the Availability Digest

www.availabilitydigest.com

Does Data Replication Eliminate Backups?

November 2010

Data replication has become the standard way to keep the database of a standby system synchronized with its production system. In such an architecture, there are two copies of the database – one at the production system and one at the geographically remote standby system.

Because there are now two independent copies of the database, is there no need to back up the database to magnetic tape or virtual tape, especially since the backup copy is only seconds or minutes old, not hours or days?

The fact is, a company would be foolish not to perform periodic backups. Data replication does not protect data. It protects system operations. It is backup that protects data.

A Review of Data-Replication Methods

Logical Data Replication

There are many techniques for data replication. Logical data replication typically replicates changes within the scope of a transaction. Therefore, the target database at the standby system is usually guaranteed to be consistent and can be used for query and reporting. Should the standby system need to take over following a source-node failure, all that needs to be done is to roll back incompleted transactions.

Block Replication

Most SANs support block-level replication, in which disk blocks are replicated from the production SAN. However, since the consistency of the database at the production system depends upon the contents of cache, and since cache is typically not replicated between SANs, the standby SAN is effectively in a corrupted state and must be cleansed before the standby system can take over (similar to running chkdsk in a Windows environment).

Asynchronous/Synchronous Replication

Both logical replication and block replication can be either synchronous or asynchronous. With synchronous replication,[1] no data is lost following a source-node failure since no change can be made to a data object unless that change can also be made to all of the data object copies across the application network (though synchronous block replication may lose data that is still in the production SAN's cache).

[1] Synchronous Replication, *Availability Digest*; December 2006. http://www.availabilitydigest.com/private/0103/synchronous_replication.pdf.

Some data may be lost with asynchronous replication[2] since data is replicated only after it has been applied to the source database – data still in the replication pipeline may be lost. Real-time asynchronous replication engines can limit this loss to fractions of a second because they replicate changes in real time as soon as they happen. Some asynchronous replication engines are scheduled and may lose minutes of data following a production-node failure.

Active/Active Systems

The ultimate in replicated databases are active/active systems.[3] In these systems, multiple processing nodes are all cooperating in a common application. Their local copies of the application database are kept synchronized via bidirectional replication. Whenever a change is made to one database, that change is immediately replicated to all of the other database copies in the application network so that all processing nodes have the same view of the application state.

A Review of Database Backup Methods

Magnetic Tape

Classically, databases have been periodically backed up to magnetic tape. This provides a level of data protection should the production system fail. A cold standby system can be brought into service by first loading onto it the last backed-up copy of the database. The applications can then be started, the network switched, and the system tested. At this point, the backup system can be put into operation to restore IT services to the users.

Today's IT service requirements for high availability have exposed some limitations of magnetic tape backup. The first limitation is the amount of data lost following a production-system failure. Since backing up to magnetic tape is an operator-intensive activity, backups are not frequent. A full database backup may be taken once per week. If will typically be followed by several incremental backups, say once per day. At some point, another full backup will be taken; and the full backup/incremental cycle will be repeated.

The magnetic tapes are typically sent to a secure storage facility for safekeeping in an atmosphere that is environmentally friendly to the tapes.

Should the production node fail, the tapes must be retrieved from storage. The last full backup must first be loaded onto the standby database, and then each incremental backup must be loaded. Only when all incremental backups have been loaded is the standby system ready to be placed into operation.

Consequently, all data since the last full or incremental backup is lost. In our example above, this can be up to a day's worth of data. Furthermore, for large databases, it can take one or more days to load the standby database. During this time, the application services are unavailable to the users. And this does not take into account what happens if a tape is lost or is unreadable.

Virtual Tape

The problems posed by magnetic tape backup are greatly alleviated by virtual tape. Virtual tape replaces magnetic tape with disk. Instead of writing backups to tape, virtual tape systems write

[2] Asynchronous Replication Engines, *Availability Digest*; November 2006. http://www.availabilitydigest.com/private/0102/asynchronous_replication.pdf.
[3] What Is Active/Active?, *Availability Digest*; October 2006. http://www.availabilitydigest.com/public_articles/0101/what_is_active-active.pdf.

magnetic-tape images to a geographically remote disk. In fact, to the system being backed up, a virtual-tape backup system looks exactly like magnetic tape drives. The system is unaware that it is backing up to disk instead of to tape.

Virtual tape brings many advantages to the backup process when compared to magnetic tape. Since backup to virtual tape is faster and can be substantially operator-free, it is convenient to back up the database more frequently. Full and incremental database backups are often done every few hours rather than days.

Restoring from disk is much faster than restoring from tape. What might take days to restore from tape can be reduced to hours using virtual tape. Furthermore, restoring from disk is much more reliable than restoring from tape. There are no problems with lost or unreadable tapes.

So Why Bother with Messy Backups?

