

Coughlin Associates

Data Protection for Small Businesses and Homes

Thomas M. Coughlin
Coughlin Associates, Inc.
408-871-8808 (phone)
www.tomcoughlin.com

July 2007

Table of Contents

Table of Contents	2
Introduction.....	2
Data Protection Techniques	3
Backup	3
Archive.....	4
Continuous Data Protection	5
Storage growth in the home and small business is exploding	6
Storage Tiering for Homes and Small Businesses.....	8
Level 0: No backup.....	9
Level 1: Backup to external storage media.....	9
Level 2: Backup to local network attached storage devices	9
Level 3: Remote storage with scheduled backup.....	10
Level 4: Remote storage service with continuous data protection	11
Level 5: Complete data protection and access.....	11
Some Comments on archiving.....	12
Providers of Internet Storage Services.....	12
File sharing services.....	12
File synchronization services.....	13
Online backup services	13
Online backup for large companies	14
Online backup for homes and small businesses.....	15
Small business and home remote archiving.....	17
Summary of online storage service offerings	17
Future developments for online storage services for homes and small businesses	19
Conclusions.....	19
Appendix A. Glossary.....	20
About the author	22

Introduction

Small businesses and homes need to protect their data from loss as much as larger businesses do. The consequences of losing important company or personal data or even temporarily losing access to that data can be devastating. This document will discuss the different types of data protection that are possible for small businesses and homes, evaluate several sources of data protection services for this market segment and talk about other features in data protection services with examples that can add additional value to users.

Data Protection Techniques

Data protection involves making copies of data that can survive the loss of the original copy of the data and can be used to recover this data whether this is the result of theft or loss of the computer or device containing the data or accidental erasure of a file. There are three basic forms of data protection that are important to small businesses and consumers. These are Backup, Archive and Continuous Data Protection.

Backup

Backup is the process of making copies of data existing in a computer or other digital device that can be recovered if the original data is lost or damaged. Backup can be done with many types of storage devices attached directly or through a network to the host device, usually a computer; that is the backup client. These target backup devices can include tape or floppy disks (for small business and homes this is now very rare), optical write-once disks and external hard drive storage. Important characteristics for backup and recovery are the backup window, the recovery point, the recovery time and recovery point.

The **backup window** is the amount of time needed to complete a backup of the client device. The **recovery point** is the point in time that the backup data can be recovered. Any data created after the most recent recovery point will not be backed up and cannot be recovered. The **recovery time** is the time it takes to recover backed-up information.

Note that backups can be **complete** or **incremental**. A complete backup backs-up all of the data on the client system to the backup storage medium. An incremental backup backs up only the changes made since the last backup (complete or incremental) of the data. Often to recover data with incremental backup one must recover each incremental backup in the order they were made up to the point where the data is to be recovered.

Note that since the bandwidth of the network can be a bottleneck to backup efficiency ways to decrease bandwidth demand such as **deduplication** (backup only of new data rather than copies of material already backed-up) are used to reduce the payload of backup data sent over the network.

Recovery of data from a backup can be performed for single files, groups of files or a complete storage volume. Usually as the amount of data to be recovered increases in total storage capacity it will take longer to recover since it takes longer to transport the data over the network.

Archive

An **archive** is a copy of data that is intended for very long periods of retention, usually in years and in some cases centuries. An archive may be **active** or online where it can be accessed relatively quickly or **cold** or offline where it can take a considerable amount of time to mount the digital storage medium and read the archived data. The average time to access archived data is the archive latency.

Archived data stored on digital tape or optical disks can last on a shelf for many 10's of years with the right choice of components and the proper environmental controls. A constant problem with a cold digital archive is staying ahead of the obsolescence of the storage media used. As time goes on digital storage technology improves and the storage media and drives change. If the data storage on a cold archived media is not migrated to new storage technology when it becomes available there is a danger that the data won't be readable as time goes by. Thus successful preservation of archived digital content must also include data migration management to deal ahead of time with the risks of format obsolescence.

True offline archive media include digital tape, optical storage and even disk drive cartridges. Note that the expected life of the archive media depends upon the quality of that media as well as the environment that it is stored in. It is likely that inexpensive off-the-shelf optical disks for instance may not provide long term data preservation (many ten's of years).

Many small businesses and homes do not distinguish between an archive and backup. For these people their backup data is also their archive. In this case the archive is usually a direct attached or network attached external storage device or a direct attached or network attached disk drive array. To provide a separate tier of backup protection and archiving in a facility or home there may be a separate set of drives used for backup and archiving. The archive in this case also acts as a way to recover data if the backup fails or becomes corrupted. Note that in this case the recovery point from an archive backup is a function of the frequency of archiving data. A true archive should keep as long a history of recovery points and/or file versions as possible.

