the Availability Digest

www.availabilitydigest.com

Backup Is More Than Backing Up

May 2009

So you think that your corporate data is safe? Your company's data is its lifeblood. Lose it without the chance of recovery, and your company will likely go out of business.

Will Your Backup Be There?

But protecting your data is more than just backing it up periodically. This was graphically illustrated by the misfortune that befell JournalSpace. As described in our Never Again article Why Back Up?, published in our last issue (April 2009) of the *Availability Digest*, we discussed how JournalSpace was put out of business because of the failure to back up its database. After five years of rapid growth, the popular blogging host lost its entire database due to the nefarious actions of a disgruntled IT manager.

JournalSpace management had been assured by its IT manager that its data was properly backed up. However, his concept of "back up" was to provide a reliable database via a RAID system. After all, since the database would survive a disk failure, in effect a second copy was always available. In JournalSpace's case, they did not lose their database due to a dual disk failure. Rather, they lost it because the disgruntled IT manager wiped it out as he was being fired.

Will Your Restore Procedure Work?

In response to this article, one reader wrote:

"*Just looking over the digest - very good stuff. I must admit the account of JournalSpace is quite horrific. And make the additional point that even if you do backups, you must test them. Just because you have backed up your data doesn't mean you can restore it!!!!! regards DarrenM*"

Repeat – *Just because you have backed up your data doesn't mean you can restore it.*

This lesson was learned through the painful experiences of a credit-collections company described in our February, 2007, article entitled Don't Wait for the Other Shoe to Drop. The company faced a catastrophic failure of two RAID disks just before the holiday season. No – problem – the IT staff had diligently backed up the database using a remote backup service, and they had a spare server that they could put into production.

At this point, they realized their problem. They had never tested the backup procedure with the spare server. The spare server was configured differently, and the backup kept giving an "internal consistency error" message. It took three weeks to fully recover the system.

Will You Know How To Restore Your Database?

Another reader responded with an even more subtle problem:

"*RE: Never Again, Why Back Up?*

Another very interesting and … unfortunate experience.

I realized something as I read this article - that even situations that are malicious in their make-up can be guarded against just like those caused by natural disaster, hardware failure or just plain user-stupidity.

Here's what I also realized – a real flaw in our catastrophe plan. First, the good and the bad (for what it's worth):

- *We do have better than average hardware redundancy:*
  - *RAID array on all servers.*
  - *Multiple power supplies on most devices.*
  - *Unfortunately, we do not have dedicated AC circuits and multiple UPS devices for the multiple power supplies…*
- *We also do not have a backup T-1 (our main weak point as we experience about 1-2 hours of downtime per year on average).*
- *We do have local daily backups shared across servers and onsite storage devices, and we do incremental backups offsite daily and keep several versions of every file and database.*
- *We periodically unit-test various restore and recovery procedures to make sure we can actually get back the data if we lose it.*

Now for the flaw:

- *The flaw is the IT guy in this case (as you will soon see).*
- *I know how to restore just about anything we have.*
- *BUT each various restore comes with its own specifics, namely 'locations', 'directions,' 'procedures,' and, of course, 'quirks' (exceptions). Let's call these the 'procedures.'*

Half of these procedures are documented – some in a binder on a shelf, some on a server, and some I just know how to do from experience. For all intents and purposes, even if I lose the server procedure docs <u>and</u> I lose the binder, I could still restore our system (files, databases and websites) with the assistance of only one other guy we have under contract.

The fundamental flaw is – me. What if I die in the building when it collapses on top of the servers? What if I am hit by a bus? Nothing is documented in full or in one central location in such a fashion that an outside service could be hired to walk in and start the restore process.

Unfortunately, my boss probably won't allow me to spend the time properly documenting the process as he doesn't see the importance of it… until a server goes down and I show up dead for work that day."

In a separate commentary, this reader goes on to say:

"I think the foolish thing that most smaller organizations do is just never follow through with the full gamut of procedural completeness (testing, documentation, rehearsed rollbacks) – not so much an attitude of 'it can't happen to us.' but more an attitude of 'I can cut a couple of corners here

and bottom line (dollars and cents), I think we can come out a little bit ahead.' Cross your fingers and go with it."

These statements say it all. If the staff member who knows how to do your database restore is not available, will you be able to recover? Even if the restore procedures are documented, are they correct? Are they clear? Documented procedures must be tested. This means having someone who is not familiar with the procedures (though technically qualified) follow them to ensure correctness and clarity. Such a test will uncover incorrect steps, steps that are not clear, and steps that, though obvious to the documenter, have been left out.

Do You Know That Your Backups Are Actually Being Taken?

Another of our readers[1] relates the story of a small company that entrusted its end-of-day backups to a young operator. To ensure that the backups were faithfully taken, the operator had to maintain a log of her backup activities and the IDs and contents of all of the backup tapes.

Unfortunately, this young operator was more concerned about her after-hour activities than she was about her backup responsibilities. When one fateful day the company lost its database and had to restore it from the backups, they found to their chagrin that there were no current backups. The log entries that she had made were all bogus.

This raises the issue of backup auditing. How do you know that your backups are, in fact, being made and made properly? This is an issue which must be faced.

Backup Issues to Consider

Recovery Point Objective

Fundamental to all backup policies is a company's recovery point objective, or RPO. This is the amount of data loss that can be tolerated following a failure. It is a compromise between the cost of backup frequency and the value of the data.