The replicated database is not only a consistent copy of the production database, but it is also only seconds or minutes old instead of hours or days old. Therefore, only seconds to minutes of data will be lost following a source-node failure (zero data loss if synchronous logical replication is used). Recovery time can be measured in minutes to hours rather than in hours to days.

The replicated backup is far superior to any backup that is achievable by tape or virtual tape. So why even bother with these backup methods? Relegate them to the dust bin of technology past.

The answer is simple. Woe be to those who elect not to back up. Here is why.

Database Corruption

Perhaps the biggest reason is database corruption. If the source database becomes corrupted, the corruption will be replicated to the standby database. Now both databases are corrupted, and the database is lost.

There are two types of replication corruption – data and structural. Data corruption is typically caused by an application bug. Both logical replication and block replication will replicate data corruption. The problem here is not so much a lost database as it is wrong data in the database. Both the production database and its standby copy are in error, and there is no way within the replication environment to repair the error Even the transaction log files, which usually provide a way to back out bad transactions and repair the database, will contain the corruption.

Structural corruption is more serious. This occurs if the structure of the database should be damaged. Structural corruption is not likely to be replicated by logical replication. However, it will be replicated by block replication since the exact copy of the contents of disk blocks are replicated. If both the production database and its standby copy experience structural corruption, the database is probably unrecoverable.

In both cases, database backups come to the rescue. By restoring the database to a certain point in time prior to the corruption, the database can be returned to a correct and consistent copy. True, some data will be lost; but this is better than losing the entire database.

A case in point:

A major bank in the South Pacific ran three redundant nodes for its critical ATM, POS, and online banking services – a production node, a disaster-recovery (DR) node, and a development node. The development node could be pressed into service as the production node if need be.

In December, 2008, an operating-system patch was made to the production system to correct a processor problem.[4] The patch had worked on earlier versions of the operating system to correct the problem but had not been tried on the current version being run by the bank. The bank installed the patch, and it seemed to work fine.

However, the patch was actually causing write errors, which corrupted the production disks. These errors eventually brought down the production system. When the bank tried to fail over to the DR site and then to the development site, it found the same problem. The corruption had been replicated to all of the systems.

Unfortunately, the bank had not made backups. It had no way to restore the database. It was able to get some data from unrelated systems and from some of its partners. Partial operations took over three weeks to restore. However, much of the database was never recovered. It took the bank months to resolve all of the disputes.

Dual Database Failures

It's hard to believe that a redundant disk can have a dual failure that will take down the entire disk subsystem. The chance that a pair of mirrored disks will fail is one-in-a-million if each has an availability of three 9s. A dual SAN, each with four 9s availability, is one hundred times more reliable. Yet such failures happen. Woe to the enterprise that hasn't backed up its database if this highly unlikely event should happen. It has lost both its production database and its replicated copy.

American Eagle, a multibillion dollar clothing retailer, experienced just this fault. In July, 2010, its web site came crashing down when its primary SAN failed.[5] As it attempted a failover to its backup system, the unthinkable happened. Its standby SAN crashed. No problem – American Eagle attempted a failover to its remote DR system only to find that its outsourcer, IBM, had not yet brought the DR system into operation. Fortunately, it did have magnetic tape backups; but it took four days to restore the purchasing functions of the web site and another four days to get its ancillary online facilities back into operation.

More often than not, a dual-storage outage is due to a maintenance error. DBS Bank, the largest bank in Singapore, started getting alert messages from its primary SAN in July, 2010.[6] The support group deduced that it was a cable problem, and a cable replacement was scheduled for the wee hours of the morning. The local service technician decided that he could probably fix the cable and started fiddling with it. The result – he took down the standby SAN as well. Gone were the bank's online services, ATM services and POS services. The bank was luckier than American Eagle – the database was still intact, and they were able to restore services in ten hours. Nevertheless, they were hit with a $230 million penalty by the Singapore Monetary Authority.

The State of Virginia was not so lucky.[7] In August of 2010, a controller board failed on the state's primary SAN. When a maintenance technician started the repair process, he pulled the controller board from the good SAN by mistake. Fortunately, the database was backed up on magnetic tape. However, it took the state over a week to rebuild the database; and it lost up to four days of data. For an entire week, twenty-six of the state's agencies were down, including the Motor Vehicle Bureau, Social Services, and the Department of Emergency Management just as Hurricane Earl was approaching.

---

[4] Innocuous Fault Leads to Weeks of Recovery, *Availability Digest*; December 2008. http://www.availabilitydigest.com/public_articles/0312/simple_fault.pdf.

[5] American Eagle's Eight-Day Outage, *Availability Digest*; September 2010. http://www.availabilitydigest.com/public_articles/0509/american_eagle.pdf.

[6] Singapore Bank Downed by IBM Error, *Availability Digest*; August 2010. http://www.availabilitydigest.com/public_articles/0508/singapore_bank_outage.pdf.

