RED WIRE SERVICES

PO Box 46454 // Seattle, WA 98146 // T 206.829.8621 // F 206.826.1223 // info@redwireservices.com

Online Backup: A Viable Offsite Data Backup Strategy for Your Business?

October 2011

SHELTER YOUR DATA

A. Executive Summary

Most organizations recognize the need for offsite data backup, but with so many vendors and solutions in this market, choosing the right system for your business can be overwhelming. To complicate matters, the majority of salespeople at data backup and recovery vendors do everything they can to make their offerings seem to fit your business, even when other data backup options may be more efficient and cost effective.

For small to medium business owners, executives, and IT managers, the search for offsite data backup systems is an overwhelming process with many new and exciting solutions — and as many vendors. Specifically, "cloud computing" has taken the industry by storm, with everyone rushing to declare their backup offering as "cloud-based," much more reliable, and of a lower TCO (total cost of ownership) than tape or other removable media.

With the tsunami of products and proponents of cloud-based backup solutions, it would seem reckless to choose a tape system for offsite storage needs in 2011 and beyond. However, finding the ideal medium and process to fulfill your offsite storage requirements is not as simple as choosing the best online backup provider. Tape is still a viable medium for offsite storage of backup data, and in many cases, is more reliable, less expensive, and easier to manage.

Cloud-based backups are novel; we just need to be careful what problems we address with them. For example, cloud backups are an excellent choice for smaller businesses and sites with a relatively small set of data. The "pay as you go" pricing model often associated with cloud backup is attractive to these small businesses, and in many cases, the service may be purchased for a fraction of the cost of a tape-based system for small offices, even when projected out over the life of the system (typically three to five years). Cloud backups are also useful for protecting critical data in larger organizations where immediate access to backup copies is a requirement. However, shipping tapes from an offsite location is still much faster when recovering a large amount of data.

Aside from these special situations, our examples in this paper show that tape or other removable media is a more efficient and cost effective solution for all but the smallest environments. For example, our mid-sized higher education institution would spend approximately $16,000 per month to utilize online backup for all of their data when a tape-based system would cost just $333 per month over five years, including physical offsite storage and transport of the tapes.

Table of Contents

B. Introduction

This white paper will explore the newer cloud-based offsite backup methodologies compared to traditional tape and other removable media, where each has its strengths, and where each should be avoided. We go beyond the hype to help you determine which solution is best for your company.

This paper will give you a peek behind the curtain, cut through the propaganda, and point you in the right direction toward your best offsite backup option. This paper will:

- Examine a few of the widely-held assumptions regarding backup using tape and other removable media;
- Compare backup using tape and other removable media to the newer methods of offsite data protection using the Internet (online or cloud backup); and
- Look at five common Small and Medium Business (SMB) customer environments: a small software development firm, a small photography studio, a small manufacturing business, a medium-sized service company, and a medium-sized higher education organization.

With this information, you will be able to formulate a reasonable plan of action and make decisions for any SMB in need of offsite data protection.

Forty-three percent of businesses that close following a natural disaster never reopen, and an additional 29% of those close down permanently within two years.
— Gulf Coast Back to Business Act of 2007

Seventy percent of small businesses in the U.S. experienced a data loss in 2009 due to technical or human disasters, resulting in an average loss of $4,700.
— AMI 2009 U.S. Small Business Annual Overview Study

Forty percent of all companies that experience a major disaster will go out of business if they cannot gain access to their data within 24 hours.
— Gartner

C. Why Offsite Data Protection?

While many small and medium-sized organizations struggle to keep their data protected onsite, it is equally important to keep at least the most critical data at an offsite location. Offsite data protection allows an organization to recover from some pedestrian as well as extreme situations, such as an extended power or network outage, malicious employees, computer viruses, cyber attacks, civil unrest, fire, flood, theft, vandalism, earthquakes, tsunamis, and other natural disasters.

For some of these situations, you may be able to regain access to some or even all your data over time, but losing data availability may be too great for your organization to recover as competitors clamor to take advantage of your misfortune. Offsite data protection is one piece of a disaster recovery strategy that must be in place to ensure your organization can predictably and reliably recover from site-wide disasters.

The statistics on organizational survival from data loss without a data protection strategy are astounding (see sidebar). Here is what W. Curtis Preston, aka "Mr. Backup," says about the need for data protection in his book, Backup & Recovery:

> *We live in the information age. Even the smallest "mom-and-pop" businesses rely on some type of computer systems to store the information they generate. Perhaps it's a list of customer phone numbers, a log of business transactions, or even details about a new product. Whatever the information is, the business would be damaged if it was lost, deleted, destroyed, misplaced, or stolen. Therefore, a complete data protection system protects against all these risks.*

While it is prudent for almost all businesses to store backup data offsite, many businesses are bound by regulations to safeguard their data, including financial institutions, brokers and dealers, and medical institutions[1]. If your business must access data after a disaster, you must have offsite backup.

[1] http://www.gartner.com/DisplayDocument?doc_cd=128123

D. Key Technologies

Before we dive into the methodology that works best in each scenario, we need to discuss and define the two main technologies involved in offsite data protection today: Removable Media (tapes or hard disk drives) and Online Media ("cloud" or multi-site private networks).

1. Removable Media

Removable media is any type of storage device that can be removed easily from computer systems. Modern examples of removable media include USB memory sticks or "thumb drives," writable CDs or DVDs, external portable hard drives, RDX cartridges, and tapes. Other forms of removable media are still in use, but these are the most common.

