



Understanding Your Data Archiving Needs

A White Paper by Ben Whitney

Introduction

Computer files are a crucial part of the daily operations in most companies; files for customer records, email, graphics, databases and business documents are among the essential components of doing business. Every year these files grow larger and more numerous, and even though many files are stale and no longer accessed, there is a possibility that they may be needed some time in the future due to unanticipated needs or external regulatory requirements.

Inevitably, the amount of free space will dwindle on your network servers to a point where the users take notice. The reaction of many system administrators is to add more storage capacity to their network by installing an additional file server. However, this is only a temporary solution as the new servers will still fill up quickly, and now both your new documents as well your old ones will always need to be backed up on a regular basis.

Backup vs. Archiving

Both backup and archiving involve making a copy of your data files, but there are differences between the two concepts.

Backup is an established IT practice because of the undeniable fact that hard drives will crash from time to time. When it happens to a user on their home computer, it is unfortunate. Many people have lost valuable documents and pictures due to a power surge or failing hard disk, and as personally tragic that can be, when this happens to a corporate server the consequences can be much more severe. When servers go down, entire departments can lose their email or network access, important deadlines can be missed, and the net result on the company can be a large financial cost. Backup is preparing your servers and workstations

for disaster recovery, and being able to restore the systems to working order in the shortest possible time. It is much easier to write back an image of a hard disk than to install and reconfigure all the software and reapply service packs. A diligent backup schedule will allow you to recover your computer in the shortest possible time.

Archiving differs from backup in that it is focused on preserving the documents, rather than the computers. It is a targeted approach for providing long-term storage of documents for their historical value. After a computer file is no longer active, it should be made read-only so that its contents can not be altered. Instead of transferring the files to a hard drive or magnetic tape, a WORM (Write Once, Read Many) medium like DVD-R provides a permanent, unalterable record of that file. It is analogous to the recording of old documents to microfiche. The papers themselves may deteriorate but the microfiche will maintain a permanent record of those documents.

Another subtle difference between backup and archiving is that a backup procedure tends to reuse the same media on a rotational basis. The old backup is typically overwritten with the new data. A consequence to this method is that if a file was not included in the previous backup set, it may no longer be recoverable. For this reason, archiving is a good practice to ensure that a file has been permanently saved.

What are the Main Benefits of Archiving?

While backup will help you recover your servers, the servers themselves will still fill up. Data archiving is a part of the solution to this problem. The process is straight-forward: take files which are no longer modified and infrequently (if ever) viewed off expensive hard drive servers and on to an archival medium.

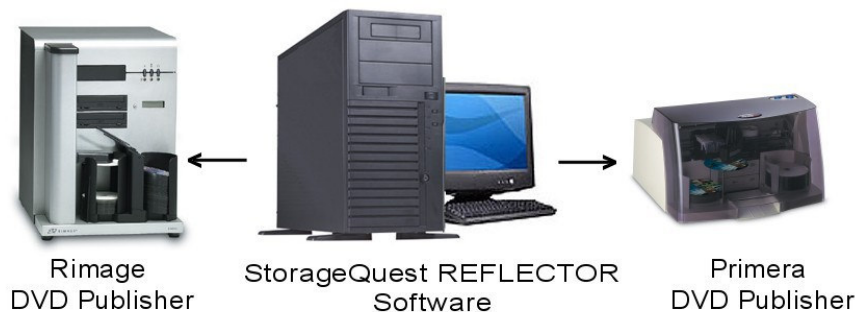
The files will remain accessible for many years in the future and serve as a permanent record of that data. Since you are cleaning old files off your servers and making room for your new files, your need to add additional file servers will be reduced.

New compliance legislation such as the Sarbanes-Oxley act of 2002 has put accountability requirements for all business data to be produced on demand several years in the future, so your files will have to be available if a request comes. By archiving your data, you will be able to certify that your files have not been altered.

Solutions for Archiving

DVD has proven to be one of the most universally accepted forms of media today. With current capacities ranging from 4.7 GB single-layer to 8.5 GB for dual-layered discs, they offer a low-cost permanent storage medium. It is recommended to use the UDF (Universal Disc Format) file system for archiving to optical media as this format maintains maximum compatibility across different optical drives and operating system platforms. Windows™, Linux, Macintosh OS-X™ and UNIX™ all support the UDF standard. By using the UDF file system, data can be stored in a standard form and be accessed many years in the future.