Some applications can tolerate hours or even days of lost data, especially if transactions are not high value or if there is a manual process that can be used to reconstruct the lost data. Other applications can tolerate only minutes or seconds of lost data, while others can tolerate no lost data.

If magnetic tape or virtual tape is used, the amount of data that will be lost is that generated since the last full or incremental backup, ignoring lost data on tape that didn't get out of the building before the fire. This typically represents an RPO measured in hours or days. Most online backup services also fall in this category.

If the RPO is to be measured in minutes or seconds, asynchronous replication is generally the solution. If no data can be lost (RPO = 0), synchronous replication must be used.

In most enterprise data centers, different applications will have different RPO requirements. The backup procedures for some applications may well be different from those of others.

Backup Availability

In the event of a disaster, will your backup be available? If you back up to tape, those backup tapes should be immediately moved to an offsite location. In the case of the terrible 9/11 disaster

[1] Thanks to Phil Grossman.

in New York City, there was at least one case in which all the backup tapes for a company's database fell with the rest of the World Trade Center.

Tape backup is the classic backup solution, but it has some unique problems. Tapes have been known to be lost due to a vehicle accident or theft as they are being transported to or from the offsite facility. When a database restore is needed, the tapes must first be retrieved. If they are stored in some third-party vault, you may have to wait a while before they can be delivered to you. Once delivered, more than one restore has failed because a tape was unreadable. How many times do you allow a tape to be reused? Do you make two copies of a backup?

If you back up via virtual tape or via data replication to disk, are you backing up to a remote site that is far enough away so that it will not be destroyed along with your primary site due to some common disaster?

Backup Audit

Backups to remote disk via virtual tape or data replication can be automated so you know that they are being done at the specified intervals. Large tape farms are also generally automatically audited so that you know that your IT staff is appropriately taking backups.

However, thousands of small businesses backing up their data via tape are depending upon their IT staff to be diligent in doing so. What safeguards do they have in place to ensure that backups are not only being done according to their backup policy but are being safely stored offsite?

In today's new world of the Internet, a very useful alternative to tape backup is to use an online backup service. These services automatically and transparently search your server for changed files and upload them to their servers over the Internet. At any time, you can download selected files to replace lost or damaged files. Many servers even store several revisions so that you can restore a file that has been erroneously overwritten. At the *Availability Digest*, we use such a service, Carbonite. Of course, you must ensure that the security of data provided by such an online service meets your security policies.

Restore Testing

Restore procedures are useful only if they work. Just like failover to a backup system, restoring a backup database must be periodically tested. If there is a backup system, this can be done by restoring the database to the backup system and ensuring that it works. How this is done depends upon the backup technique used.

- If magnetic tape is used, the latest backup (the latest full backup plus any subsequent incremental backups) must be retrieved and used to reload the database. If a third-party offsite tape storage service is used, this is a good opportunity to ensure that the provider can meet its service-level agreement for prompt delivery (do you have one?). Furthermore, this is the best kind of audit that can be made to ensure that backups are being properly taken.

- If virtual tape or data replication to a backup disk subsystem is used, that subsystem should be mounted on a backup system and ensured that it is correct.

- If an active/active system is being used, one in which there are two or more geographically-distributed copies of the database in active use, then it is known that the "backup" is, in fact, proper; and no testing is required.

Restore Documentation

As our reader confessed, proper restore procedures are only useful if they can be properly performed. True, the IT operator charged with backup/restore responsibilities can probably do it flawlessly. But if he is not available, is there someone who can act in his place?

This can be assured only with proper documentation. How do you know that the documentation is proper? Only by having someone not familiar with the documentation attempt to do a restore with no help whatsoever. If he stumbles, correct or clarify the written procedures.

Backup Documentation

Of course, all of the restore capability in the world is not going to do any good if there is no database backup available. What happens if the one IT person you have on board who knows how to back up gets (God forbid) hit by a truck? Is there someone who can take over his role?

The way to ensure the survivability of a backup capability is to make sure that backup procedures are documented and tested just as are restore procedures.

Protecting Against Malfeasance

This leads us to our final backup issue, and that is nefarious or accidental acts that can destroy a database and its backups no matter how well they are planned and executed. Though it may be hard to do, one must be paranoid about backup security. Not only is it necessary for management to be well aware of backup policies and their proper execution, but the possibility of a distraught employee or perhaps even an external hacker destroying the backup data must be carefully considered.

At the very least, the backup data must be secured. Best practices would dictate that it would take the concurrence of at least two people to modify a backup database in any way. This makes sense if the backup is a remote disk subsystem since a security scheme requiring two independent passwords can be devised.

However, what about the case in which an employee is charged with taking magnetic tapes offsite and perhaps storing them in his house, which is often done. If he is the one that is disgruntled, good luck on getting the backups returned when needed. Perhaps the policy ought to be that two employees are involved, taking possession of backup tapes on alternate days. Then, at least, the exposure is limited. Of course, storing tapes using a third-party storage service solves this problem.

A corollary to this relates to active/active systems. Any operator command that can be issued against all databases in the application network should be disabled. Access to each copy of the database should be separately protected and accessible only to different personnel,

Summary

If your company's survival depends upon its data (as many do), proper backup/restore procedures are a matter of corporate life or death. This requires not only that the backup strategy support the required RPO but that all backup and restore procedures be documented, tested, and audited.