Removable media is the simplest form of offsite backup or data protection. The person responsible for backups instructs the backup system to write a copy of their data to the removable media on a regular basis, and when complete, the media is sent offsite via a parcel carrier or data service courier, taken to the bank safe deposit box, or even brought home by an employee.

Advantages of Removable Media	*Disadvantages of Removable Media*
Simple	Employees must diligently follow a strict schedule of moving physical media, which often becomes a low priority against more pressing tasks.
In general, backups are written quickly and can be restored quickly, once media is retrieved.	Extra moving parts and shipping of media can lead to an increased chance of media failure and ultimately failed recoveries, especially with commodity hard disk drives.
Relatively inexpensive	Restoring data only available offsite takes a long time (must retrieve media first).

Let's take a closer look at the most commonly used removable media for offsite backups: tapes and hard disk drives. These mediums are more common because they can hold a significant amount of data, can be written to quickly, and are relatively inexpensive. CDs, DVDs, and USB thumb drives store only small amounts of data and are impractical and hard to manage for all but the smallest data sets.

a. Tape Backup

Tape backup is the traditional backup methodology for both primary and offsite backup, but it is quickly losing ground to disk-based backup. Because of the proliferation of cloud-based solutions, many no longer consider tape to be a viable backup option, but it is ideal for many environments.

Tape does have its problems: sequential-only access, time spent managing potentially hundreds of physical tapes, the complexity of tape libraries (to more easily manage a large number of tapes), and a relatively high cost of drives and related equipment.

Tape, however, has its place and shines for cost per GB and longevity of media, as data can be archived and restored decades later, and the speed is *greater* than commodity disk drives.

Many believe that tape is an inferior and unreliable backup medium, some advocating that 70 percent or more of tape recoveries fail. However, we have never seen a tape recovery fail in a system that was properly maintained and managed. Furthermore, hard drives are actually ten or more times *less* reliable (based on bit error rates).

While tape backup has been around for over 60 years, today's tape is far superior to previous generations. The first LTO generation, LTO-1, held just 100GB of data and had a maximum throughput of 20MB/sec. Today's latest LTO-5 generation operates at 140MB/sec and holds 1.5TB uncompressed, a speed unsurpassed by all but the fastest — and most expensive — enterprise hard disk drives.

Another important factor is that tape is built for backups and archives. It's not designed to be randomly accessed or with a focus on high speed, as hard drives are, but rather longevity and reliability. LTO tapes have a bit error rate of one in about 10^{16} bits, an order of magnitude higher than enterprise disk (FC/SAS), and two higher than commodity disk drives (SATA). Tape backup drives also read just after they write to prove the data was written correctly, and the magnetic medium on which tape data is stored is more stable than disk[2].

Tapes are actually "greener" than hard disk-based solutions as well. While idle disk drives consume nearly as much power as when active, idle, unmounted tapes use no power at all. Also, while you may have hundreds of tapes, you are likely to have just a few tape drives, each of which uses about two to three times the power of a hard drive in use.

Therefore, if you have three tape drives operating constantly, you may consume about nine hard drives worth of power. However, you can store as much data as hundreds of hard drives. This results in dramatic power savings for data that can be moved to tape.

b. Commodity Hard Disk-Based Backup

While tape may be faster in many scenarios, hard drives have an edge when random access to data is needed. When restoring a small subset of data, hard disks are quick and efficient at finding the specific data you need and restoring it quickly.

Increasingly, sites use hard drives where tapes were used in the past, shipping the disks offsite to protect against onsite data loss. Offsite disk has similar characteristics as offsite tape, but with additional drawbacks. Hard drives were never designed for being shipped with data or stored for years in idle mode.

Hard drives at most have a five year warranty, and in general, five years is the maximum useful life of drives. When using hard drives for removable storage, keep in mind that they must be replaced at least

---

[2] http://www.backupcentral.com/mr-backup-blog-mainmenu-47/13-mr-backup-blog/380-tape-more-reliable-than-disk-for-long-term-storage.html

every five years, and ideally all archives should be kept on two removable drives in case one should fail. While tape tends to be more reliable in the long term than hard disks, two copies of all tape archives are recommended (one onsite and one offsite).

Even with these drawbacks, hard drives can be a reasonable means for offsite backup for smaller sites. While hard disks were never designed for offsite backup, they work well for moderate retention periods. With 3TB disk drives available today, more companies can fit their important data on one drive, keeping their offsite media simple and affordable.

c. RDX Media

While not that new to the market RDX cartridges are still considered a "new" technology. RDX cartridges are essentially specialized laptop hard drives with additional features that make them more competitive with tape media.

Unlike commodity disk drives, RDX is designed for backup purposes. RDX drives can survive much more physical abuse and drops than standard disk drives, have up to a 30-year shelf life, and maintain the random access benefits of hard drives.

While the characteristics of RDX hard drive media are different than tape, with a maximum transfer rate of 45MB/sec and random access benefits, we will categorize both tape and RDX as "Tape/Removable Media Backup," as they are within the same order of magnitude of cost and performance. RDX media is approximately three times the cost of tape, but single drives ("docks") are much less than single tape drives. When utilizing a library of RDX cartridges, the cost is on par with a tape library, making RDX's advantage for sites needing only one drive, with a maximum total data usage of 1TB (size of one full backup).

RDX media is significantly slower than tape (45MB/sec versus 140MB/sec for LTO-5) and has the same error rate as commodity hard drives (an uncorrected error occurs every 10^{14} bits)[3]. It is still an excellent choice for small environments that need more robust media than commodity drives and will last decades on the shelf.

