Understanding the Benefits of Online Backup and Data Synchronization

An Osterman Research White Paper

Published September 2011

Sponsored by

BACKBLAZE
Executive Summary

Organizations store enormous amounts of data. For example, a company with even a relatively small number of users can generate many terabytes of content each year in the form of word processing documents, spreadsheets, presentations, videos, emails, images, scanned documents and many other content types. Similarly, individual users can generate many gigabytes of content each year that they must preserve for corporate and/or personal requirements. This content must be backed up for a variety of reasons, including disaster recovery, business continuity, legal and regulatory compliance, theft, and just good data protection practice.

However, there are several factors that complicate the problem of backing up data, including an increasingly dispersed workforce as a result of more employees working from home, more devices on which users can generate and store content, and more stringent corporate governance requirements that require organizations to preserve content for long periods of time. Add to this the variety of problems that could potentially make data temporarily inaccessible or destroy it permanently, including natural disasters, power outages, lost devices, hardware failures and user errors.

Moreover, many organizations and users do not back up their data on a regular basis because they don’t have the time, backup processes are too time consuming, or they simply forget to do so.

KEY TAKEAWAYS

There are four primary takeaways discussed in this white paper:

- Business records and other content, as well as personal information, represent critical information that must be preserved, sometimes for very long periods or indefinitely in some cases.

- A failure to properly back up data and make it readily available to users can have disastrous consequences in any organization.

- Traditional backup capabilities are lacking in several key respects, not least of which is the fact that many organizations do not keep them in geographically remote locations to protect against natural and other disasters.

- Cloud-based backup and synchronization services can directly address the problems that organizations and individuals face in the context of their data management requirements and should be seriously considered as a way to protect data and to make it easily available.

ABOUT THIS WHITE PAPER

This white paper discusses the critical importance of protecting corporate data, as well as the benefits of cloud-based backup and synchronization services. It also discusses the relevant offerings from Backblaze, the sponsor of this white paper.
Users Generate and Store Enormous Amounts of Data

Users create large amounts of content in the form of emails, word processing documents, spreadsheets, presentations, video files, image files, audio files, instant messaging conversations, notes and all sorts of other content – and they store this content for long periods of time. For example, let’s assume that the typical user conservatively produces the following:

• 30 emails per day (a total of 350 kilobytes)
• Three word processing documents per day (a total of 300 kilobytes)
• One spreadsheet per day (50 kilobytes)
• Two presentations per week (500 kilobytes)
• One video file every two weeks (10 megabytes)

That means that during a year’s time (250 workdays), the typical user would very conservatively generate 320 megabytes of content. However, it’s important to note that this content is typically shared among many users and sent to others inside and outside the organization, all of whom will store their own copies of this data. The result is that the typical user generally stores 25 to 50 gigabytes of data. With data stores growing at roughly 40% annually, this means that users are adding roughly 10 to 20 gigabytes of storage every year. Some of this content is produced by users themselves, but most is content that they receive and store for future use.

DATA STORES CONTINUE TO GET LARGER OVER TIME

Not only do users and corporate systems generate large amounts of data, but also this content must be stored for long periods of time for a variety of reasons:

• Users have a variety of individual requirements to store the content they generate, either because of corporate policies to store certain types of information or simply for personal reasons so that they can refer to older content or have a record of what they have done in their job.

• More importantly, virtually every organization has an obligation to preserve its business records for purposes of e-discovery, regulatory obligations or industry best practice. For example, in the United States:

  o Accountants of publicly held corporations must retain certain records and workpapers relevant to the audit or review of such corporations’ financial statements for seven years.

  o Employers must retain documents related to hiring, promotion, demotion, transfer, layoff or termination for one year from the date of record production or the personnel action involved, whichever occurs later. If a charge of discrimination has been filed, or a civil action brought against an employer, the employer must retain all personnel records relevant to the charge or action until the final disposition of the case.

  o Certain types of financial services firms must maintain a record of each order to purchase or sell securities for a period of seven years from the date the order record was created; for the first two years, these records must be kept in an “easily accessible”
Understanding the Benefits of Online Backup and Data Synchronization

What this means is that records must be stored for significant lengths of time, resulting in data stores that continue to grow in size over time. In an era of increased requirements for corporate governance and greater government oversight, we expect data stores to continue to grow at an even faster rate in the future as more data is produced and less data is deleted.