Disk drive systems used for active archiving typically use SATA hard drives since these provide higher storage capacity for a given price and provide good reliability under less frequently accessed conditions expected in an archive application. Although not likely to be found in small businesses there are even SATA storage arrays called MAIDs (**massive array of inactive disks**) that keep most of the SATA drives in the array powered down most of the time, providing significant power and heat load savings for a data center.

Continuous Data Protection

Continuous data protection refers to techniques for continuously copying data from a computer or other client device over a network to a storage system that is remote from the original client. The purpose of making this copy is that if something happens to the original content (or a backup or archive of that data) the data from the remote site can still be used to recover with a very recent recovery point. Ideally, the recovery of the offline data should be as rapid as possible to avoid significant downtime for the content and applications.

Perhaps the ideal continuous data protection system will be ***mirrored*** (that is when data is written to the original client storage system it is also written to the remote mirror storage system). Mirroring generally must be done almost real time.

If the client storage system fails due to some sort of man-made or natural disaster at the site of the original storage system the mirror storage system can be substituted for it. For the mirrored system almost all of the data that was on the original storage system can be recovered from the mirror in the case of a failure of the original storage system.

These sorts of mirror systems are relatively expensive and generally require expensive dedicated internet lines between the client storage system and the remote mirror target storage system. As a consequence such storage systems are beyond the means of most small businesses and personal users.

Another option is to copy material to a remote storage device at scheduled intervals with recovery points that are very close to real time. If the remote backup is done periodically, then only the data on the original storage system up to the point of the last backup can be recovered. As the remote backup approaches simultaneity with the creation of the original copy, remote backup becomes continuous data protection (CDP). True CDP tries to copy changes as soon as they occur rather than at specific scheduled times. By moving copies of changes to data to a remote site frequently a CDP system provides much greater protection of data from catastrophic events than a traditional backup which may only happen once a day.

The official definition of CDP from SNIA (Storage Networking Industry Association, a storage industry trade group) specifies that any backup technology that calls itself CDP must be able to restore any version of any stored file from any point in time. This is very ambitious and CDP as it is used here aims towards the SNIA definition but almost always comes up short from the perfect CDP data protection definition.

The number of Internet connections in homes world wide are skyrocketing. Likewise almost every business gets an edge from high bandwidth Internet

connectivity. This has lead internet service providers to provide some of the advantages of true continuous data protection through online data centers. These vendors offer backup or copying of user data to their data center where it can be accessed by the user and used to recover data in the case of a failure of the user storage due to drive failure, viruses as well as natural and man-made disasters.

Storage growth in the home and small business is exploding

The explosion in digital storage capacity offered by hard disk drives, flash memory and optical disks has increased the amount of digital storage and stored content in a typical home or small business. Furthermore the low cost of significant amounts of digital storage has put storage into more devices that we have on our desks or carry with us. All the content that is created or stored on these devices needs to be aggregated, supported and protected.

In the home, digital still and video cameras are replacing film cameras and analog tape video. The result of downloading these digital files to computers is that valuable family records are placed on hard disk drives that are part of laptop or desktop computers. Small business records and collateral are increasing becoming digital records and assets and these are also resident on hard disk drives in computers. If the hard drive fails or the files are corrupted and there is not a backup copy of the content, then these unique images or records can be lost forever.

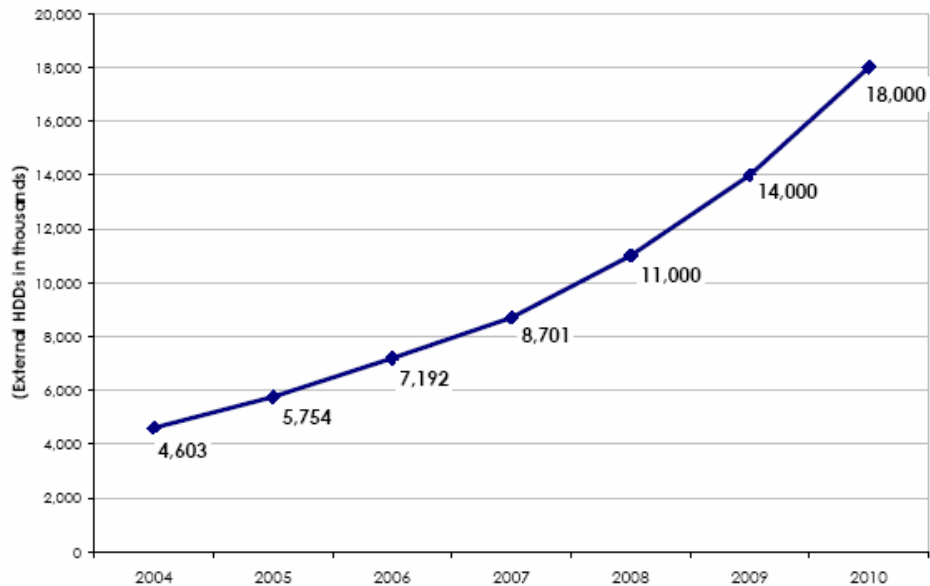
The growth in stored content and high speed networks in the home give incentives to the development of backup and recovery services to serve this market. **Table 1** gives estimates for a reasonable growth in personal content storage requirements for a home having a digital still camera, a digital video camera, and generating personal computer files. Annual personal storage capacity requirements are easily several hundred GB per year today and growing as the resolution of content increases. By 2010 the cumulative unique personal content in a home may reach close to 2 Terabytes and if we presume that people start to create life logs in the next decade this could swell to hundreds of TB of personal cumulative content in a typical tech-savvy home.