[3] http://www.quantum.com/ServiceandSupport/SoftwareandDocumentationDownloads/RDX/Index.aspx

2. Online Media ("Cloud" or Multi-Site Private Networks)

For the purposes of this white paper, the term "online media" is any data storage mechanism that is available over a network to backup systems or servers. This storage may be available via the Internet ("cloud storage" is a current buzzword for this) or at second site available via a private network within a multi-location company.

Online media works much like removable media, except that the physical handling of the media is not required. The party responsible for offsite data protection instructs the backup system to write to the online media on a regularly scheduled basis, and by virtue of being in a physically different building, city, or even country, the online backup fulfills the need for offsite data protection.

Advantages of Online Media	*Disadvantages of Online Media*
Offsite protection is automatic; no one needs to remember to ship media	More complicated to implement
Small restores from offsite are quick and easy; no waiting on media to arrive	Backup to and restores from online media can be slow depending on network/Internet connection speed[4]
Can easily be remotely monitored/audited	Online transfer and storage fees for Internet storage can be very expensive
Fewer moving parts means greater reliability	

The following are technologies related to online backup: compression and replication.

a. Compression and Deduplication

Sending or receiving backup data over a network can be very expensive in terms of transmission time, bandwidth used, and storage used at the remote site. Compression can lessen the impact of these drawbacks by reducing the amount of data that must traverse the network.

Compression works by first analyzing data before transmission and intelligently removing unnecessary data such as blank space and repeating information. When this is done at the file level, it's generally considered compression. When done at a more global scope (per tape, per disk, or for all data in a data store), it is referred to as "deduplication." Deduplication is a newer technology that inspects large stores of data for repetitions and replaces them with references to the original data.

For example, assume you email your team of ten employees a large document, and in turn, each of them saves this document to their location on a file server. When the backup system with deduplication

[4] This can be easily remediated by storing one backup copy locally as well as online. Initial backups can be accelerated by "seeding" data by copying one full backup to hard drives or tapes and sending them to the remote location. After seeding is completed, only the much smaller file changes need to be copied over the Internet.

technology runs on the file server, it notices after backing up the first copy of the file that other copies exist. When the backup system encounters the second copy, it doesn't copy the file but rather just creates a link to the first backup copy. Deduplication essentially gives you ten times the backup capacity over older methods in this example. Deduplication also makes it possible to retain many more backup sets over time, as only files (or even parts of files) that change count against your total available storage.

As the storage and bandwidth savings can be significant, compression and deduplication are used by most online backup tools. However, this technology is of no help for files of random nature, including photographs, videos, music, and pre-encrypted data.

b. Replication

Replication is simply keeping two copies of a data set synchronized across two or more locations. Replication can be used as a simple and effective way to store an offsite copy of your onsite data array for disaster recovery purposes.

3. Encryption

In this world of identity theft, we must all be diligent in keeping our customer and employee information safe. Encryption allows us to "lock" our data so that only we, or those we designate, can view the data. As a general rule, all data that leaves your organization should be encrypted.

E. Recommended Validation Schedules

No matter what method is selected, onsite and offsite backups must be validated regularly. Statistics show that, in particular, tape backups can be unreliable when they are needed the most; close to 40% of organizations have a tape restore fail each year[5], likely due to lack of testing. It is important to keep in mind that any backup system can be problematic if not validated regularly.

We recommend that customers adhere to the following validation schedule:

Weekly

Review log files, performance statistics, and usage data at least *weekly* to identify anomalies in archived data. Inspect anomalies closely to identify underlying data corruption. Most systems also send email or other alerts when a failure occurs, and these should be resolved within one business day.

Monthly

Validate your data recovery processes, both onsite and off, *monthly* by restoring some data from each class of storage and backup process, such as recoveries from:

- Offsite media
- Onsite media
- Incremental or differential backups
- Full backups
- Specialty data (data other than regular files: databases, email archives, etc.)

Quarterly

Each *quarter*, validate your backup and recovery processes by recovering one system in your environment from a particularly selected time period. Rotate the time periods and system used each quarter.

Yearly

Yearly, perform a full-scale recovery of all systems from offsite archives to temporary systems. If your disaster recovery strategy includes recovering to a remote location (and it should), this is an ideal time to test that plan as well.

[5] "Data Protection Manager – From The Customer Perspective," Microsoft, 2006.

F. Case Studies

In this section, we examine five sample environments, explaining the challenges and benefits of removable media and online media in each. We also determine an ideal configuration for each environment and compare the costs of each.

1. Sample Organizations

Organization Profile	Data Set Size	Weekly Data Change (Δ)	Standard Retention	Internet Speed up/down (Mbps)
Small Software Co.	1TB	10%	2 weeks	2/10
Small Photography Studio	2TB	0.25%	2 weeks	1/7
Small Manufacturing Co.	300GB	1%	2 weeks	0.75/6
Mid-Sized Service Co.	1TB	5%	2 weeks	10/10
Mid-Sized Higher Ed.	12TB	5%	6 weeks	100/100

Table 1: Data characteristics of sample organizations

For each organization above, there is an associated data set size, weekly data set change, standard retention, and speed of internet bandwidth.

Data set size represents the total amount of data the organization wishes to protect; generally, this is the size of one full backup of all servers in an organization.

Average weekly data change (Δ) is the percentage of the data set that changes on a weekly basis. For most established environments, Δ can be determined by the size of a differential backup taken seven calendar days after a full backup or by taking the sum of seven days of incremental backups. Averaging a month's (or even a year's) worth of differential or incremental weekly backups provides even better insight, as many organizations have end-of-month, end-of-quarter, and end-of-year events that can significantly increase Δ and even the data set size.

