastrophe**

To defend against site failure by providing automated off-site data protection, such solutions offer fast remote data synchronisation of data on disks — across the street, across town, or across the globe. In case of a catastrophic failure at the primary site, the system administrator must be able to quickly redirect application servers to access data from replicas located in the backup data centre. Administrators can specify a variety of policies to control the replication process, giving them a very granular, policy-driven mechanism for keeping an extra set of data off-site for disaster protection.

### **Summary**

Remote data replication is no longer an option, but a necessity for corporate enterprises. The amount of data at risk and the cost to replace that data – if in fact it can be replaced – highlight the need for a data protection solution that extends

beyond a building, a campus or even a country. While most customers recognise the need for this type of data protection, they are squeezed by the continued tight IT budgets and the cost and complexity of existing solutions.

Simplified disaster recovery solutions must be based on a customer-centric philosophy, addressing the cost and complexity issues that have kept many customers from successfully deploying remote data replication solutions. The complexity problem can be solved today by using fully integrated replication solutions built on performance-optimized appliances. By deploying data replication, mirroring, backup and multiple snapshot storage applications on an affordable appliance platform, a comprehensive DR and replication solution can address the key customer requirements of flexibility, ease of deployment and scalability.

*Ravi Chalaka is vice president of marketing, MaXXan Systems, Inc.*

<http://www.maxxan.com/>