As a consequence of the growth of valuable content stored on computer digital storage in homes and small businesses the market for local backup storage has grown and will increase considerably in the next few years. The majority of external backup devices in the home are currently direct attached USB or Firewire (IEEE 1394) hard disk drive based boxes. **Figure 1** shows projections for the growth of external storage out to 2010.

Table 1. Projected growth in home personal cumulative content based on reasonable estimates of content resolution and use (Coughlin Associates, 2006).

<i>Year</i>	<i>2004</i>	<i>2005</i>	<i>2006</i>	<i>2007</i>	<i>2008</i>	<i>2009</i>	<i>2010</i>
Photographs							
Average Storage/Picture (GB)	0.003	0.003	0.003	0.003	0.004	0.004	0.005
Number of Pictures Per Year	500.0	550.0	605.0	665.5	732.1	805.3	885.8
Annual Photograph Capacity (GB)	1.3	1.6	1.9	2.3	2.8	3.4	4.1
Video							
GB/Hour	13.0	14.3	15.7	17.3	19.0	20.9	23.0
Number of Hours	11.1	12.2	13.4	14.8	16.3	17.9	19.7
Annual Video Capacity (GB)	144.3	174.6	211.3	255.6	309.3	374.3	452.9
Documents							
Important Documents Per Year	20	22	24	27	29	32	35
Average Capacity/Document (GB)	0.1	0.1	0.15	0.2	0.2	0.25	0.3
Annual Document Capacity (GB)	2.0	2.2	3.6	5.4	5.8	8.0	10.5
Total Capacity Per Year (GB)	146	177	215	261	315	382	463
Cumulative Capacity (GB)		323	538	799	1,114	1,496	1,959

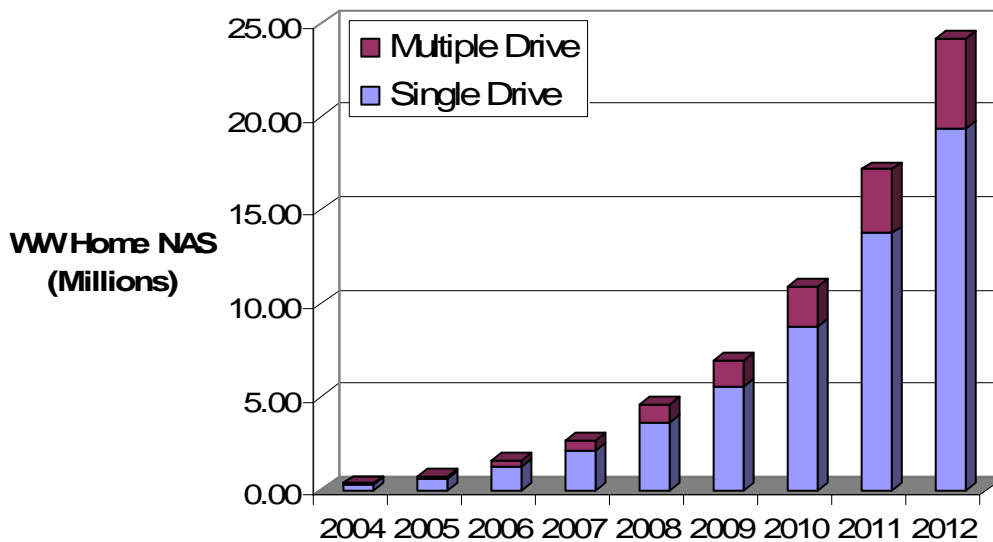
Figure 1. Growth projections for external storage boxes¹



¹ Tom Coughlin, The DNA of the Digital Home: Trends in Home Network Storage, Coughlin Associates and The Diffusion Group, 2005

External storage devices that provide network storage (**network attached storage** or NAS) using wired (Ethernet) or wireless (WiFi) are becoming more common. These NAS devices allow sharing content stored on the NAS device between computers and devices attached to the network as well as backing up these computers and other host devices to the NAS to protect the content should the original copy on the initial storage device be lost due to a crash or file corruption. **Figure 2** shows projections for the growth of networked storage out to 2012.

