

BSI Reliable Storage

About half of all companies that lose their data never reopen their doors and 75% are out of business within 3 years. Every company therefore needs to have a comprehensive strategy for protecting its data. Bristol Systems Inc recommends a layered approach to protecting your data that can be adjusted to fit your budget.

There are several potential problems to worry about and there are several techniques to deploy. In the next section we survey the problems separately from potential solutions, and then in the following section we discuss potential solutions.

Problems

There are a variety of ways that data can get corrupted or destroyed. You need to protect your data against all of them in some way.

Problems begin with user errors. It is all too easy to accidentally delete or corrupt a file. Lacking a simple way to recover an old version, many users just redo their work. If you can easily get back an older version of your file, then you may save yourself some work.

Frequently a short power outage or a drop in voltage causes a computer to crash losing the work in progress. Such an outage can also cause corruption and even permanent damage on any hard disk in your network. It is not unknown for someone to accidentally unplug a power cord or throw a circuit breaker causing the same problem. Sometimes people reach down and turn off their computer using the power switch causing the same damage. (To turn off a computer safely, close all files, shut down all applications, and then use the shut down command from the operating system. This takes more time, but it is worth it.)

Hard drives are amazingly reliable, but when they fail, all data on them are usually lost. While a backup mitigates this, the time to restore and the cost of lost data not captured by the last backup are still expensive problems. Also, most companies have their data spread across many hard drives – when you consider all the company's servers and workstations. Of course, the more hard drives you have, the more likely that one of them will fail. Hence, you'd better plan on a drive failure!

System failures – when all or part of an entire system is out of order – are relatively frequent. Computers today are quite complex, and the probability of a component failure is relatively high. Such a failure may or may not cause damage to the data stored in the system; however, it almost always makes that data inaccessible until the system is repaired.

Fires, earthquakes, floods, and other disasters can wipe out an entire data center. Most likely such a disaster will also destroy backups of any type that are kept at the same site. While the physical equipment can be replaced (albeit expensively), all is lost unless there is adequate backup data stored at an undamaged site. Obviously, there is considerable variation on the size of the problem when addressing disaster recovery. If the entire business needs to be functional within hours or even a few days after the disaster, then nothing short of a duplicate, functioning data center will do the job. In this case, the issue becomes how to mirror an entire data center.

Sometimes the problem isn't really one of destroyed data, it is one of recovering very old data.

Finally, backups of data can take considerable storage and management. One usually compromises and saves very little of older data. The problem becomes just how much data to archive for retrieval much later than it was created. The longevity of the media is an important consideration in this case.

The following table summarizes the various approaches to these problems:

	Corrupted Or Lost File	Hard Disk Failure	System Failure	Disaster (Fire, Earthquake, Etc.)	Archives
UPS	Helps to avoid corruption and failures due to power fluctuations and outages.			N/A	N/A
RAID	N/A	Keeps the system running without loss of data if one drive fails	N/A	N/A	N/A
Tape Drive	Slow recovery with medium recovery probability. Labor intensive. Tapes age and must be refreshed.			Must be stored offsite.	Cost per gigabyte is optimal.
Removable Disk	N/A	Fast swap in RAID systems	Fast move of all data to duplicate hardware.	Implements sneakernet for large files when bandwidth is low.	Zip disks and CD's are cost effective for small data sets. Hard drives are too expensive.
Local Mirrored Server	With time delay or versioning, provides fast user friendly file recovery and audit trails	Switching to mirror maintains system performance while repairs are made.		Can recover if disaster happens to offsite facility.	N/A
Offsite Mirrored Server		Not as fast as a local mirror, but for large amount of data can be much faster and more reliable than even local tape backups. Can copy to removable drives so that the mirror is preserved.		Has potential for archives especially when online access is needed.	
Redundant Offsite Storage	This is primarily to address geographic disasters such as earthquakes that might damage both the site and the offsite facilities. It also addresses a disaster happening to the primary offsite facility, particularly if archives are stored there. If a mirrored server is used, it can also address some types of network outage.				

UPS. The first line of defense is to protect against the most common causes of damage: power fluctuations and outages. Get a good uninterruptible power supply (UPS) and make sure all attached servers can safely power down when the UPS batteries run low. Make sure your UPS system is regularly tested and maintained.

