Archiving to AWS vs Iron Mountain

The Diminishing Point of Return Comes Quickly

January 2019

Table of Contents

Amazon’s Glacier is a cloud storage offering designed for the storage of data which doesn’t require fast access or have a recovery time objective (RTO). Iron Mountain is a similar service for storing data offsite, but uses physical data tape storage from the customer’s data center unlike cloud storage approach of moving data directly from disk across wide area networks (WAN) into a remote, cloud storage facility.

Introduction

One of the most common words to frequent conversations in the computer industry today is "cloud." Seemingly, this single word would represent a singular approach, but that is far from the truth. There are many types of applications/uses for the cloud. There has been a major push to take software applications to the cloud (Software as a Service/SaaS). An even broader approach takes the compute process as well as applications to the cloud (cloud compute). One of the most popular uses for the cloud is transcoding and distribution – the ability to convert a single media file from its source format into multiple formats which allow the file to be played back on various devices such as tablets, smartphones, PCs, etc. — and then distribute it to users. Backing up files to the cloud has also become a popular process, especially for small data sets such as on smartphones or personal computers.

A relatively new approach to cloud use involves backing up *large* data sets to the cloud. This holds great appeal for organizations wanting to move data offsite for disaster recovery. Cloud backup itself is not new. Smart phones and personal computers have used cloud backup services for many years to backup small amounts of data. What *is* new is attempting to move hundreds of terabytes, or even petabytes, to the cloud. The bandwidth to move large data sets, and more importantly restore large data sets, is both cost and performance prohibitive. In moving data to the cloud, the upload or backup process can be accomplished by "trickle feeds," slowly moving small amounts of data until all data has been moved. If this approach is used for disaster recovery, a "trickle restoration" is insufficient when some form of disaster has wiped out an entire server, server groups, or an entire data center. It could take a week or more to restore 100TB over a relatively fast internet connection. Amazon offers their Snowball and Snowmobile services for bulk upload, and this approach may have merit for some customers, but these products don't significantly change the economics of cloud backup.

This paper is designed to examine the true cost of storing data in the AWS Glacier Cloud versus in an Iron Mountain Storage Facility. To source data for this paper, we used Spectra Logic as a "real-world" example (Spectra uses both Amazon and Iron Mountain), and obtained actual quotes from both vendors. To allow users to make calculations and decisions based on their *own* data sets, we start by looking at the cost of storing 6TB of data (a single LTO-7 tape) using Amazon Glacier versus Iron Mountain. While it's unlikely that an organization would store a single tape offsite, this analysis makes it easy to compare costs for any size of storage requirement. We will then examine the cost of storing a data set starting at 200TB, a much more typical data center scenario. Rather than focusing exclusively on storage costs, we analyze restore costs, restore performance, TCO, true geographic separation, and examine lesser discussed topics such as vendor lock-in and the value of genetic diversity in storage mediums. The goal of this paper is to allow data users, large and small, to make decisions about data storage providers based on the whole of parameters which are so important in assuring long-term digital preservation.

Iron Mountain

Iron Mountain has become the de facto standard in offsite data storage, protecting organizational assets since 1951. By offering onsite pickup and drop-off services, Iron Mountain has streamlined the process of storing and retrieving an organization's data on tape offsite. Based on weekly pick-up and drop-offs via secure transportation, customers have reliable and predictable storage transportation. Customers can access their data in as little as three hours. Media is stored in secure containers, in an environmentally controlled environment, to ensure long-term media survival (typically 30 or more years for tape).

What is Amazon Glacier

Amazon Glacier offers customers the ability to store and access their data over the internet. Glacier is part of Amazon's Web Services, which is commonly referred to as the premier public cloud offering. It is primarily used for archiving and long-term backup, but is fully integrated into Amazon's other public cloud offerings. Amazon Glacier provides three options for access to archives, from a few minutes to several hours with standard access time being between three to five hours. This is the wait time to download the data you have requested. Depending on the size of the data set and network connection, the amount of time until your data is ready to be used will vary. The more an organization pays for network bandwidth, the faster data will be downloaded.