Figure 2. Projected growth of home network storage²



Storage Tiering for Homes and Small Businesses

Larger companies have the option to create dedicated communication lines for very sophisticated continuous data protection (CDP), version control, file synchronization, system restore, etc. These storage solutions are very feature rich and robust but often require experienced and trained personnel to run them and involve a significant cost in infrastructure and hardware to put into place. These products are out of the range of typical homes or small businesses.

Companies providing sophisticated enterprise level backup and CDP software include IBM Tivoli software, EMC Retrospect and Symantec Backup Exec.

Because of the need for simplicity of operation and low cost home and small business users have had to rely upon usually simpler file protection capabilities.

² Tom Coughlin, 2007 Storage Visions Conference

In this section we divide uses of storage for file sharing and backup by small businesses and homes into six levels of storage tiering. These are

- **Level 0)** Internal hard disk drive alone—no backup,
- **Level 1)** Backup to external storage media,
- **Level 2)** Backup to local network attached storage devices
- **Level 3)** Remote storage with scheduled backup,
- **Level 4)** Remote storage service with continuous data protection,
- **Level 5)** Complete data protection and access including remote storage service with full version control, multiple computer synchronization, perpetual archiving and remote browser access

Note that functions of these levels of storage tiers can be combined when advantageous. For instance a local or local network storage device can also work with an external storage service to provide enhanced data protection.

Level 0: No backup

This is the state of no backup of the data on a computer's hard disk drive. All storage is local, there is no archive and file transfers are done via email attachments.

Level 1: Backup to external storage media

Direct attached magnetic tape drives, CD/DVD writers, external USB, Firewire (IEEE 1394) or eSATA disk drives or USB flash memory devices can be used to backup files. Backup to these devices is often a manual or semi-manual process since at least the storage media needs to be changed on a regular basis. External local storage devices are made by Maxtor, Seagate, Iomega, Buffalo, Lacie and many other companies.

Mirroring between two hard disk drives is a way to create an automatic backup copy of the data on a hard disk drive. This is direct attached storage and backup is automatic so if the original drive fails the mirrored drive can immediately take over. This provides an extra level of protection and a failover capability but it is not full continuous data protection and the process does not protect against damage to the computer as a whole. This process might not protect the user from data corruption and other threats.

Level 2: Backup to local network attached storage devices

There are many devices manufactured by various storage vendors that allow automated backup over a local wired or wireless network in a home or small business. Software can be scheduled to backup to the local network storage device (network attached storage or NAS) on a scheduled basis. This form of

backup provides local copies of content, protecting the user from drive crashes and loss of the initial device with the content. They can also allow recovery of the lost data up to the last backup recovery point.

External NAS storage products allow both backup of files on attached computers and other devices that can connect to the network. In addition NAS storage can allow sharing of content to devices attached to the network. This provides both local file sharing and data protection. However there is a need for broader file sharing outside of the immediate local network or protection of the saved content in case of some catastrophic event such as a fire or flood that may destroy all of the storage devices in the local network.

If configured properly a local NAS device on a network that is connected to the Internet can be used to share files outside the local network. Some companies have also offered to pair up backup storage on NAS or external direct attached boxes with external storage services (these represent a mixing of these storage levels).

To protect this networked stored content from local human or naturally caused catastrophes users can make a copy of all the content onto a portable storage device and make sure to take it offsite on a regular basis. This will reduce the potential loss of data due to a local catastrophe but this approach requires a lot of labor and discipline and still doesn't ensure protection of the data if the remote storage location is subject to the same catastrophe as the original location or if the catastrophe happens before the copy can be taken off site.

Examples of companies making local network attached storage devices include Maxtor, Seagate, Iomega, Buffalo, Lacie and many other companies.

Level 3: Remote storage with scheduled backup

In order to effectively protect content from catastrophes or to share the content outside of the local network some remote storage service is required. There are several simple remote storage services available on the Internet that offer simple remote storage with scheduled backups. These backups do not happen often enough to call them continuous data protection (e.g. once a day) so if the user needed to recover backed up data they would only be able to recover up to the latest scheduled backup. As a consequence the more recent files or file versions could be lost.

Companies that offer such simple remote storage and scheduled backup include Dell Online Backup, FilesAnywhere, Spare Backup, Mozy and Xdrive.

Level 4: Remote storage service with continuous data protection

Large companies can set up dedicated networks and storage systems for CDP and instant fail-over but for small businesses and homes an Internet-based storage service provider is more likely to be used to provide CDP.