THE NUMBER OF DEVICES CONTINUES TO GROW
Further complicating the data storage issue is the increasing number of devices that individuals use to generate and store content: desktop computers, laptop/notebook/netbook computers, smartphones, tablet computers, USB sticks, etc. This not only increases the difficulties associated with storing data using conventional approaches, but it also tends to increase the overall quantity of content that users and organizations store. Further, users must be able to access current versions of data across all of these devices.

ORGANIZATIONS MUST STORE DATA FROM MANY SOURCES
Similarly, organizations must store all of the content on file servers, backup tapes, disk-based backup systems, archiving systems, SharePoint repositories, application databases and other data stores. Moreover, corporate systems that generate purchase orders, documents related to supply chain management, invoices, logs and other files account for a large share of the content that organizations must retain.

This has led to a situation in which data is produced by multiple systems and is stored in a large number of locations, making data management and backup that much more difficult. For example, the following figure from an Osterman Research survey conducted earlier in 2011 shows that the typical organization stores data in a diverse group of locations.
A GROWING PROPORTION OF USERS ARE WORKING REMOTELY

An important trend that is having significant implications on data storage, backup and content synchronization is the growing proportion of users who work remotely. The availability and deployment of unified communications systems and cloud-based communication technologies make it easier for organizations to let users work from home or from other remote locations on a regular basis. Employees are motivated to do so because of reduced commute time, lower vehicle operating costs and overall quality-of-life issues. Employers are motivated to permit telework for two important reasons:

- They can significantly reduce their real estate costs and property taxes by operating smaller offices. For example, a 250-employee company in which staff members average two days per week working from home needs to provide workspace for only 150 employees. If we assume a real estate cost of $20 per square foot per year and 150 square feet required per employee, reducing the office space for 100 employees can result in annual real estate savings of $300,000 annually in addition to reduced power, cooling and other costs.

- Moreover, productivity can be substantially higher when employees work from home. For example, if just 40% of an employee’s commute time was devoted to work, an employee with a fully burdened annual salary of $80,000 would generate more than $1,500 in additional productivity each year for his or her employer.
Despite the advantages to both employees and employers of working remotely, there are a number of additional problems created from telework, including the sheer number of venues at which data is stored, the increased difficulty of protecting sensitive and confidential data, the difficulties associated with reliably backing up this data, and synchronizing data across multiple systems and locations to avoid version control problems.

**What Could Happen to Your Data?**

Not only are there problems with managing data generated by multiple systems, stored in many locations and accessed by geographically distributed employees, but there are a number of major and minor problems that can negatively impact organizations and the individual users within them – and the data to which they must have access:

- **Natural disasters**
  Natural disasters impact just about every location on earth at one time or another. In the United States, for example, there is an annual average of 1.8 hurricanes, 818 tornadoes (although there were 1,458 tornadoes in 2011 as of June 8th), 75,000 people driven from their homes by floods, and 63.8 earthquakes of magnitude 5.0 or greater. Fires, although most not a disaster from “natural” causes, are another major issue for businesses and other non-residential organizations – in 2009, for example, there were 16,500 fires in stores and offices and 5,500 fires in institutional properties in the United States, not including the large number of residential fires.

- **Power outages**
  Major power outages in the United States are on the increase. For example, the North American Electric Reliability Council (NERC) Disturbance Analysis Working Group has found that there were 41 power outages impacting 50,000 or more customers during the period 1991-1995, 58 such outages during 1996-2000 and 92 such outages during 2001-2005. While power outages normally impact data access and backups temporarily, these outages can damage computer equipment and can destroy data.

- **Lost devices**
  The proliferation of smartphones, laptop computers, tablets and other mobile devices that contain corporate and other data has created another growing problem, namely lost devices. For example, a 2008 survey by Credant Technologies found that more than 3,000 laptops and 55,000 mobile phones were left in London taxis during just a six-month period. Even as far back as 2006, a Pointsec Mobile Technologies survey found that during a six month period users left 8,701 devices in Washington, DC and Baltimore-area taxis, while 3,106 devices were left in taxis in the San Francisco Bay Area. The problem is getting dramatically worse as the number of mobile devices proliferate.

- **Other problems**
  Although all of these problems can and do create data loss, there are much more common causes of data loss. For example, a Google study discovered there is a 36% chance of hard drive failure within a five-year period. An analysis of Apple’s Time Capsule failure rate found 7.3% of these high-capacity devices failed after approximately 19 months of use. Even solid-state drives have a reasonably high failure rate – as high as 2.39% according to
One source estimates that 4mm DAT tapes have an annual failure rate of 11.1%.

Add to this the occasional leaky sprinkler pipe above a user’s PC or a server room, the cup of coffee spilled on a laptop keyboard, or the USB stick that gets left behind at a client’s office never to be seen again.