In late November of 2018, Amazon announced a new tier of storage -- Amazon Deep Glacier -- that delivers storage for as low as $0.00099 per GB per month and is expected to be available some time in 2019. Similar to Amazon Glacier, the new Deep Glacier storage tier is not designed to be used as a storage tier when data needs to be accessed in any quick fashion. The Deep Glacier tier of storage will make data available within 12 hours, and at that point, users can begin their download. Data can be made available sooner, but that speed of access comes at a price. At the time of this paper, only the storage prices have been published, but nothing has been published on the cost to retrieve data, or the cost to retrieve data sooner than the 12 -hour announced time of access.

Cost Comparisons for 9TB of Data Storage

Let's examine the data lifecycle of a single tape that represents a critical imaging scan that is a total of 9TB in size. This body scan has the potential to be involved in the cure for cancer and must be permanently retained in the event that reevaluation is needed in the future. In a traditional data center, there is a good chance this data would be stored on tape (most likely LTO tape technology). LTO is by far the most commonly used data tape technology today. If LTO-7 Type M tape technology is used, this data set will be stored on a single LTO-7 Type M tape. For security and data availability, a second copy is being made for disaster recovery purposes. The question becomes: What is the best method of storing this second copy of data? We examine the options of storing the 9TB image in the cloud (Amazon Glacier and Deep Glacier) or the more traditional method of vaulting the data in an offsite data repository (Iron Mountain).

A cost analysis is performed with the following parameters:

- Base 10 calculations when converting TB to GB
- An LTO-7 Type M tape at list price at time of publication is used ($107 per cartridge)
- Iron Mountain's list pricing is used
- Amazon Glacier and Amazon Deep Glacier published pricing is used
- This is a storage-only calculation
    - No cost for bandwidth to get data to the cloud
    - No cost for a tape library to get data onto tape

LTO-7 Type M Tape Cartridge						
Year 1						
Offering	**Capacity**	**Cost of Tape**	**Capacity in GB**	**Cost per Year**	**Cost per month**	**Cost per GB / per month**
Iron Mountain	9TB	107.00	9,000	$ 119.00	$ 9.92	0.00110
Amazon Glacier	9TB		9,000	$ 496.80	$ 41.40	0.00460
Amazon Deep Glacier	9TB		9,000	$ 108.00	$ 9.00	0.00100

Years 2+						
Offering	**Capacity**		**Capacity in GB**	**Cost per Year**	**Cost per month**	**Cost per GB / per month**
Iron Mountain	9TB		9,000	$ 12.00	$ 1.00	0.00011
Amazon Glacier	9TB		9,000	$ 496.80	$ 41.40	0.00460
Amazon Deep Glacier	9TB		9,000	$ 108.00	$ 9.00	0.00100

5 Year Total for Iron Mountain	$ 167.00
5 Year Total for Amazon Glacier	$ 2,484.00
5 Year Savings for storing in Iron Mountain over Amazon Glacier	$ 2,317.00

5 Year Total for Iron Mountain	$ 167.00
5 Year Total for Amazon Deep Glacier	$ 540.00
5 Year Savings for storing in Iron Mountain over Amazon Deep Glacier	$ 373.00

To store a single LTO-7 Type M tape (9TB of uncompressed data), Iron Mountain will charge $12.00 to store the tape for 1 year – equivalent to $1.00 per month for storage. Converted to a cost per GB, this would cost $0.00011 per GB per month. To store the same 9TB of data in Amazon Glacier, Amazon will charge $0.0046 per GB per month which equates to $27.60 per month or $331.20 per year. Using Amazon Deep Glacier to store the same 9TB of data, the charge would be $0.001 per GB which equals $9.00 per month or $108.00 per year. An organization would save $2,317.00 over 5 years by storing a single tape using Iron Mountain storage services versus Amazon Glacier services and a $373.00 savings over using Amazon Deep Glacier. The five-year savings is compelling. If this data is to be archived for decades, or indefinitely as the above scenario is set, the savings become game changing.

As with any technology, new generations and advancements occur on a regular basis. Tape is no different. The LTO tape technology roadmap shows a strong future with projections through LTO-12. In 2017 the LTO-8 tape technology became publicly available with an astonishing 12TB of uncompressed capacity on a single tape. With the release of a new tape technology the initial price of media is high but with the additional capacity it provides a solid option for organizations. The cost of LTO-8 media is expected to be greatly reduced when supply issues have been resolved. This is expected to happen in 2019 where the cost of a single LTO-8 tape is expected to be 35%-45% less than the current cost per tape. For the purpose of this paper we have used current list price of media, but cost savings are expected to be greater as LTO-8 media prices reduce.