Continuous Data Protection for small businesses and homes should provide access to user data over the Internet even if the original copy is lost and also recovery of the data onto new storage devices from the Internet. These operations should be convenient enough that the small business or home can quickly access their vital data.

These remote CDP services may or may not include version control and perpetual archiving and don't include file synchronization between computers. Examples of CDP and close to CDP products (either immediate copies of short scheduled updates) for the home or small business market with various frequencies of data retention include Carbonite, Data Deposit Box and Tilana Reserve. Some other products come close to CDP with very frequent scheduled backup (such as Spare Backup) or through continuous file synchronization such as BelnSync and Microsoft's Foldershare.

Level 5: Complete data protection and access

Once data is on a storage service connected to the internet it can allow sharing and even synchronization of files between computers. This is the way that some online storage services can allow the sharing of pictures and other files between friends and family. The data on the remote storage service should encrypt the data and otherwise protect it from unauthorized access.

A remote storage offering that includes multiple computer file synchronization, automated continuous backup to provide CDP with full version control and a perpetual archive and also remote browser access can be called a complete data protection and access service. However this does not include file sharing such as that offered by some sites for photos and other files such as Snapfish, Sharpcast, etc.

Only one company, Tilana Reserve offers a product that meets all the requirements of Level 5 although some other products such as Data Deposit Box and Xdrive have some but not elements of Level 5.

Some Comments on archiving

We should explore archiving a bit more to understand the various ways that this can be done. The archive can be physical storage media that is kept in a secure place on a shelf and available only after mounting back into a drive or storage system. This is a cold archive.

The archive can also be online and available to the user. This is an active archive. If an online backup can keep both new files and older files and older file versions it can act as an online or active archive. In order to be considered an active archive the stored content must be maintained and protected for a long period of time. This is the type of archiving being offered by some online service providers.

An active archive can also be a way to offload files not used on a regular basis on a computer but which the user wants to maintain for an extended period with ready access if needed. This is a sort of ILM (information lifecycle management) for small storage users. An online service can provide continued access to older files that then don't need to be stored on the host computer system. With such offloading on a remote site online backup can become online archiving.

For many homes and small businesses the backup is effectively the archive, in this case an active archive rather than a cold archive. For such users it is beneficial if the maintenance of the stored data is made simple and if data can be retained on the archived storage system after they are removed from the original system should the user choose this option. This allows the user to migrate only the data he wants to a new computer while maintaining either local copies on a network storage device or on a remote storage service or both.

Providers of Internet Storage Services

This section looks at many of the remote storage services that are now available or soon will be available due to the widespread use of the internet. We will focus on services that are generally accessible to homes or small businesses. These will be broken up into categories including file sharing, synchronization of files between computers, online backup and online archiving.

File sharing services

Many companies are offering online file sharing services, particularly for sharing photographs. Companies offering photograph file online sharing include: Flickr, Kodak EasyShare, Snapfish and Yahoo Photos. This list is only a sampling of the many products providing files sharing.

Other companies that offer general online file sharing are Box.net, Microsoft Live Folders, SimDisk S-Drive, Sharpcast, OmniDrive and XDrive. A particular approach to remote file sharing is offered by several services that allow establishing a URL for any remote file (usually a photograph or video) for the purpose of content sharing such as through social networking web sites. Services that provide this sort of selective sharing include Dropboks.com, mediafire.com and localhostr.com.

The offering and price for these services vary and are summarized in the table located in the summary section below.

File synchronization services

Companies that offer synchronization of files between computers are BelnSyn, FilesAnywhere, Foldershare, Mac Backup and Sync, Sharpcast, Streamload Media, Max SyncBack and Tilana Reserve.

The offering and price for these services vary and are summarized in the table located in the summary section below.

Online backup services

Companies offering these online services can be broken down into those catering to large corporate users and those focusing on homes or small businesses. The offering companies are discussed for each market segment below.

Following is a characterization of various markets for online backup:

Following is a set of definitions of various markets based on a blend of the categories used by the industry participants that I interviewed.

- High End/Enterprise: Larger companies with many employees (>500) and often many branch offices
- Medium Size Companies: Mid-size companies with fewer employees (100-500) and no or few branch offices
- SMB/SME Market: Small companies with <100 employees and no or very few branch offices
- HOME OR SMAL HOME OFFICE/Consumer Market: Home based or small businesses as well as consumers, usually with 1-5 employees or consumers

The home and small business market referred to in this report is primarily the home or small home office market as described above as well as the low end of the SMB market.