Standard retention notes how long each organization keeps backup data available. Organizations that retain two weeks of data, for example, can recover deleted data up to two weeks after the deletion.

Internet speed simply specifies the bandwidth available to the company in both directions: up (or to the internet or remote site), and down (or from the internet or remote site). These values are critical when considering online backup.

2. Data Protection Options

To provide a cohesive view of the different data protection options for each sample organization, we provide two primary options to compare and contrast for each organization: removable media and online backup.

3. Assumptions

We assume that each organization utilizes an onsite disk archive system for onsite backups; other onsite technologies should yield similar results.

a. Online, Offsite Backup

For online backup, we assume that 50 percent of the upstream Internet connection bandwidth can be allocated for continuous data protection use and downstream bandwidth (i.e., if an organization has a 2Mbps upstream and a 10Mbps downstream connection, we dedicate 1Mbps upstream and 1Mbps downstream for data protection use).

We assume that the online backup system uses compression and deduplication and only replicates data changes. We also assume that the initial backup does not take place over the local Internet connection, but by shipping drives or uploading via a commercial service (known as "seeding").

Costs of online backup storage vary widely, depending on the amount of data you access as well as how much is stored. To keep the math simple, we'll use the price of $1/GB per month. We feel that this is a reasonable number charged by middle-of-the-road providers for storing your data offsite and providing the necessary tools and software. Note that this per-gigabyte charge does not include labor required to implement the backup solution or maintain your backups.

b. Removable Media Offsite Backup

For removable media backup, we assume the site uses a fairly modern, standard backup tape system utilizing LTO-4 tapes and a single tape drive within a simple tape library housing 10 or more tapes (such as the Dell PowerVault 124T - $8,000, modestly equipped). For small sites, we assume a single LTO-4 drive priced at $2,500. Finally, LTO-4 media is priced at $40 per tape.

LTO-4 drives approach 350GB/hr in real life, store 800GB per tape (1.6TB with "2:1" compression), and support encryption. Newer LTO-5 drives and media are available, but LTO-4 is more prevalent and more cost effective.

While the characteristics of RDX hard drive media are different than tape, with a maximum transfer rate of 45MB/sec or 162GB/hr and random access benefits, we will categorize both tape and RDX under "Tape/Removable Media Backup" below, as they are within the same magnitude of cost. RDX media is approximately three times the cost of tape, but single RDX drives ("docks") are much less than single tape drives. When utilizing a library of RDX cartridges, the cost is on par with a tape library, meaning that RDX's advantage is for sites needing only one drive, with a maximum total data usage of 1TB (size of one full backup).

As of October 2011, 1TB RDX cartridge costs $300, and a USB docking station/drive costs $400.

4. Key Environments

Small Software Company

Organization Profile	Data Set Size	Weekly Data Change (Δ)	Standard Retention	Internet Speed up/down (Mbps)
Small Software Co.	1TB	10%	2 weeks	2/10

In the Seattle area, many companies fit this profile: smaller shops with 5-20 developers and onsite support staff. A tech-savvy staff with a high amount of data used per employee sets them apart from other similar-sized organizations.

As the focus of this company is creating software, their amount of data change tends to be higher than other companies of a similar size (10 percent per week). Using the data from the table above**Error! Reference source not found.**, we can model the performance of removable versus online media for this organization.

a. Option 1: Online Backup

Amount of data change per week

`1,024GB[6] (Data Set Size) * 10% (Δ) =` **`102GB/week`**

Utilizing 50 percent of their 2Mbps upload capacity, we have:

`(102GB[7]/week) ÷ (1Mb/sec) =` **`10 days`** `of synchronization` **`per week`**

Since this customer has enough data change to require 10 days of synchronization in a week with only seven days, we need to look at alternatives:

- Upgrade Internet connection beyond 2Mbps or install a dedicated connection just for backups
- Invest in more sophisticated duplication tools to lessen the amount of data change per week

The cost to store the 1024GB data set (1TB) online for the two-week retention period is approximately **$1,229/month**:

`(1024 + (10% * 2)) * $1 =` **`$1,229`**

b. Option 2: Tape/Removable Media Backup

Since tape backup requires an initial investment in equipment that can be reused (capital expenditure), we will amortize the cost over five years. While most tape drives and associated equipment will last longer than five years, most organizations replace equipment every five years as a precaution. Also note that rarely used tapes should last at least ten years, and up to 30 years when stored properly.

---

[6] 1TB = 1,024GB = 1,048,576MB

[7] 1MB = 8Mb, thus 205GB * 8 = 1,640Mb

Using LTO-4 technology, we can store our entire data set on one tape and an entire week of incremental backups on another tape. As such, we only need two tapes for each week of backups we require and can then send these tapes offsite. Ideally, we need one set of tapes onsite, one in transit, and one securely offsite at all times to ensure all copies are never in a single location. Therefore, we'll need six tapes on an ongoing basis; we will round up to 10 tapes to replace worn tapes and one tape drive.

The initial hardware cost for the drive is approximately $2,500 and an additional $400 for 10 tapes. Thus, for about five years of protection, we have an initial cost of only $2,900 —or only $48/month! We should also factor in offsite storage costs, which could range from free (bank vault already in use, second location) to $150 a month for media pickup and storage companies. Even accounting for expensive transport and storage (at $100/month), tape costs only **$148/month** in this scenario versus $1,229 for online media, or a savings of nearly 88 percent!