LTO-8 Tape Cartridge

Year 1

Offering	Capacity	Cost of Tape	Capacity in GB	Cost per Year	Cost per month	Cost per GB / per month
Iron Mountain	12TB	240.00	12,000	$ 252.00	$ 21.00	0.00175
Amazon Glacier	12TB		12,000	$ 662.40	$ 55.20	0.00460
Amazon Deep Glacier	12TB		12,000	$ 144.00	$ 12.00	0.00100

Years 2+

Offering	Capacity		Capacity in GB	Cost per Year	Cost per month	Cost per GB / per month
Iron Mountain	12TB		12,000	$ 12.00	$ 1.00	0.0000833
Amazon Glacier	12TB		12,000	$ 662.40	$ 55.20	0.0046000
Amazon Deep Glacier	12TB		12,000	$ 144.00	$ 12.00	0.00100

5 Year Total for Iron Mountain	$ 300.00
5 Year Total for Amazon Glacier	$ 3,312.00
5 Year Savings for storing in Iron Mountain over Amazon Glacier	$ 3,012.00

5 Year Total for Iron Mountain	$ 300.00
5 Year Total for Amazon Deep Glacier	$ 720.00
5 Year Savings for storing in Iron Mountain over Amazon Deep Glacier	$ 420.00

When a new tape technology comes out, an organization using Iron Mountain would pay the same amount for their storage because Iron Mountain bills per tape cartridge, not by the capacity it holds. When LTO-8 tape technology becomes more cost advantageous, an organization will be able to gain 33% on their storage capacity without realizing any additional storage costs. It is important to note that these calculations are not using compression which will further move finances in the favor of "per tape" storage cost versus "per GB" storage cost. Public cloud providers use the "per GB" method of calculating monthly storage costs based on the amount of data stored during the month. As data sets increase in size, the bill received from the cloud vendor will increase as well.

As the above example points out, storing one LTO-8 tape (12TB of uncompressed data) via Iron Mountain remains steady at $12.00 per year – equivalent to $1 per month. Due to the increase in data being held on that tape however, we now see cost per GB drop significantly to $0.000083 per GB per month. To store 12TB of data in Amazon Glacier at $0.0046 per GB per month, storage costs will increase to $55.20 per month – equivalent to $602.00 per year. An organization can save $3,012 over 5 years by storing a **single** LTO-8 tape using Iron Mountain storage services versus Amazon's Glacier storage services. When organizations use Amazon Deep Glacier at $0.001 per GB per month, the storage costs will be $12 per month, and $84.00 per year. A total savings for 5 years would equal a $420.00 savings using Iron Mountain vs Amazon Deep Glacier.

Capital Expense (CapEx) versus Operational Expense (OpEx)

The above cost analyses do take into account the cost of each physical tape (CapEx), but they are primarily aimed at OpEx, the operational cost of moving and storing data. When comparing any "on-premise" data storage approach to a similar "cloud" storage approach, the conversation of CapEx vs. OpEx is an important one. Is it more affordable to own storage equipment (on-premise) or outsource that cost (cloud)? There are many pros/cons that could be covered, but we will start with cost.

In our first example above, it would be improbable that one would buy an automated tape library to deal with a single tape. If the backup data set is truly 9TB, even though the cost of storing a single tape for 5 years in Iron Mountain ($160) would be much less than 5 years in the cloud ($2484 using Amazon Glacier), the cost of a small tape library to deal with a single tape would negate the savings for many years. Many factors have to be examined. Do you currently have a tape library or will you be purchasing a new tape library? What will the salary be for the individual maintaining the tape library? What will the salary be for the individual maintaining the system sending data to the cloud? While it would be impossible to set every parameter for every data center in this white paper, it is important to introduce the cost of the tape library itself.

Total Cost of Ownership Example

For the next example, we take a “real world” look at comparing Iron Mountain and Amazon Glacier using our own Spectra Logic data center in Boulder, CO. The following example uses real costs, data sets, ecosystem requirements, recovery time objectives (RTO), and service level agreements (SLA). Spectra uses this storage as a disaster recovery option, and only accesses this data on an as-needed basis (rarely).

