STORAGE SWITZERLAND

HOW TO GUARANTEE 100% RECOVERABILITY OF YOUR BACKUP

Storage Switzerland, LLC

George Crump, Senior Analyst

The single biggest challenge for IT personnel involved in the data protection process is making sure that their backups are recoverable every time. Management and users won't remember the ninety-nine successful recoveries but they will always remember the one failure. As a result data center managers are looking for ways to improve their confidence in the recovery process - they have to know recovery is going to work every time.

It's in vogue for backup vendors to market with the message "It's all about recovery". This is true, that in the overall data protection process, data recovery is the most important step. But recovery confidence is gained through the use of tools that will verify that ability to recover data. In this area vendors are found surprisingly lacking.

The Problem With Legacy Verification Methods

The confidence challenge stems from the fact that most legacy data protection software is limited in the ways it can verify if backed up data can truly be recovered. And the methods they provide to recover are not very accurate. The old standby is backup verification. For legacy systems this means a block by block comparison of data on the backup media with data that's on the application server. There are two significant problems that make this type of verification a problematic process.

The first problem is that it effectively doubles the backup window since every block of data must be brought across the network again and compared to the first backup. Most data centers don't have the available backup window to double the amount of time that this comparison takes. For this reason alone verification is often the first feature that's turned off. At this point the backup administrator is flying blind, or at least flying on hope.

Even if time can be allotted for this process, there's a second problem with the old verification method - it's not always an accurate representation of recoverability. Since it is comparing blocks from the source with the target it's unable to detect data corruption. For example, if the original source blocks were corrupted prior to backup and the backup accurately captured those blocks, then the verification will simply confirm that there is a 100% chance of your being able to recover data - corrupted data.

Legacy Testing

The other form of verification involves actually performing a recovery. Most data centers, thankfully, don't have the requirement of recovering servers every day. But this method of verification means that recovery has to be tested and practiced. The problem is that the testing is too laborious and requires finding a test server that's properly configured and ready for data to be restored. This process with legacy systems can take hours. Virtualized server environments do simplify this somewhat, but only with respect to finding a test server, since a new virtual server can be created relatively easily. All the data still has to be recovered.

The time and effort required to test this ability to recover has made these tests into major events that happen quarterly or even yearly for many data centers. This means that if an error has crept into the data protection process there's no way to know until the next testing window comes up, which could be months away. This type of testing also means that there is often a significant cost and work stoppage with the testing process itself. In short, it becomes an event that no one on the IT team looks forward to or expects to work correctly. Change is the only constant in a data center and change breaks processes, especially processes as intertwined as data protection. Since these errors cannot be ironed out at the moment when they occur and since identification has to wait until the next testing event, it's harder to remedy them.

The issues with true verification and the lack of testing are responsible for most organizations' lack of confidence in their ability to recover. This lack of confidence is well founded, since confidence comes from successful experiences. If IT faces major events filled with failure then it's no wonder that their confidence level is so low.

Achieving 100% recovery confidence will require a new generation of data protection capabilities. These will focus on more intelligent verification of protected data and more intelligent recovery testing so that the test process can be an ongoing one, not a once-a-quarter exercise of hoping it all works.

Next Generation Verification

The next generation of backup applications needs to go a step further by understanding the file systems and applications they are protecting in order to verify 100% recoverability and simplify the testing process so that it can be done almost in real-time. Companies like AppAssure are delivering a new level of granularity with their backup technology which understands the individual specifics of the server applications themselves.

The first step in this advanced verification process is to perform a 'mount-ability' check. This check, as the name implies, makes sure that the file system or protected application can be mounted. It's relatively straight forward and one that a modern backup application should perform at the end of each backup process.

Next generation verification goes beyond making sure the volumes can be mounted but also that the data on those volumes is useable and is itself mountable by the applications. This again means an application awareness on the part of the backup software to run specific application data consistency checks after the completion of the backup.

Next Generation Testing

In a perfect world, applications will never fail and disks will never stop spinning. The fact is that given enough time, failure is inevitable. A backup product is a critical component in protection from that eventuality because it protects all of the data that needs to be recovered when something goes wrong.

It is difficult to gain confidence in the ability to recover data while waiting for a recovery process to take place. As the saying goes “practice makes perfect”, but practicing a recovery is no small feat in the data protection process. Next generation testing means having the ability to recreate a server as a virtual instance on a moment’s notice.

Next generation testing, available from companies like AppAssure, allows for servers to be exported as virtual instances or virtual machines (VMs), a process called “creating a virtual standby”. With the click of a button the VM can be created, restored to and then tested against to make sure that recovery will function properly. This process takes minutes instead of the days involved in the legacy testing methods described above. Problems or changes to the environment that may have caused a recovery failure can be quickly caught and corrected while those changes are still fresh in everyone’s minds.

Summary

Recovery confidence comes from making sure that the data being captured by the backup application is in a usable format. The next generation of backup applications though, has to move beyond a bit comparison and make sure that the data being captured is actually usable by the file system or the software application. This provides a day-to-day confidence in the data being protected.