For RDX media/drives, we are looking at $3,400 for one dock and 10 pieces of media, or $57/month over five years (**$157** including offsite storage).

Additionally, as seen under Option 1 above, online backup is not possible in this environment without increasing bandwidth or purchasing advanced data reduction (deduplication) software. With LTO-4, we can copy data at 350GB/hr, which means our full backups will finish in about three hours, versus 10 days with online backup.

c. Option 3: Hard Drive Backup (non-archive quality)

Before we discuss removable hard drive backups, please understand the risks of using this media, as discussed in section "b. Commodity Hard Disk-Based Backup."

Using hard drives for this type of environment is very similar to tape. We can simply purchase six USB hard drives and rotate them as we would tapes (two onsite, two in transit, and two offsite). Unlike tape, hard drives have no given duty cycle as far as how many passes are considered safe before they wear out; however, with disks, it's only a matter of time before one fails. Due to this likelihood of failure, we will purchase eight drives to have a few spares on hand.

For under $100 each, 1TB drives are easily attainable. These drives, in general, should be good for three to five years, especially with our spares, so we are up and running for just $800 or about **$14/month**!

While these inexpensive drives won't match the speed or reliability of tape, they will likely reach 180GB/hr, finishing our full backup in less than six hours.

d. Feasibility/Conclusion

As per the explanations above, for this small software development company, online backups are not feasible without upgrading from a standard internet connection or investment in additional backup software. Additionally, at a price of $1,229/month, many businesses in this space would be priced out of the market.

With the downsides of external hard drive backups (reliability/durability of media), it is hard to recommend external hard drives for all but the most cash-strapped companies.

Tape is the winner in this environment: fast, affordable, and able to meet all requirements.

Small Photography Studio

Organization Profile	Data Set Size	Weekly Data Change (Δ)	Standard Retention	Internet Speed up/down (Mbps)
Small Photography Studio	2TB	0.25%	2 weeks	1/7

Small photography shops are in a tough spot. All use digital photos these days, and even with those using compressed photographs, the data sizes are huge for companies of their size (usually just one to three people). Their photos are highly critical to their survival, however; lose them, and recurring revenue goes out the window. Budgets are also tight, as they clearly do not have the resources of larger firms, even though their amount of critical data rivals these firms.

a. Online Backup

Amount of data change per week:

```
2048GB * 0.25% (Δ) = 5GB/week
```

Utilizing 50 percent of their 1Mbps connection gives us 512kbps, and we have:

```
(5GB/week) ÷ (512kbps) = 23 hours per week
```

This is completely feasible; all of the weekly data change can be absorbed online in just a single day.

The cost to store the 2TB data set online is approximately **$2,058/month**:

```
((2048 + (0.25% * 2)) * $1) = $2,058/month
```

While storing data online is possible, the cost is likely prohibitive.

b. Option 2: Tape/Removable Media Backup

With such a low data change rate, this company can easily survive with just a single tape or RDX drive with eight pieces of media (two onsite, two offsite, two in transit, and two spares).

With RDX, the capital investment is about **$2,800** for a USB docking station and eight cartridges. For a single LTO-4 tape drive and eight tapes, the investment is **$2,820**.

Both options clearly can handle the size and rate of change for this backup set.

Accounting for $100 for media storage per month, both tape and RDX come in over five years at **$147/month**. Due to the ease of using RDX for smaller data sets, RDX probably has the edge here.

c. Option 3: Hard Drive Backup (non-archive quality)

For this 2TB data set, 2TB external hard drives are a good fit. These drives are attainable for about $125 a piece. Keeping one onsite, one offsite, one in transport, and a spare, we need four drives for an investment of $500. These drives, in general, should be good for three to five years, especially with our spares, so we are up and running for just $500, or about **$8.33/month** or **$108.33/month** with storage.

d. Feasibility/Conclusion

In this environment, external non-archive quality hard drives are a reasonable option. Cost is a huge factor for this environment, and the additional $39/month for "proper" archive equipment will be significant for these businesses.

The focus for these customers should be a solution that is as easy as possible, as creative folks tend not to be the most linear thinkers, and thus rotating media offsite may be the last thing on their minds but critically important nonetheless. They are very likely to store media at an employees' home or at the bank for additional cost savings.

Small Manufacturing Business

Organization Profile	Data Set Size	Weekly Data Change (Δ)	Standard Retention	Internet Speed up/down (Mbps)
Small Manufacturing Co.	300GB	1%	2 weeks	0.75/6

a. Online Backup

Amount of data change per week:

```
300GB * 1% (Δ) = 3GB/week
```

Utilizing 50% of their 768kbps (0.75Mbps) gives us 384kbps, and we have:

```
(3GB/week) ÷ (384kbps) = 18.2 hours per week
```

This is completely feasible; all of the small manufacturer's data can be synchronized in less than one day.

The cost to store the 300GB data set online is approximately **$306/month**:

```
((300 + (1% * 2)) * $1) = $306/month
```

b. Option 2: Tape/Removable Media Backup

With such a low backup set size and data change rate, this company can easily survive with just a single tape or RDX drive with five pieces of media (one for onsite, one offsite, one in transit, and two spares).

With RDX, the capital investment is about **$2,000** for a USB docking station and five cartridges. For a single LTO-4 tape drive and five tapes, the investment is **$2,700**.

Both options clearly can handle the size and rate of change for this backup set.