When looking at a true Total Cost of Ownership, or TCO model, there are a number of other factors that must be considered above and beyond the cost of storage alone. When dealing with a cloud provider, these costs are not as easily identified as with traditional, hardware-based solutions due to the vast amount of options, billing parameters, recall fees, and early deletion fees associated with public cloud offerings. For this TCO model we attempt to identify all costs associated with using Iron Mountain and Amazon Glacier for a backup/disaster recovery storage workflow. Due to the fact that the new Amazon Deep Glacier tier of storage has not published costs to access and download, that tier of storage is not used in this example. This TCO analysis is performed with the following parameters which represent Spectra’s needs:

Data Capacity Requirements:

- 200TB starting capacity
- Steady growth rate of 5TB each month.
    - Note that Spectra assumes that it will keep versions of projects and files, which is why the steady growth rate.

- 1 Year analysis for data storage
- Begins in January and ends in December (12 months later)
- Ending capacity 255TB capacity

Iron Mountain Storage Specifics:

- $3,100 in LTO-7 Type M media (29 tapes – enough media for full year)
- Using small containers that hold 10 tapes per container
    - Cost = $10.00 per container
- Contract states that Spectra cannot exceed 8 containers (80 tapes)
- Weekly pickup and drop-off by Iron Mountain from Spectra costing $185 per month
- For a special request to bring back a full tape set from Iron Mountain it would take 3 hours and there would be an additional charge of $170.00

Cloud Storage Ecosystem Specifics:

- Amazon Glacier for cloud storage
    - $0.0046 per GB per month for storage
- Trickle data to the cloud using 150 Mbps connection. (This is the best price Spectra could find to connect from its Boulder facility.)
    - Cost = $259.90 per month
- Ability to transfer 1.62TB per day, if using full bandwidth potential for 24 hours
- To pull full data set out of cloud after one year (255TB capacity), it would require 260 days for a full retrieval of the data set without increasing the network connection.
- It would cost over $20,000 in retrieval costs from Amazon Glacier to access the full data set based on per GB retrieval fees.

Total cost of ownership results

Iron Mountain cost analysis

Month	Capacity in TB	Number of tapes	Number of Containers	Cost for containers	Weekly pickups	Total for month
January	200	23	3	$ 30	$ 185	$ 3,318
February	205	23	3	$ 30	$ 185	$ 215
March	210	24	3	$ 30	$ 185	$ 215
April	215	24	3	$ 30	$ 185	$ 215
May	220	25	3	$ 30	$ 185	$ 215
June	225	25	3	$ 30	$ 185	$ 215
July	230	26	3	$ 30	$ 185	$ 215
August	235	27	3	$ 30	$ 185	$ 215
September	240	27	3	$ 30	$ 185	$ 215
October	245	28	3	$ 30	$ 185	$ 215
November	250	28	3	$ 30	$ 185	$ 215
December	255	29	3	$ 30	$ 185	$ 215
					Total cost for 1 year	$ 5,683

Amazon Glacier cost analysis

Month	Cloud cost per GB	Storage Costs	Internet Costs	Monthly Total
January	0.0046	$ 920	$ 260	$ 1,180
February	0.0046	$ 943	$ 260	$ 1,203
March	0.0046	$ 966	$ 260	$ 1,226
April	0.0046	$ 989	$ 260	$ 1,249
May	0.0046	$ 1,012	$ 260	$ 1,272
June	0.0046	$ 1,035	$ 260	$ 1,295
July	0.0046	$ 1,058	$ 260	$ 1,318
August	0.0046	$ 1,081	$ 260	$ 1,341
September	0.0046	$ 1,104	$ 260	$ 1,364
October	0.0046	$ 1,127	$ 260	$ 1,387
November	0.0046	$ 1,150	$ 260	$ 1,410
December	0.0046	$ 1,173	$ 260	$ 1,433
			Cost for 1 Year	$ 15,677

Total Cost of Ownership Comparison

		Iron Mountain		Amazon Glacier		
Month	Capacity in TB	Weekly pickups	Total for month	Storage Costs	Internet Costs	Monthly Total
January	200	$ 185	$ 3,318	$ 920	$ 260	$ 1,180
February	205	$ 185	$ 215	$ 943	$ 260	$ 1,203
March	210	$ 185	$ 215	$ 966	$ 260	$ 1,226
April	215	$ 185	$ 215	$ 989	$ 260	$ 1,249
May	220	$ 185	$ 215	$ 1,012	$ 260	$ 1,272
June	225	$ 185	$ 215	$ 1,035	$ 260	$ 1,295
July	230	$ 185	$ 215	$ 1,058	$ 260	$ 1,318
August	235	$ 185	$ 215	$ 1,081	$ 260	$ 1,341
September	240	$ 185	$ 215	$ 1,104	$ 260	$ 1,364
October	245	$ 185	$ 215	$ 1,127	$ 260	$ 1,387
November	250	$ 185	$ 215	$ 1,150	$ 260	$ 1,410
December	255	$ 185	$ 215	$ 1,173	$ 260	$ 1,433
		Total cost for 1 year	$ 5,683		Cost for 1 Year	$ 15,677