WHAT HAPPENS WHEN DATA IS LOST?
When data is lost there are three primary consequences that can occur:

• The least harmful consequence is that users must rekey in lost data (assuming they are even able to do so), resulting in major productivity losses as users create their work a second time. This can lead to major delays in projects, missed deadlines, missed sales opportunities, unhappy clients and other serious problems.

• Lost data can also have more serious consequences, including the inability to recover lost intellectual property, an inability to recreate data that has been generated by past employees, missing email and records of other communications, lost business and tax records, etc. This can create serious problems on a number of levels, with ramifications that may greatly exceed the productivity losses that can result from lost data.

• While some statistics, such as those indicating that the majority of businesses that lose data go out of business a short time later, appear to be more hype than reality, businesses that lose data for any reason are at a severe disadvantage relative to those that do not lose data.

WHAT DOES ALL OF THIS MEAN?
Summarizing the preceding few pages reveals that there are four fundamental issues that any decision maker should take into account when formulating a data management plan:

• Users and organizations are creating and storing large and growing quantities of data.

• This data must be stored for longer periods and for a variety of reasons, including those driven by legal considerations, regulatory requirements, corporate best practice and individual user needs.

• Data is stored in a growing number of locations, including the large number of traditional data stores within the workplace, as well as mobile devices, portable storage devices, employees’ home computers and other venues.

• Loss of data carries with it enormous consequences that, in some cases, could put an organization out of business.
Best Practices for Backup and Data Synchronization

There are seven key practices that any organization – regardless of its size or the industry in which it participates – should follow. These are also good advice for individual users who need to protect personal records, images, videos and other content:

1. **Protect data**
   It is absolutely critical that all data be protected by making backups of this data on a real-time or near real-time basis. *Good* practice dictates that all data is backed up to tape or disk, but *best* practice requires that all data is backed up to a geographically distant location, such as a secondary location or the cloud. Doing so is critical to protect against a variety of eventualities, such as a hurricane, tornado, flood, fire or other catastrophe that heavily damages or destroys a corporate location, as well as the data it houses. Backing up to a distant location is also critical in case a corporate location becomes inaccessible or unusable due to a snowstorm, police action, terrorist attack, or a utility failure (such as a broken gas main) that causes an evacuation of a particular area; or in the event of a power outage.

2. **Encrypt data**
   It is essential that data is encrypted behind the corporate firewall and/or on the client before being sent over the Internet. This will ensure that data is protected both in transit and at rest. It also ensures that in the unlikely event of a data breach at the remote location, user data cannot be accessed.

3. **Make backups automatic and continuous**
   Backups should be automatic so that users or administrators are not required to intervene in the backup process. This will ensure that backups are not dependent on individuals remembering to backup content stores. As a corollary to this point, backups should be continuous and as close to real time as possible so that newly created data is not lost in the event of a power failure or other problem.

4. **Make data easily accessible in case of disaster**
   Not only is it critical to back up data, preferably to a distant location, but it is just as critical to make this data available easily and quickly for purposes of server restoration, giving users the ability to access their individual data when working at a new location, or bringing a new facility online rapidly. For example, if users need to move to new, temporary offices after an earthquake, it is imperative that users in the new location or those working from home have access to all of their data as quickly as possible.

5. **Make data easily accessible to users in near real time**
   Another important part of any data protection and management strategy is the ability to make all data easily accessible to users in as close to real time as possible. For example, data backup and synchronization capabilities are very important for users who collaborate in geographically separate locations so that they can have access to the latest versions of documents, spreadsheets and other content. An inability to provide rapid access to the most current information can lead to employee downtime while waiting for content to be sent through email, version control problems and other impediments to efficient work.
6. **Make backup and sync as simple and painless as possible**
   Finally, but certainly not least important, is making backup and synchronization as painless as possible for users and IT staff alike. Because these capabilities can be a major source of difficulty and a significant consumer of time for IT administrators, they should be as painless as possible to deploy, configure and manage.

7. **Storage should be unlimited or inexpensive**
   Because data storage requirements will likely be greater than most individuals or organizations anticipate, it is important to use a provider that will provide either unlimited storage or storage that will be and remain inexpensive as data stores increase in size.