Accounting for $100 for media storage per month, RDX comes in over 5 years at **$133/month** and tape at **$145/month**. Given that the company has a secure offsite location to store media (like a bank vault), the price drops to just $33 and $45 a month, respectively.

c. Option 3: Hard Drive Backup (non-archive quality)

Using hard drives for this environment is very similar to tape or RDX. We can simply purchase five USB hard drives and rotate them as we would tapes (one onsite, one in transit, and one offsite). We'll also want two spares, for a total of five drives.

For under $100, 1TB drives are easily attainable. These drives, in general, should be good for three to five years, especially with our spares, so we are up and running for just $500, or about **$8.33/month** or **$108.33/month** with storage.

d. Feasibility/Conclusion

In this scenario, all options are roughly the same price ($159–$306/month). Due to the ease of management, online backup for this environment is attractive. While it's nearly double the price of cheap hard drive backup, it has some very compelling advantages, namely:

1. A service provider can easily and inexpensively audit backups remotely.
2. Employees save time and money by not shipping or transporting tapes/cartridges offsite.
3. Small restores – retrieval of a single deleted file, for example – are very quick,
4. Since this company has a relatively small data set, even large restores from the internet complete quickly (at 80 percent utilization, or 4.8Mbps, 300GB may be restored in 6 days).

Mid-Size Services Organization

Organization Profile	Data Set Size	Weekly Data Change (Δ)	Standard Retention	Internet Speed up/down (Mbps)
Mid-Sized Service Co.	1TB	5%	2 weeks	10/10

a. Online Backup

Amount of data change per week:

```
1TB * 5% (Δ) = 51GB/week
```

Utilizing 50 percent of their 10Mbps connection gives us 5Mbps, and we have:

```
(51GB/week) ÷ (5Mbps) = 24 hours per week
```

This is completely feasible; all of the data can be synchronized in about one day.

The cost to store the 1TB data set online is approximately **$1,126/month**:

```
((1024 + (5% * 2)) * $1) = $1,126/month
```

b. Option 2: Tape/Removable Media Backup

With 1TB of data and given normal compression expectations, one full backup for this environment will fit on a single tape, and a second tape can hold an entire week of incremental backups. Thus, eight LTO-4 tapes or RDX media will cover the needs of this midsized service company; two tapes onsite each week, two offsite, two in transit, and two spares.

With RDX, the capital investment is about **$2,800** for a USB docking station and eight cartridges; for a single LTO-4 tape drive and eight tapes, the investment is **$2,820**. The prices are very close in this environment. The price per piece of media for RDX is much higher than tape, but the tape drive cost is much more expensive than the RDX dock, bringing the two technologies on par, cost-wise. Should the company require more retention in the future, tape would rise at a much lower cost until a need for more advanced tape handling equipment is warranted (such as when backups cannot fit on one tape).

Both options clearly can handle the size and rate of change for this backup set.

Accounting for $100 for media storage per month, both RDX and tape come in over five years at **$147/month**. Given that the company has a secure offsite location to store media (like a bank vault), the price drops to just $47 a month.

c. Option 3: Hard Drive Backup (non-archive quality)

Using hard drives for this service company is very similar to tape or RDX. We can simply purchase eight USB hard drives and rotate them as we would tapes (two onsite, two in transit, and two offsite). We'll also want two spares, for a total of eight drives.

For under $100, 1TB drives are easily attainable. These drives, in general, should be good for three to five years, especially with our spares, so we are up and running for just $800 or about **$13.33/month** or **$113.33/month** with storage.

Given the risks with this type of storage medium, and looking at the investment over five years, RDX or LTO media seems like a much better choice for the small cost difference of about $34/month. However, this is an option for the most cash-strapped companies and is by far better than no backup at all.

d. Feasibility/Conclusion

For this midsize firm, it is clearly more cost-effective to use more traditional media for offsite backups. Even with the transportation charges, the offsite backup system costs just $147/month, compared to $1,126 for online storage.

Many salespeople would argue that online backup is still superior for this environment, as it is less labor-intensive. This may have merit in some environments, but the management effort of online storage is not insignificant, either. For a cost savings of nearly $1,000 per month, you can easily hire a competent company to manage the tapes for you.

Clearly, removable media is the winner for this environment.

Mid-Size Higher Education

Organization Profile	Data Set Size	Weekly Data Change (Δ)	Standard Retention	Internet Speed up/down (Mbps)
Mid-Sized Higher Ed.	12TB	5%	6 weeks	100/100

a. Online Backup

Amount of data change per week:

```
12TB * 5% (Δ) = 615GB/week
```

With 50Mbps bandwidth available, we have:

```
(615GB/week) ÷ (50Mbps) = 28 hours
```

Clearly, the internet bandwidth is not a problem in this environment, but at $1/GB, the monthly storage cost of online media is nearly **$16,000 a month!**

b. Option 2: Tape/Removable Media Backup

Compared with the other environments we've looked at, the 12TB of data at this institution is quite large. As such, to properly handle tapes, we will need a tape changer, such as the Dell PowerVault 124T, which runs about $8,000. With 12TB of data, we'll also need to utilize 12 tapes per week, plus an additional tape for incremental backups (615GB), for a total of 13 tapes per week. Since we need six weeks of retention, that is 72 tapes total, and we should add a few spares, taking us to 75 tapes total.

With 75 tapes, and the tape library, we are looking at a total cost of $11,000, or $183/month over five years. Adding in offsite storage, we get to **$333/month** (assuming $150/month in storage fees for this number of tapes).

Given the large amount of media required, RDX is not likely to work well. RDX would require some kind of library, which is hard to find. Assuming a $3,500 library, such as the Tandberg Data RDX QuikStation™ (8900-RDX), and 75 pieces of 1TB media, we are looking at an upfront investment of $26,000 or $433/month over five years. Adding in offsite storage, we are at **$583/month**, nearly double the price of tape media.