When examining the storage costs associated with keeping a disaster recovery copy in Iron Mountain (Orange) vs Amazon Glacier (Blue), it is clear that choosing Iron Mountain is much less expensive than using the public cloud. With a cost savings of just under $10,000 per year, organizations looking to secure their data at an affordable price can choose Iron Mountain and realize substantial savings over public cloud offerings.

Total Solution Analysis

In the next example we look at a full solution setup where an organization is starting from scratch and deciding which option works best for their organization.

In this example we look at a 5-year total cost of ownership with the following assumptions and parameters:

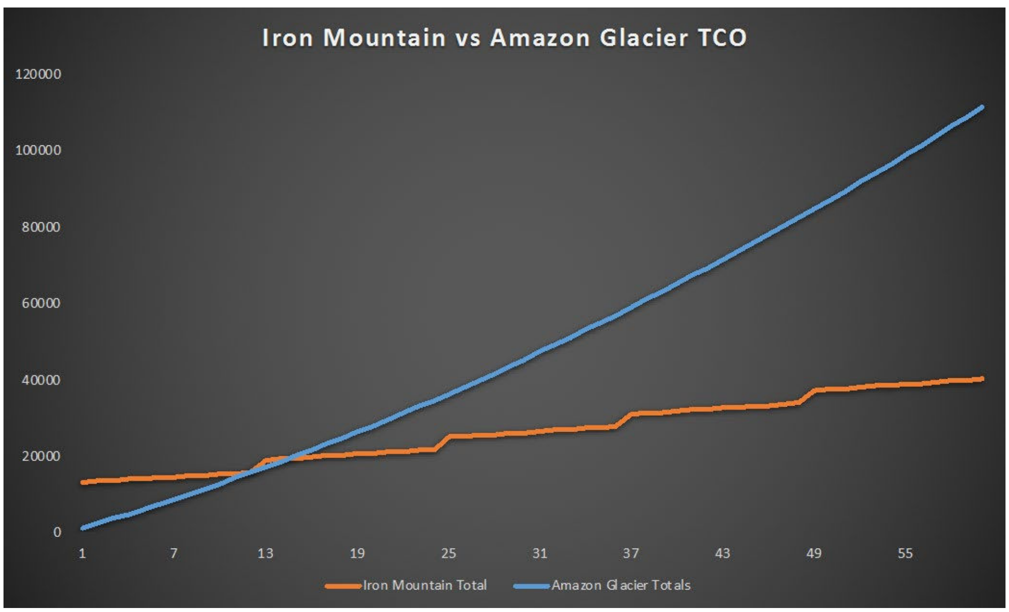
- 5-year model starting with 200TB of data
- Growing by 5TB per month

- Tape library purchase in month 1 for Iron Mountain example
    - Spectra Stack Tape Library
    - 2 * LTO-8 Tape Drives
    - 56 LTO-7 Type M tapes
        - Tape purchases occur yearly when tapes are needed
    - Slot licenses for 60 Tapes
    - Total cost of hardware = $14,500
    - Total cost of media = $6,306
    - Annual maintenance on the Spectra Stack tape library next business day
- Iron Mountain offsite data storage service
    - Same pricing as used in Spectra Logic example – Iron Mountain list pricing
    - Using small containers that hold 10 tapes per container
        - Cost = $10.00 per container
    - Contract states that Spectra cannot exceed 6 containers (60 tapes)
    - Weekly pickup and drop-off costing $185 per month
- Amazon Glacier Public Cloud Storage
    - Same pricing used in Spectra Logic example – published pricing

Year	Yearly Total for Iron Mountain	Yearly Total for Glacier
2019 Total	15,634.60	15,676.80
2020 Total	5,736.99	18,988.80
2021 Total	5,796.99	22,300.80
2022 Total	5,886.99	25,612.80
2023 Total	5,976.99	28,924.80
Grand Total	39,032.56	111,504.00