### Why Consider Cloud-Based Backup and Sync?

There is a growing number of cloud-based backup and/or synchronization (backup/sync) services that will automatically copy content to remote storage systems and make it available from any Internet-connected computer with the appropriate access credentials. There are several reasons that any organization should consider using such a service:

- **Data will be protected**
  A cloud-based backup/sync service, by automatically and continuously migrating content to remote storage systems, will provide the inherent protection of data that organizations require. Further, because the data is moved automatically to a remote location (and further backed up from there), the best practice of keeping corporate data in at least two geographically separate locations can be achieved. This means that even if one or more locations that house corporate data becomes inaccessible due to a catastrophe or a less serious incident, all data is protected and can be accessed from any other location.

Cloud-based backup/sync services are superior to traditional practices that back up data to tape or other media and then store these backups in a remote location because of two inherent limitations of this practice:

- These backups, even if performed nightly, will not capture data stored between backup windows. For example, if the nightly backup occurs between 2:00am and 6:00am each morning and a disaster occurs at 3:00pm one day, up to nine hours of data will be lost and is potentially not recoverable.

- Moreover, shipping or driving tapes and other media to a remote location is expensive and time-consuming, and can expose this content to loss from theft or negligence.
• **Corporate governance obligations will be easier to satisfy**
  Cloud-based backup/sync services make corporate governance obligations easier to satisfy by automatically making backups of critical business records that must be retained for legal, regulatory or other reasons. By migrating these records to cloud-based storage and establishing the appropriate retention periods for them, organizations can be assured that they will have preserved critical content for the appropriate length of time. While a backup of this content is not an archive in the sense that the information has not been indexed and therefore made easily searchable, at least the retention of this content – a key requirement – will have been accomplished.

• **Employee productivity will be improved**
  Another important benefit of a backup/sync service is that employees are given ready access to content regardless of their location. For example, an employee who creates a document at work and then wants to continue working on the document after he or she arrives at home can easily access that document. The updated document will then be available when the employee is at work the next day.

  Other employee-focused benefits of a backup/sync service include employees not having to recreate lost content, not having to ask others for copies of files, or experiencing version control problems that can lead to lost productivity and other problems.

• **Cloud-based storage can be less expensive**
  Cloud-based backup/sync services can also be less expensive than their on-premises counterparts when all of the costs of the latter are included. These include the cost of the tape drives, servers or other hardware and software elements necessary to perform the backups; the media on which the backups are stored; the transportation of media to a remote location for safekeeping; the IT labor required to manage and troubleshoot backup systems; and the opportunity cost of having internal IT staff manage what is, for all intents and purposes, a relatively mundane aspect of data management.

• **There are other, intangible benefits**
  Cloud-based backup/sync services can also provide a number of benefits that may be more difficult to identify or quantify, but are nonetheless important to consider. For example, losing data can create a perception of corporate mismanagement – customers and business partners may fear that their data may be at risk and may have somehow become compromised as a result. Employees may become angry when they realize that they might have to recreate days’ worth of content in advance of a meeting or presentation. Even if IT regularly backs up content to tape or disk, finding and restoring this content – if IT even has the bandwidth to perform these tasks – can take several hours.

**Questions You May Want to Ask**

Osterman Research recommends that decision makers consider asking at least some of the following questions when evaluating vendors of online backup and/or synchronization services:
QUESTIONS TO ASK OF YOUR SENIOR MANAGEMENT

• Are you comfortable with the notion of having corporate data stored in the cloud? You should be, so if not, why not?

• How vulnerable is our organization to data loss?

• If you (senior management) are not comfortable with having corporate data stored in the cloud, what is our disaster recovery and business continuity plan for recovering data assets after a natural disaster, power outage, fire, flood, etc.?

• Have you conducted a direct and opportunity cost analysis of providing disaster recovery and business continuity capabilities using on-premises/internally managed systems?

• Have you evaluated the costs of online backup and/or data synchronization services relative to on-premises/internally managed systems?

QUESTIONS TO ASK OF A PROSPECTIVE PROVIDER: COMPANY BASICS

• Is your company financially viable?

• How long have you been in business?

• How many customers do you support and how has this changed over the past six months? The past year?

• What size and type of customers do you support?

• Can you provide referenceable customers that are similar to our organization?

• What corporate certifications or audits do you offer?

• What happens to our data if we want to switch to another provider?

QUESTIONS TO ASK OF A PROSPECTIVE PROVIDER: WHAT’S INCLUDED?

• What is the minimum number of users that you require for an account?

• How much data can be stored per user? In aggregate?

• Do you provide unlimited storage for a fixed price?

• If there is a maximum amount of storage for our company, can this storage be allocated differentially across your users, or is it a fixed amount per user?