While the hardware we assume in this paper is good enough for comparison, in reality, RDX would not be feasible in this environment, as 8TB is the current maximum storage capacity (12TB may fit with compression). While this may work on paper, in practice, some excess online capacity is likely needed. The same with LTO-4 / Dell 124T: with 16 tapes maximum and a total storage capacity of 12.8TB, some excess capacity and additional tape drives are likely a requirement.

c. Option 3: Hard Drive Backup (non-archive quality)

Due to the sheer amount of data and moving parts in this environment, non-archive quality hard drives are highly *not* recommended. However, should we try to implement this, our costs would be $7,500 up front plus $150/month for storage, or around **$275/month** over five years (a savings of just 17 percent).

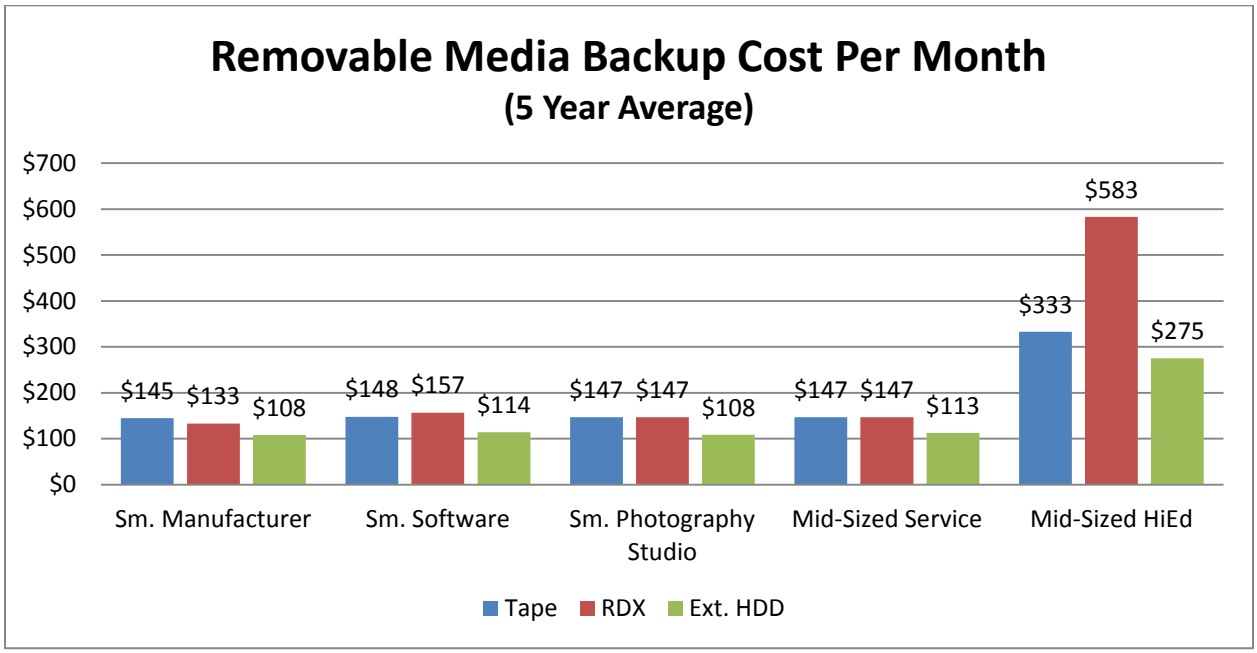
External hard drives are not feasible for several reasons here. Aside from reliability, consider the number of free USB ports on a given server, the ability to easily span multiple drives, and the time required to manage dozens of drives simultaneously.

d. Feasibility/Conclusion

This use case shows that tape is the clear winner. Sure, commodity hard drives are cheaper, but not drastically (just $58 less per month). Commodity hard drives connected via USB or eSATA would also be much more difficult to manage than tape or even RDX, as these technologies are designed for multiple pieces of media per backup, where external hard drives are not.

