Data Protection that Works!

"This gentlemen, is a football." Famous words from a famous coach that may still ring true for CPA's and their clients when it comes to basic computer operations.

It has been over 20 years since the advent of the personal computer shifted responsibility for data from centralized IT departments to end users. Since that time, almost everyone who uses a personal computer has known of someone who suffered a catastrophic data loss or perhaps even suffered it themselves.

This, combined with the growing dependence of business on their information systems raises serious internal control and going concern questions for auditors. What financial loss would your client suffer if their computer systems failed? Does your audit consider the viability and possible business consequences of their backup and recovery strategy? The importance of computers to business has become such that I maintain that CPA's owe it to their clients to review, validate and report on the viability of their data processing system.

Yet, in a recent survey of small businesses, we found that **none** had a working data back up and recovery strategy. So, for CPA's and their clients, "this gentlemen, is a football."

In the early days of personal computing, the biggest risks related to hardware failure. Amazing improvements in manufacturing technology have reduced this risk to almost nil Today, weather and natural disasters seem to be capturing top position in the sources of risk category, but there is another more sinister contender that is about to take first place—theft. Criminals are beginning to recognize the value of information stored on computers for extortion, identity theft, and business intelligence. Today computers, like purses, are being stolen for the value of their contents, and not for their intrinsic values. This means that physical and logical access controls to the computer system must also be extended to any copies of the data or programs as well. This requirement can challenge even the most carefully crafted plan.

The first element in a data processing recovery strategy is backing up the system—programs, patches, and data files.

In the early days of personal computers, this was easily accomplished by creating a secure duplicate copy of specific folders from the principal operating drive. Unfortunately, most modern windows applications scatter essential files all over the hard drive in such a manner as to make their identification for backup purposes almost impossible. Moreover, vendors now provide regular software "patches" that are essential to the continued safe use of their program. This means that recovering a program's complete functionality by reinstalling the original CD's is not possible. All of the patches, which may or may not be readily available, will have to be obtained and installed in the correct order to restore the program.

An effective backup strategy must either encompass the entire operating environment, or be completed in two steps—a data step, and an application step.

Separating the data and application backups makes intuitive sense. After all, the application portion of the system is generally more static, while the data portion changes by the minute. Thus, application backups do not need to be performed as frequently as data backups. This can save time and resources in the backup creation process. It may or may not save time during a data restore operation. But, creating a dual step backup process requires a good understanding of the data that is going to be backed up, because unless ALL the data, no matter where it is located, is backed up, the restoration of the operating environment cannot be assured.

Many users find the complexity of assuring that the data backup actually captures all data to be so overwhelming that they merely opt for a single operating environment backup. As speed of backup devices increases, and the cost goes down, this is not unreasonable.

Let's examine the two models from a process perspective to understand what's involved and the relative advantages and disadvantages of each.

The operating environment encompasses everything in the system, which may or may not even exist on a single hard drive. Backup of this environment is best done by taking a snapshot, or image, of the content of all of the relevant hard drives in the environment at a single instance. There are numerous drive imaging tools that can be used for this purpose. The disadvantage of this method is the amount of backup media required along with the amount of time to create the images. A second complicating factor is assuring the synchronicity of the data. If a work environment is not 24/7, this can be obtained by closing the system at set intervals for backups. Otherwise, special tools will be required to back up files during periods of use while maintaining synchronicity. For imaged backups, data recovery merely involves taking the time to restore the imaged drives and resuming processing at the point of backup. Transaction logs or other documentation should be retained to allow the recreation of transactions that occurred between the current date and the recovery date, otherwise the recovered system will be present and synchronized, but still out of date.

A data backup would generally focus on selected drives and folders known to house data. This would be a smaller subset of the entire operating environment, and can often selectively target certain files based on the nature of the data and processing taking place. Consequently, an image backup may not be necessary resulting in a smaller archive and less time invested in creation.

How often should you back up? The rule of thumb is whenever you get to a point where you would not want to have to manually recreate the data from the last backup. Depending on data strategy and budget, this could be anywhere from simultaneous backups where the data is written twice, to hourly, daily, or weekly cycles.

Testing the backup strategy – at least once a year, IT and users should coordinate a complete recovery test to validate the policy and process will actually work, as well as to refresh the parties on their respective responsibilities. This may be best accomplished on a test system, installed clean. Otherwise, you will not be accurately simulating a worst case recovery situation.

We had a client who had a nice backup system in place for 3 years before we got there. The system had been established by an IT consultant and consisted of a weekly full system backup followed by daily incrementals with rotating tapes over a several month cycle. When we attempted to validate the backup policy, it turned out that the tapes had long since gotten full, and nobody had authorized an overwrite—so each night at the close of business, when the user dutifully switched tapes to the next one in the cycle, the backup would fail because the tape was full, and nobody checked the error logs. In truth, there were no backups and nobody knew it!

Now that you have the backup, how do you protect the data? Loss from fire or hurricane, while clearly unattractive, is much better than loss from theft. Backups, ideally, should be encrypted and stored in a protected area where they should be inventoried at regular intervals. We like the idea of an off-site safe, but it needs to be such that you can get access to it anytime day or night. We had a client once who had their employees take turns taking home the backup. However, this created a huge problem for the company, who could not recover the data if the employee was unavailable or uncooperative. It also opened a new door of opportunity for additional access to the data by unauthorized persons.

In use data will be harder to protect but basic precautions can be put in place—usernames, passwords, system logging, and when possible, data encryption should all be used. Hardware should be isolated from the internet and regular testing of firewalls should be done to validate the protection is still working.

The greatest risk of loss will come when data is moved to lap tops or USB drives. HP, and other manufacturers now provide hardware encryption for hard drives but it may not be turned on. USB drives can be protected with tools like TrueCrypt, a free, open source tool that provides very advanced security features. Regardless, protecting mobile assets will require additional policies, training and enforcement.

We have tried to provide a brief overview of issues and principles for managing an effective backup and recovery strategy. Many of the implementation details will be system and company specific, which is beyond the scope of this article. If you feel you would benefit from a review of your current backup strategy, or would like help to develop one, please feel free to contact us.