*Appendix A - Full monthly breakdown of all costs

This example points to a common phenomenon in comparing on-premise storage to cloud storage – on-premise storage for year 1 is significantly higher in CapEx than cloud and vice versa. OpEx is significantly lower in year 1 for on-premise and vice versa. Factoring in the cost of a tape library still shows a significant savings over utilizing a public cloud such as Amazon Glacier over 5 years. With the given data set, hardware, and retention period, Amazon Glacier is still more than twice the cost of installing a tape library and utilizing Iron Mountain. The breakeven point for on premise cost occurs in the 15th month, or just over a year into this example.



There are some other advantages to the CapEx model, for instance with only the cost of additional media the user can make a duplicate set of tapes to reside at their facility. All tape media can be encrypted with the library's built in encryption/key management system, and in many cases the same library can be used for other applications.

What Do You Do When a Disaster Hits?

When organizations experience an event resulting in significant data loss, it is commonly referred to as a disaster. This could be the result of human error, natural disaster, or cyberattack. It could affect a single server, group of servers or take out an entire data center. The earlier calculations focus on how we make copies of data and move it out of the way of a possible disaster; but, it's only half of the equation. It's the half we usually think of *before* the disaster strikes. Disaster recovery planning requires that the second half of the equation be considered ahead of time as well – How do we get the data back *after* the disaster?

Recovery of data is the only reason we've done any of this, so that's where true forward thinking and planning come into play. By understanding the amount of downtime which can be absorbed without impacting day-to-day operations, the size of the data set, the tools necessary to recover the data, the means by which the data will be transferred back, and the cost associated with recalling part, or all, of your data, organizations can make informed decisions and be prepared for any disaster.

Amazon Glacier Retrieval Considerations

One of the major limiting factors to accessing data stored in the cloud is network bandwidth. This is an insignificant problem when accessing a few files, but becomes a much larger consideration when large data retrievals are required. One must keep in mind that existing corporate bandwidth used for daily operations will now be shared with the restoration process. Rarely if ever is an organization willing to stop all typical use of their WAN connection until a large data restoration is completed. The calculations above are based on a 150 Mbps trickle feed rate to move data into the cloud. If this same bandwidth were wholly dedicated to a full disaster recovery of 440TB of data, that would be a 284-day recovery period! Even using a wholly dedicated 10 Gbps bandwidth, which is highly unlikely and cost prohibitive to be available solely for data recovery, the restore period would run for over 100 hours. When downloading large data sets, the potential to over utilize available bandwidth is virtually a given, and would cripple existing corporate operations.

Another important consideration in cloud data retrieval is the cost associated with accessing *your* data. When using a public cloud such as Amazon Glacier, users don't actually own their data but rather rent it from Amazon. If users need it back, they have to pay for it. Data that costs less than half a cent per GB to store will cost 5 cents per GB to retrieve. You have to pay a 1,000% increase on your storage costs to recall that data.

Iron Mountain Retrieval Considerations

Iron Mountain provides a very straightforward process for data recovery. Each Iron Mountain customer has a set contract, laid out at the beginning of the service agreement that explains all costs associated with recalling and transferring data back to the organization. There are no third-party considerations (such as bandwidth providers) or impact on existing operations associated with data retrieval. While important for any data recall, this is mission critical in crisis mode operation during a disaster recovery.

If time is not the highest of importance, a request can be made to bring back any number of tapes on the next scheduled delivery at no additional cost to the organization. If data is needed sooner than the normal service will accommodate, Iron Mountain offers two expedited services for customers. For $130 a delivery can be scheduled for up to 1 truck load of tapes, containing tens of petabytes of data, to be delivered within 24 hours. For the Spectra Logic example, above, the entire data set could easily be retrieved in a single trip. In the event that data needs to be accessed sooner, a critical trip can be scheduled at a cost of $170, and a truck load of tapes can be delivered within 3 hours.

Large Data Set Retrieval Costs

There is no question that a single file or small data set restoration is significantly less expensive and very simple using cloud storage. As this paper points out in opening statements, that's a great fit for the cloud. It's the restoration of large data sets that break the cloud model. Looking at the Spectra Logic example mentioned, let's examine the impact of a disaster recovery 33 months into the model. Spectra would have accumulated 360TB of data. Having experienced a major disaster in which all data was lost, they would need to recall the full 360TB as quickly as possible.