[7] The State of Virginia – Down for Days, *Availability Digest*; October 2010. http://www.availabilitydigest.com/public_articles/0510/virginia.pdf.

Point-in-Time Restoration

A major advantage of backups is recovering accidental or malicious deletions of files or tables. With a backup, you can return to some prior point in time and recover a lost or corrupted file or table.

JournalSpace was a major blogging site. In December, 2008, JournalSpace summarily fired its IT manager for stealing from the company.[8] On his way out, the disgruntled employee did a slash and burn on JournalSpace's entire SQL database, overwriting it with garbage. Only then did management discover that the manager had never fulfilled his duties to back up the database. After exhausting all options to recover the database, JournalSpace went out of business the next month.

I had my own experience with lost data. I use Carbonite to back up the Availability Digest web site. After publishing the September issue, I moved to a new computer and used Microsoft's Easy Transfer to transfer my files over my wireless connection from my old computer to my new one. A little over a month later, I started on the October Digest only to discover that Easy Transfer had not transferred the web site files. No problem – I went to my Carbonite backup to get them only to discover that Carbonite did not have them either. A call to Carbonite customer service "explained" the problem. Carbonite holds backups for only 30 days and then irretrievably deletes them. The customer representative stated that after all, Carbonite is a backup service, not a storage service! Backups are no good after they are discarded. Fortunately, I was able to upload the files from my hosted web server.

Archiving

Backups are the only way to archive information for long-term storage. This is often required by corporate policy or by regulatory requirements.

Security Auditing

Archiving is also an important facet of security auditing. Should you or the auditors discover suspect activity, the archive is the only way to find out when it started, who did it, and what the impact was.

Protecting Data That Is Not Protected by Replication

Data replication doesn't always protect everything. Often, to make the most efficient use of processing and network capacity, only that data that is deemed critical is replicated. There is no standby copy of the less critical data, though it is important to the smooth operation of the enterprise.

Do you know what data is replicated and what is not? As new applications are added, do you keep track of the protection of that data?

The only safe way to protect all data is to back it up periodically so that it can be recovered following its loss.

[8] Why Back Up?, *Availability Digest*; April 2009.
http://www.availabilitydigest.com/public_articles/0404/journalspace.pdf.

Peace of Mind

Finally, there is peace of mind. As one person told me, no matter how safe he feels with replication, he does not want to have to stand in front of the Board of Directors and explain how he lost the company's data.

Backing Up – Magnetic Tape or Virtual Tape?

So backup is imperative. There is great risk to your data and to your company if you do not perform periodic backups.

But what kind of backup? We have seen from some of the above stories that even when companies backed up their databases on magnetic tape, it could take them days to restore operations following a major failure. Virtual tape improves recovery time significantly, typically reducing recovery time from days to hours. It also improves reliability of backup since there is no concern about lost or unreadable tapes. In addition, since backups can be made more frequently, the amount of data lost can be reduced from days to hours.

Magnetic tape does have a role, though, and that is long-term storage. Maintaining backup tape images on disk is typically valuable only for a limited amount of time. Depending upon the application needs, that time may be months or more. There comes a point when the need to rapidly access backed-up data for point-in-time restoration or for audit purposes loses its value. At that time, old data should be moved to magnetic tape to save money and energy.

A recent study by The Clipper Group[9] has concluded that in a typical long-term archiving scenario, a disk archive will cost about 23 times as much as a tape archive and will burn 290 times as much energy. If a virtual tape library with a 20:1 deduplication factor is used, disk archiving will still cost about five times that of tape backup.

What About the Cloud?

Cloud computing is the new paradigm today. Many cloud providers offer storage services in the cloud. The advantage of using cloud storage is that there are no operational worries – just pay the bill.

Unfortunately, cloud storage has yet to be proven a reliable backup medium. We read every couple of months about some cloud storage provider that has lost part or all of the data it is holding.[10] Unless you can reconstruct your data somehow, be aware of the cloud – it can be dangerous to your health.

Kodak is a particularly honest online storage provider. On its web site, it urges customers to keep a copy of each image they upload to the site in a separate and secure place.

Summary

Data replication does not protect data. It protects system operations. Should a system fail, rapid recovery can be made to a standby system with a current application database that has been maintained in synchronism with the production database via data replication.

[9] David Reine, Mike Kahn, Disk and Tape Square Off Again – Tape Remains King of the Hill with LTO-4, *Clipper Notes*; February 13, 2008.
[10] The Fragile Cloud, *Availability Digest*; June, 2009. http://www.availabilitydigest.com/public_articles/0406/fragile_cloud.pdf.

However, if the production database gets corrupted, or if a file or table is lost, data replication provides no protection. If there is a simultaneous failure of both the production and standby databases, data replication provides no protection. Only backup provides this protection.

Therefore, the database must be backed up. Near-term backups should be kept on disk for rapid recovery and reference. Long term archiving of data should be on magnetic tape for economy.