Products like the StorageQuest REFLECTOR provide services for archiving your data – locating it on your network and recording it to DVD. Using the REFLECTOR with a Rimage or Primera disc publisher provides a very convenient way to archive your data in UDF format by automatically recording and printing labels on your DVDs. The data will automatically span as many discs as needed, and all files on each disc will be readable without any special software required.



Using a portable medium such as DVD also allows you to make off-site backups of your data, increasing your ability to recover from unforeseen events. Simply write two copies of each DVD and transport one copy to an off-site location.

Should You Keep Your Archives Online?

Now that you have moved your infrequently-used data off of your hard disk servers and on to a permanent archival medium, you can decide if you want to keep your data online. That decision will depend on two factors: how frequently the data will need to be accessed and how much time is required to spend locating and restoring the correct files.

By adding a read-only optical library to your network, you can provide online access to all your archived data. Products like the StorageQuest MSM 150 will enable plug-and-play connectivity to optical libraries, such as the PowerFile® 200-slot DVD library. By putting your DVDs in an inexpensive optical library, your archived files will remain instantly accessible through a network share when you need them. As for storage capacity, a 200-slot DVD library can hold 1.7 TB of read-only files, while additionally providing a dust and light-free environment to prolong the life of the media. Connecting an MSM 150 to a PowerFile optical library will allow you to easily turn your archive of DVD media into a NAS (Network Attached Storage) resource.



The Importance of Proper Labeling and Cataloging

You will need to properly label your discs. A DVD-R without a label, or any other archival medium for that matter, will not yield any clues to its contents unless that disc is put in a drive. Once the discs have left the recorder, either a marker or a label should be applied to identify the disc. Software programs such as the StorageQuest REFLECTOR will automate the process of recording and printing unique identifying labels on each disc, reducing the possibility of errors made in manually labeling.



At a minimum, the items you should record on a label include name of the archive, a description, date written, the number of discs in the set and the total number of discs.

When retrieval is needed, you will also need to be able to locate which disc your files are on with minimal effort. It is important that any solution is able to maintain a catalog of the location of your files and provide facilities to locate them when they are needed.

Safe Storage of Archival Media

When you have produced your archive media and you need to start thinking about a long-term storage plan, you need to store your data in the right environment. Since leaving your DVDs in the sun in the back window of your car isn't a good idea, you will need to provide an environment which will encourage data longevity and security.

To provide an optimum environment for your discs, keep them in a low-humidity, dust-free and dark environment. Physical security should come in the form of a locked cabinet in a secure server room. Your physical security needs may vary, so take proportionate measures to the value of your data.

Manufacturers of DVD-R discs state that their media will last between 30 and 100 years. Make sure that you use high-quality archival media to ensure the maximum data longevity. Look for terms such as *Archival Media* or *Medical Grade*. For example, TEAC manufactures "Medical Grade" DVD-R discs with a scratch-resistant coating designed for long-term archiving.

The Big Picture

Files in a network migrate from the desktops and inboxes of the users to file servers and high-performance RAID servers. As the files age, the contents often become useful only for historical (or perhaps nostalgic) value. Moving the files to a permanent archival medium, combined with automatic labeling, cataloging and storing the files in an archival environment provides your business maximum confidence that your data will be safe under any foreseeable circumstances. Products like the StorageQuest REFLECTOR and the MSM 150 can assist you with the process of archiving your data.

Adding an optical library will give you the additional advantages of instant network access to your data when the need for it comes, and keeps your media organized and secure. By planning and implementing an archiving policy, you can ensure the longevity of your data for years to come.

Referenced Products

StorageQuest, Inc. - StorageQuest is a developer of high-quality storage management hardware and software. The [StorageQuest REFLECTOR Network Backup & Archive Manager](#) software features powerful CD and DVD network backup and archiving features. The [StorageQuest MSM-150](#) offers plug-and-play network connectivity to any optical library. <http://www.StorageQuest.com>

Primera Technology - Primera is a leading manufacturer of automated CD and DVD publishers. <http://www.primera.com>

Rimage Corporation - Rimage is a leading manufacturer of automated CD and DVD publishers. <http://www.rimage.com>

PowerFile - PowerFile is the worldwide volume leader of DVD storage libraries. <http://www.powerfile.com>