Amazon Glacier Restoration

Data set size in TB	Data set size in GB	Cost to retrieve data per GB	Total cost to retrieve data
360	360,000	$0.05	$18,000.00

It would cost $18,000 to recall the full 360TB. More importantly, it would take a total of 233 days to recall the data based on the 150Mbps network connection currently in place. For an organization that needs their data immediately, waiting for over 200 days would most likely drive them out of business. An option is to increase network bandwidth, but this is very costly, and not easily done. When dealing with bandwidth contracts and terms of service, it is impractical if not impossible to significantly increase performance for a few weeks.

Iron Mountain Restoration

Trip Type	Deliver Time	Data Retrieved	Total cost to retrieve data
Critical Trip	3 Hours	360TB	$170.00

It would only cost Spectra Logic $170 to do, via Iron Mountain, the full 360TB restoration, and they would have access to their tapes within 3 hours where they can begin accessing their data. With two LTO-8 drives that can read data at 360MB/sec it would take less than 6 days to restore all the lost data. No other changes, charges, contracts, or unexpected bills would be associated with their retrieval of data.

Still Headed to the Cloud?

As stated earlier in this paper, there are many reasons to utilize cloud services. Although the cost for long-term storage, and the bandwidth required for disaster recovery of that storage, don't seem to be competitive from a cost perspective, there are still mandates by some organizations to, "put everything in the cloud."

When archiving to the cloud, there is still a simple way to make the experience less costly, more predictable for budgeting and assure your organization is not "locked in" to a given cloud provider – the often-unfortunate experience of "vendor lock-in." Topping the list of *cloud best practices* is to keep a copy of the data locally. Even a small tape library, with tapes ejected to onsite vault storage, will offer multiple advantages for minimal cost.

First and foremost, maintaining a local copy of data allows organizations to switch cloud vendors should their prices rise, quality of support fall, or execution of SLAs be missed. After years of a "trickle feed" approach to slowly moving data to the cloud, most organizations are not prepared to bring the data back. The performance figures just mentioned make that clear – showing a wholly dedicated 150 Mbps download requiring 233 days to bring back 360TB of data. While a full disaster seems unlikely you're willing to take that risk, is it really unlikely that you will never change cloud providers? The time to retrieve data and the cost to retrieve data remain the same regardless of the reason it is brought back, and it may well keep an organization locked into the first cloud provider they worked with regardless of service. By keeping a local copy of data, cloud data sets may be deleted when a contract expires with no additional cost for restore or bandwidth. Simply move the local data set to the new cloud provider. It's also a compelling insurance policy should the cloud provider go out of business or suffer their own catastrophic loss via natural disaster or cyberattack.

Cyberattack is the second reason for keeping a local copy of data on tape. Cyberattacks are becoming more and more common. From "ransomware" that encrypts disk-based data and demands payment, to international attacks on commercial organizations such as Sony Pictures, any data resting on disk that is accessible is vulnerable. As many cloud facilities store data on inexpensive disk, this data is not truly offline from an attack. With tape, an 'air gap' can be created ensuring that the media is not attached to the network, which prevents the data from being hacked. Tape is only online when the tape cartridge is mounted in the drive. When tape is stored offsite, it is fundamentally secure from cybercrime. In addition, the "genetic diversity" offered by having two wholly different types of media protecting data gives better assurance that it will survive such attacks. Malware aimed at disk storage cannot penetrate tape storage.

Finally, we come full circle to cost. It is easy, fast and inexpensive to restore single files or small data sets from the cloud. If that is the majority of restores, users are good the majority of the time. When users *do* find that you need a large restoration, what does that unexpected cost do to your budget? That's a good time to rely on a simple, quick restoration from onsite tape.

One might argue that the cost of Amazon's services could come down over five years. That is certainly likely. However, the media cost of the tape library system will also come down over the five-year period, this has not been reflected in our models. Furthermore, over the next five years the user could upgrade to support LTO-8, LTO-9 and possibly LTO-10 to gain ever higher throughput and capacity at lower media costs.

Next Steps?