For these markets, particularly the medium and high end enterprise markets recovery characteristics are critical. Recovery characteristics are often defined as a class of service based on characteristics related to how available the data is, how fast it can be copied, how far back the recovered data should go and the time for data recovery. Several classes of service are shown in **Table 2** based on SNIA classifications. These classes are often referred to as how many 9's are in the data availability for the numbers of decimals of reliability. A contract with a supplier of backup and recovery solutions would refer to the guaranteed class of service as a Service Level Agreement (SLA). The cost of adding each "9" in data recoverability increases roughly exponentially. Most small business and homes can afford at most 3-9's data value classification and many get by on 2-9's.

Table 2. SNIA data value classification (Often used for SLA definitions)

Data Value Class	Data Availability	RPO (Data Loss Risk)	RTO (Max. Recovery Time)	DPW (Copy Data Time)
1 – Not Important to operations	90%	1 week	7 days	Days
2 – Important for productivity	99%	1 day	1 day	12 hours
3 – Business important information	99.9%	2 hours	2 hours	10 minutes
4 – Business vital information	99.99%	10 minutes	15 minutes	None
5 – Mission critical information	99.999%	1 minute	1.5 minutes	None

Source: Derived from SNIA "Implementation Guide for Data Protection," March 2004

Online backup for large companies

All the big name data storage system and software providers provide backup capability. This is one of the fundamental uses for storage systems. Some of these backup approaches still involve magnetic tape but increasingly these are disk drive based solutions.

Companies that provide products for this market include: IBM Tivoli, Symantec Backup Exec, EMC², Network Appliance, Hitachi Data Systems, Fujitsu, Sun and many other smaller companies. These software suites can be used for online backup and disaster recovery mirroring as well as other applications involving large data centers.

There are also companies offering continuous data protection of various types for enterprise environments. Some of these include Asempra, TimeSpring Timedata, as well as all the major storage hardware and software companies mentioned above.

There are several companies that offer online backup services for larger businesses. The major providers of enterprise level online backup services include Asigra, Arsenal, Connected, Livevault (part of Iron Mountain and now called Digital Iron Mountain) and eVault. Combined with many smaller providers such as Amerivault US Data Trust and many regional and special market niche players the enterprise and larger business backup and recovery market was probably over \$6 B in 2006. These companies usually partner with appropriate data center providers to provide online backup services with guaranteed SLA levels using their software for larger businesses.

Online backup for homes and small businesses

Valuable personal content cannot be replaced unless it is backed up and to protect this data from physical damage or destruction to the home or small office it must be backed up to a remote location. The growth of this personal content will be a major incentive to the growth of consumer and home or small office remote backup, as the cost model approaches sustainable levels for a very small company and consumer market.

Some of the larger storage providers also offer products that service home or small business backup such as the EMC² Insignia Retrospect (formerly Dantz) software that is widely used for local backup of computers to direct-attached or network attached storage devices.

Other local storage backup software including BackUp My PC, Double ImageO, Genie-Soft Backup Manager Home, Easy BackUp, Microsoft Backup Edition, Norton Save & Restore, NovaStor NovaBACKUP, ShadowBack, Sonic's Simple BackUp and WinBackup.

Many companies have developed products for the home or small business online backup and recovery market. This is a fledging market with probably less than \$250 M in total revenue in 2007. This revenue is split between a great many small service providers. Although the market is not very large yet it has great growth potential as families and small businesses acquire more and more

valuable content that they wish to protect. Although the barriers to entering this market are currently not very great these types of customers are very price sensitive and generally don't have sophisticated hardware and software support available to them. As a consequence small business and home online backup and recovery must be very cost effective, offer valuable features for the price and be easy to use and maintain. Automation of commonly used functions is a key enabler for this market segment.

Remote online backup and recovery services for homes and small businesses are enabled by the low cost for building online disk drive arrays using ever higher capacity and inexpensive SATA disk drives (currently as high as 1 TB in capacity) and the availability of sufficient bandwidth for most customers to effectively use these online storage pools. Remote storage is used to keep files that can be accessed without carrying them on your computer or even to be able to travel and use computers in remote facilities without having to haul one around while one travels. The files kept or synchronized on online remote storage can also afford a backup of these files although this is generally an incomplete and a rather manual process.

One of the original providers of personal and small business remote file storage services was XDrive (now part of AOL). Such remote data storage can be used for manual backup of files. Recovery then consists of copying the files back to the disk drive in the host system. Other companies providing remote data storage include Google's GMail, Yahoo Briefcase, GoDaddy's Online File folder and Tilana Reserve.

Some companies offering online remote storage also include backup software as part of their offering that allows regularly saving content from home computers or storage devices that are accessible through a home network and recovering it if there develops a problem with the original content. Note that although this software is often sold as the solution for the external storage solution an external storage product generally will not be restricted in what backup software it can use.