5. Summary

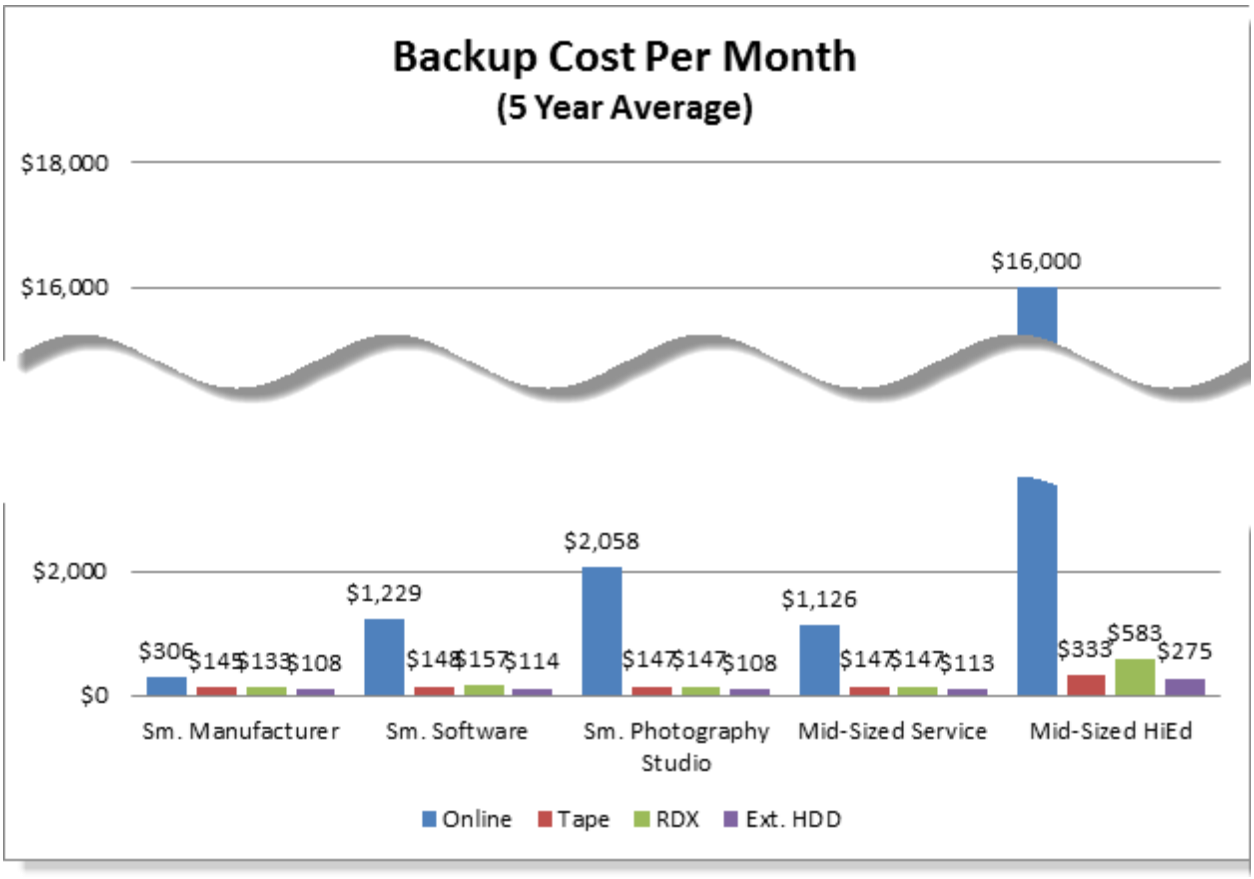
Comparing the three types of removable media we covered: tapes, RDX cartridges, and commodity external hard drives, costs are roughly the same for each in the environments studied. The one standout is that, as the amount of data increases, tape actually becomes cheaper than RDX. Tape is also the most reliable and speedy option, and as such, it is an obvious choice for organizations with larger data sets.



Online backups were not a fit for most of the use cases we examined but may be practical in environments where cost is less of a factor and onsite technical expertise is limited. A thorough, in-depth environment and needs assessment must be done in order to recommend one method over another.

G. Conclusion

In closing, the feasibility of online storage for offsite backups is the *exception* and not the rule as a fit for small and medium sized businesses. The following chart shows that only the small manufacturing environment has a reasonable price tag for online backup; in all other environments online, or cloud, backup is nearly ten or more times the monthly cost of removable media options.



This is not to say that online backups are never appropriate. Other drivers, such as recovery time objectives (time window that a failed system/application must be recovered within), may sway larger SMBs toward some type of online backup. In most of these cases, online backup is prudent for select, highly critical applications and data.

Therefore, based on the specific requirements of a particular business, a reasonable plan of action for data protection could include both removable media and online backup protection. However, it is important to look at data protection options from a holistic standpoint and truly examine the costs and feasibility of online backup before making a decision to go into the "cloud." By choosing a firm that looks at these variables and provides a plan that works best for your business, you can confidently choose the best data security solution, which may not necessarily be cloud backup. Don't discount tape as an outdated relic; tape could be the best option to save critical data and restore lost files.

Need help setting up a backup, archive, and disaster recovery strategy? Need help selecting the proper tools, media, and software? Never want to worry about backup and recovery again? Call Red Wire Services today for a complimentary assessment at (206) 829-8621, email sales@redwireservices.com**, or visit us online at http://www.redwireservices.com.**

RED WIRE SERVICES

PO Box 46454 // Seattle, WA 98146 // T 206.829.8621 // F 206.826.1223 // info@redwireservices.com

About Red Wire Services

Red Wire Services shelters small and medium-sized businesses from technological disasters and their impact, including lost revenue, embarrassment, and possible business collapse. Unlike other outsourced technology firms, we focus solely on data protection services to ensure your data is as safe as possible.

We work with you to identify your most critical data and business systems and then implement a customized plan to protect you from disasters — large and small.

About The Author

Nick Webb is the founder of Red Wire Services, LLC and has made a career of designing solutions that improve data availability and enable disaster recovery. Nick brings more than 10 years of experience planning, implementing, and maintaining best practice solutions for systems management for a wide range of organizations. His firm, Red Wire Services, specializes in preparing medium-sized businesses to survive technological disasters — while also helping unprepared organizations recover their data and systems so they can return to operation.

Nick has a broad technological background and a proven track record in projects involving all major operating systems, virtualization, networking, backup, and storage systems, allowing him to plan for and protect your data no matter where it is stored. Highly regarded by clients and teammates, Nick is driven by quality yet able to calmly assess client requirements. He brings advanced levels of insight, organized planning, and effective communication to bring you the best technical and business solutions for your current and future needs.

Nick holds a Bachelor of Science degree in Computer Science from the University of Idaho and is a certified LPIC-2 Advanced Linux Professional. He enjoys giving back to his profession by mentoring newcomers in the field, is an active member of the League of Professional System Administrators, and has been a presenter at Linux Fest Northwest since 2010.