Protect against this by having a high quality uninterruptible power supply (UPS) protecting each computer – or at least on every file server. A low end UPS can be purchased for under \$100. Such a UPS will protect your computer for a few minutes during a power outage, which will give you a chance to save your data. More expensive UPS's can protect your computers for a longer period of time; they can also automatically shut your computer down after a few minutes or after the battery runs down. The best UPS's also protect against power fluctuations.

RAID. Appropriately configuring a redundant array of inexpensive disks (RAID) can address a single or small number of simultaneous disk failures. It is worth the small amount of cost to implement RAID storage for most companies.

Tapes. A backup system based on tapes works reasonably well, but it is surprisingly expensive if one takes into account the costs of labor and tape media. Tape backups also do a poor job at quickly recovering a file that was accidentally corrupted or deleted. Many users prefer to redo their work rather than to bother and wait for a system manager to attempt to recover the file from a backup.

Mirrored Servers. BSI Mirrored Servers address many of the deficiencies of tape backups. In addition, they provide much additional functionality. BSI Mirrored Servers are not just copies of your servers. They deploy a versioning scheme so that you can recover the version of a file that was mirrored at a particular time. Mirroring is sort of like doing an incremental backup with a detailed catalog of what files were backed up when. Old copies are only deleted based on a user-defined aging policy. The beauty of this is that files can be mirrored automatically at relatively high rates so that most saved changes to files are captured. Excess copies are slowly deleted. For example, changed files can be mirrored every minute, then all but the last version of the day can be deleted every week, and every quarter all but the last version of the month can be deleted. Other, more sophisticated, policies can be implemented based on the type of file, the server or directory on which it resides, etc.

Local versus Remote. Mirrored servers can be local or remote or both. Typically, the mirroring schedule and the aging policy of a local mirrored server is different from a remote one. Local mirrors are usually optimized for high performance in case it needs to be deployed as a file server when the primary file server goes down. Remote mirrors are usually optimized to minimize the network bandwidth needed to access them.

The Bandwidth Tradeoff. Remote Mirrored servers have the disadvantage of consuming considerable network bandwidth. It is easily possible to have enough changed files that a backup to a remote mirrored server can take more than 24 hours consuming 100% of the bandwidth to the remote site. For example, a T1 can move about ½ a Gigabyte per hour maximum. Thus 50 Gigabytes of compressed data would take more than 24 hours to move over a T1. One can get more sophisticated by moving only the changes and then reconstructing the mirrored file on the remote server. This adds cost and complexity to the mirrored server, but it is a good solution. Some companies simply opt for more bandwidth. This is a good solution when bandwidth is cheap.

Archives. Mirrored servers provide archiving, audit trails, and directories in addition to their backup and recovery functions. The archival data and catalogs represent valuable company data

that itself should be backed up. In the case both local and remote mirrored servers are deployed, there is a duplicate copy of this data that will serve this purpose except for the unlikely possibility that both servers get simultaneously corrupted. In any case, either full backups to tape or a redundant offsite mirrored server is recommended, both kept at a third, geographically distant, location. BSI provides at its remote storage sites, high-end tape archiving facilities for the use of its clients. These archival copies are stored in fireproof and earthquake proof vaults at a geographically distant site.

Summary. The absolute minimum protection against data loss or corruption is to have a backup copy of your data. Traditionally this has been implemented using tape backups; although, some small businesses use Zip Drives or CDRW drives. These solutions are prone to user error, since they are not automated and often are not implemented and verified consistently. They also have significant hidden costs. Adding a UPS and using RAID storage can dramatically lower the probability of loss and are almost always cost effective improvements. A BSI Mirrored Server automates the backup process, allows end users to recover old versions of files, ages and deletes unneeded files, provides audit trails, provides easy archiving, and provides fast and reliable bulk recovery. It can be deployed locally to provide a redundant server, or it can be deployed remotely to provide disaster recovery. High-end tape archiving allows not only archival copies of the data to be made; it also provides insurance against disaster happening to both the customer site and the remote storage site. Bristol Systems Inc. can help you customize your BSI Reliable Storage system to meet your needs and your budget. Call 714.532.6776 and ask for a BSI Reliable Storage expert.