Some of the online backup and recovery products (also usually offering remote online file storage) are FilesAnywhere, iBackup, Spare Backup, Streamload and Tilana Reserve. Many of these products offer or will offer software to make the backup and recovery functions relatively automatic.

Carbonite, Mozy and Tilana Reserve offer automatic remote backups. Carbonite and Tilana Reserve offers a true CDP, where files are immediately protected with file change transfers as they occur. Mozy only offers incremental backups on a scheduled interval. Neither Carbonite or Mozy products offers online archiving capability although Tilana Reserve does.

Amazon Simple Storage Service (Amazon S3) is an API that has enabled a new crop of online backup and storage services targeting SMB/Consumer, including JungleDisk, ElephantDrive, MediaSilo, and others. Iomega offers an automatic scheduled backup to remote storage service called “iStorage.”

An interesting and natural development in this market is the inclusion of online remote backup as an option or as part of a home or small business network storage device. This allows keeping local backup copies as well as provides a form of continuous data protection where if all the local copies of content are destroyed the remote copy could be used to reconstruct the original data (at least up to the last snapshot or saved backup version on the remote site). Seagate (originally as the acquired companies Maxtor and Mirra) offered a bundled local storage and remote storage capability by partnering with MyFabrik to provide remote storage and backup to supplement external storage boxes.

Small business and home remote archiving

In the case of Carbonite or Mozy, items deleted from the protected computer are also deleted from the remote space, only retrievable with the help of a support call for a matter of days (usually 30), after which time the items are gone forever. Thus these products don’t offer true file archiving. Tilana Reserve offers unlimited numbers of versions retained on the remote site.

There are some relatively expensive services that allow you backing up and archiving your files remotely, and have full browser access to anything you store there. These include Files Anywhere. This product offers multiple computer file synchronization, version control (retention of multiple versions of files) and scheduled automatic backups to a web accessible remote storage (however the frequency of these backups is not frequent enough to be true CDP). Another product, Data Deposit Box offers similar features except that it doesn’t offer file synchronization.

Summary of online storage service offerings

Table 3 summarizes and compares various remote storage service options currently available in the market. The software chosen for this table all offer automatic backup. This table excludes services that only offer file sharing, remote computer access or remote storage capability. It also excludes services that cater to higher end enterprise server applications. Even with these exclusions it is clear that this is a very busy market with lots of players offering different sorts of products.

In Table 3 synchronization refers to synchronization between two or more computers via a remote site, versioning and archiving refers to long term

retention including saving multiple versions of files, file sharing is sharing of specified files with others over the Internet while protecting access to non-shared content and remote CDP means the data is backed up on the remote service on a frequently scheduled basis (ideally each time a file is changed) and allow recovery of the data back to the original computer or a new computer.

Table 3. Comparison of remote storage services.

Service	Synchro- nization	Versioning/ Archiving	File Sharing	Remote CDP	Pricing	Comments
Backup2Net	No	No	No	No (scheduled)	Free \$179.40	1st 100 MB For 10 GB
Backup Solutions	No	Yes (up to 10 versions)	No	No (scheduled)	~\$179/yr ~\$890/yr	4 GB 30 GB
BelnSync	Yes	No	No	Yes (as synchro- nization)	\$60.00/yr \$100/2 yrs	5 GB
Carbonite	No	No	No	Yes	\$49.95/yr \$89.95/yr	Unlimited
Data Deposit Box	No	Yes	No	Yes	\$2/GB/mo	
eVault Small Business Edition	No	Yes	No	No (scheduled)	~\$1,020/yr	5 GB
FilesAnywhere	Yes	Yes	Yes	No	\$89/yr \$159/yr \$489/yr \$1,650/yr	5 GB 8 GB 25 GB 80 GB
FirstBackup	No	No	No	No (scheduled)	\$48/yr \$144/yr	50 MB 1 GB
Foldershare (Microsoft)	Yes	No	Yes	Yes (as synchro- nization)	Free	
iBackup	No	Yes	No	No (scheduled)	\$99.50/yr \$2,000/yr	5 GB 100 GB
iStorage	No	Yes	No	No (scheduled)	~\$500. 0/yr ~\$1,720/yr	5 GB 15 GB
Mozy	No	No	No	No (scheduled)	\$59/yr	50 GB
PCVault	No	Yes	No	No (scheduled)	\$71.40+/yr	2 GB, price can be higher
SOS Online Backup	No	Yes	No	No (scheduled)	\$93/yr	2 GB
Spare Backup	No	Partial	No	Partial (once every 15 min.)	\$59.88/yr	50 GB
Tilana Reserve	Yes	Yes	No	Yes	TBD	TBD
Xdrive	No	Yes	Yes	No (scheduled)	Free \$9.95/mo	5 GB 50 GB