• How much does storage cost per gigabyte beyond the basic amount included in each account?
QUESTIONS TO ASK OF A PROSPECTIVE PROVIDER: INFRASTRUCTURE
• How many data centers do you operate?
• Is data replicated between data centers for redundancy?
• What are the specs and certifications for the data center(s)?
• Is the data center SAS 70 Type II certified? If so, for how many years?
• Are you using your own technology or another vendor’s?
• How scalable is your infrastructure?
• Do your data centers have backup generators, redundant telecommunication links, etc.?

QUESTIONS TO ASK OF A PROSPECTIVE PROVIDER: RELIABILITY
• What Service Level Agreements do you offer?
• How much downtime has your system experienced during the past month? Six months?

QUESTIONS TO ASK OF A PROSPECTIVE PROVIDER: SECURITY
• How physically secure is your data center?
• What intrusion detection systems are in place?
• Do you encrypt all data at the client computer? If so, do you provide a private key option?

QUESTIONS TO ASK OF A PROSPECTIVE PROVIDER: SUPPORT
• Do you provide 24 x 7 technical support?
• If 24 x 7 support is provided, is it live support during non-business hours or do you have to submit and email and receive a response the next business day?
• Do you provide different means of communication including phone, email and online chat?
• Is only one IT administrator able to contact support with questions or can any employee with a question or issue contact a support representative?

It is important to note that not every question will be important to every potential customer of an online backup or data synchronization service, nor will an affirmative answer to every question discussed above be necessary. This is particularly true when online backup services, for example, are used only for disaster recovery purposes – they may be called upon only rarely, and so things like rapid response to technical support questions may be unnecessary.
Summary

Users and organizations are creating and storing growing quantities of data that must be preserved for long periods, and that must be available in the event of disaster or another problem that could make primary content stores unavailable. As a result, every organization should implement a data management plan to ensure that critical business data is available continuously from any location. Cloud-based backup-sync services will make automatic and continuous copies of data to remote storage where it will be protected and made available on demand from any location without intervention from individual users or IT staff.

Sponsor of This White Paper

Backblaze online backup automatically backs up all laptop and desktop data. The service provides unlimited storage for just $3.96/computer/month.

Key features:

- Automatic, continuous, and incremental online backup
- Unlimited storage and bandwidth
- Fully encrypted
- Backs up external drives
- Restore files via the web or FedEx DVD or hard drive
- Computer location service for lost/stolen systems
- Windows XP, Vista, 7 and Mac OS X 10.5+

Built upon our highly efficient cloud storage, the Backblaze online backup service enables easy-to-use, rapid, cost-efficient backup of all corporate laptops and desktops. Backblaze was selected as a winner of the AlwaysOn Global 250 private companies for game-changing technology and market value. CNET said, “Backblaze takes the guess work out of backing up your data,” and The New York Times said “Backblaze’s backup application is one of the easiest to use.”

Get started with a free online backup trial today or contact us to learn more about business online backup.
Understanding the Benefits of Online Backup and Data Synchronization

© 2011 Osterman Research, Inc. All rights reserved.

No part of this document may be reproduced in any form by any means, nor may it be distributed without the permission of Osterman Research, Inc., nor may it be resold or distributed by any entity other than Osterman Research, Inc., without prior written authorization of Osterman Research, Inc.

Osterman Research, Inc. does not provide legal advice. Nothing in this document constitutes legal advice, nor shall this document or any software product or other offering referenced herein serve as a substitute for the reader’s compliance with any laws (including but not limited to any act, statute, regulation, rule, directive, administrative order, executive order, etc. (collectively, “Laws”)) referenced in this document. If necessary, the reader should consult with competent legal counsel regarding any Laws referenced herein. Osterman Research, Inc. makes no representation or warranty regarding the completeness or accuracy of the information contained in this document.

THIS DOCUMENT IS PROVIDED "AS IS" WITHOUT WARRANTY OF ANY KIND. ALL EXPRESS OR IMPLIED REPRESENTATIONS, CONDITIONS AND WARRANTIES, INCLUDING ANY IMPLIED WARRANTY OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE, ARE DISCLAIMED, EXCEPT TO THE EXTENT THAT SUCH DISCLAIMERS ARE DETERMINED TO BE ILLEGAL.

http://www.erh.noaa.gov/cae/svwx/tornado bystande.htm
http://central.tli.umn.edu/The%20Grid_Amin_Stringer.pdf
Failure Trends in a Large Disk Drive Population, Google, Inc., February 2007
http://www.guardian.co.uk/technology/blog/2010/jul/16/apple-time-capsule-failure-calculated
http://www.circadianforce.com/electronicbackup.htm
http://www.continuitycentral.com/feature0660.html