Future developments for online storage services for homes and small businesses

As homes and small businesses grow in their unique digital content the value of protecting, sharing and using this data grows. In addition to the services discussed earlier in this paper there are many other things that can be done once such content is online. For instance:

- The remote stored content could be viewed by mobile devices such as PDAs or mobile phones (Tilana Reserve has announced a service offering this capability).
- With appropriate protection and limits to access to the content it may be possible to create games and other entertainment that include material from unique personal or family content to give such products a more personal touch.
- The service could be extended to non-computer devices such as portable devices, for instance automatic copying of photographs and videos taken with a mobile phone or to synchronize video and audio files for a automobile entertainment system with web access (probably wireless...).
- If file sharing capability is included then select files of data on say family history could be shared between family members and even allow additions and synchronization of such data between the member's computers.

Conclusions

Personal and small business data is growing at a phenomenal rate. It is expected that by the next decade there will be more personal (i.e. made by individuals) than commercial data being generated and accumulated in the world. For this reason solutions are needed to protect this content as well as make it available to those with a legitimate reason to access it.

A model of personal and small business storage tiering for data protection is presented showing the progression in technological solutions and greater storage security as we move up the levels of the tiers. The highest level offers solutions providing full continuous data protection of the data with easy recovery, perpetual archiving of content including multiple file versions, synchronization of important data between computers, remote browser access and ideally capabilities to share content with others as appropriate.

There are many offerings by companies to provide software and hardware for local and remote backup, archiving and recovery of data or file synchronization. These products have been examined and a summary table made of remote services playing to this market that offer synchronization, backup and recovery (esp. CDP) and archiving capability. There are many players in this market and their offerings and prices vary considerably. Because of the pricing pressure in these markets and the need to make products for individual consumers that are straightforward and easy to use it is difficult to make a product that serves all these needs.

Tilana Reserve offers a product that represents clear innovations in putting together many of the needed elements for personal and small business data protection and synchronization of files. We expect to see continued advances in this market to serve the growing body of digital content created by individuals, families and small businesses and to make such content available for permission-based advanced features and applications.

Appendix A. Glossary

Classical backup and recovery have evolved considerably as the demands for data protection have changed. Data protection now recognizes the requirements for various levels of data availability as well as the ability to recover from human, device or natural disasters. Various terms are used to describe these various niches for the comprehensive data protection market. Following are a set of definitions that we use in discussing general developments in data protection and in particular in remote backup and recovery.

- **Recovery Point Objective (RPO)** is the point in time to which you wish to recover data (this may be seconds, minutes, hours, days and even weeks). This determines how often you want to backup and how far back you want to retain the backups for fast restore.
- **Recovery Time Objective (RTO)** is the time in which you want to have the data recovered. Unless there is a local copy of the data to be recovered (such as with CDP)
- **Single Instance Storage** refers to backing up changes where one knows where changes in a file take place but having to back up significantly more than was changed in order to recover the changed file. The term refers to a technology that was common in Microsoft Windows 2000
- **Commonality Factoring** or **deduplication** refers to backing up only the changes and information on where these go rather than unchanged parts of the file. Commonality factoring significantly reduces the cost of backup by reducing bandwidth requirements
- **Service Level Agreement (SLA)** defines the level of backup or recovery performance expected from a service such as up-time, RPO and RTO.
- In **Continuous Data Protection (CDP)** you desire to recover the backed up data almost immediately (low RTO). This usually requires a backup device very close to the device being backup that the backed up device can fail over to. CDP is a method to implement business continuity.
- **Remote backup and recovery** involves sending data offsite to a distant location in order to ensure that the data can be recovered over the long distance connection. Remote backup is a method to enable disaster recovery.
- **Archiving** differs from backup and CDP in that it usually refers to the long term retention of data, not necessarily immediately available.

Appendix A. Service provider company information

About the author



Thomas M. Coughlin
408-871-8808 (phone)
www.tomcoughlin.com

Tom Coughlin, President, Coughlin Associates has been working for over 30 years in the data storage industry. He has over 60 publications and six patents to his credit. Tom is active with IDEMA, the IEEE Magnetics Society, IEEE CE Society, and other professional organizations. Tom is Chairman of the 2007 Santa Clara Valley IEEE Section. He is the founder and organizer of the Annual Storage Visions Conference, a partner to the annual Consumer Electronics Show. Coughlin Associates provides market and technology analysis as well as Data Storage Technical Consulting services. Publications include an annual report on data storage for consumer electronics and another on digital

storage for entertainment creation and distribution. He also publishes a quarterly newsletter (**The Data Storage Technology Newsletter**). For more information go to www.tomcoughlin.com