The intention of this paper is to show the major differences between archiving to the cloud versus archiving to a remote repository. It would be impossible for one paper to address the needs of all organizations. The calculations performed for this paper are very straightforward. Specific data centers, content repositories, data sets and client needs are rarely so straightforward. Spectra offers onsite consultation with extremely detailed calculators and site analysis to give a dollar-for-dollar, service-for-service analysis of your specific data center. Our Solutions Architects will compare the cloud vendor of your choice with the offsite repository (or onsite vault) of your choice to provide exact costs and SLAs for your specific environment. This complimentary service can be scheduled through your Spectra Sales Associate or Spectra Professional Services Associate at your convenience.

About Spectra Logic

Spectra Logic develops data storage and data management solutions that solve the problem of long-term digital preservation for organizations dealing with exponential data growth. Dedicated solely to storage innovation for 40 years, Spectra Logic's uncompromising product and customer focus is proven by the adoption of its solutions by leaders in multiple industries globally. Spectra enables affordable, multi-decade data storage and access by creating new methods of managing information in all forms of storage—including archive, backup, cold storage, private cloud and public cloud. To learn more, visit www.SpectraLogic.com.

303-449-6400 • 800-833-1132 • 6285 Lookout Road • Boulder, CO 80301 USA • spectralogic.com

Appendix A

Month	Year	Hardware	Media	Support	Capacity in TB	Number of tapes	Number of Containers	Cost for containers	Weekly pickups	Yearly Total for Iron Mountain	Cumulative Iron Mountain Total	Cloud cost per GB	Storage Costs	Internet Costs	Yearly Total for Glacier	Cumulative Cloud Total
Jan-19	2019	8,230.00	4,824.60	-	200	23	3	30	185	13,269.60	13,269.60	0.0046	920	259.9	1179.9	1179.9
Feb-19	2019				205	23	3	30	185	215.00	13,484.60	0.0046	943	259.9	1202.9	2382.8
Mar-19	2019				210	24	3	30	185	215.00	13,699.60	0.0046	966	259.9	1225.9	3608.7
Apr-19	2019				215	24	3	30	185	215.00	13,914.60	0.0046	989	259.9	1248.9	4857.6
May-19	2019				220	25	3	30	185	215.00	14,129.60	0.0046	1012	259.9	1271.9	6129.5
Jun-19	2019				225	25	3	30	185	215.00	14,344.60	0.0046	1035	259.9	1294.9	7424.4
Jul-19	2019				230	26	3	30	185	215.00	14,559.60	0.0046	1058	259.9	1317.9	8742.3
Aug-19	2019				235	27	3	30	185	215.00	14,774.60	0.0046	1081	259.9	1340.9	10083.2
Sep-19	2019				240	27	3	30	185	215.00	14,989.60	0.0046	1104	259.9	1363.9	11447.1
Oct-19	2019				245	28	3	30	185	215.00	15,204.60	0.0046	1127	259.9	1386.9	12834
Nov-19	2019				250	28	3	30	185	215.00	15,419.60	0.0046	1150	259.9	1409.9	14243.9
Dec-19	2019				255	29	3	30	185	215.00	15,634.60	0.0046	1173	259.9	1432.9	15676.8
	2019 Total									15,634.60					15,676.80	
Jan-20	2018	-	1,527.15	1,539.84	260	29	3	30	185	3,281.99	18,916.59	0.0046	1196	259.9	1,455.90	17132.7
Feb-20	2018				265	30	3	30	185	215.00	19,131.59	0.0046	1219	259.9	1,478.90	18611.6
Mar-20	2018				270	30	3	30	185	215.00	19,346.59	0.0046	1242	259.9	1,501.90	20113.5
Apr-20	2018				275	31	4	40	185	225.00	19,571.59	0.0046	1265	259.9	1,524.90	21638.4
May-20	2018				280	32	4	40	185	225.00	19,796.59	0.0046	1288	259.9	1,547.90	23186.3
Jun-20	2018				285	32	4	40	185	225.00	20,021.59	0.0046	1311	259.9	1,570.90	24757.2
Jul-20	2018				290	33	4	40	185	225.00	20,246.59	0.0046	1334	259.9	1,593.90	26351.1
Aug-20	2018				295	33	4	40	185	225.00	20,471.59	0.0046	1357	259.9	1,616.90	27968
Sep-20	2018				300	34	4	40	185	225.00	20,696.59	0.0046	1380	259.9	1,639.90	29607.9
Oct-20	2018				305	34	4	40	185	225.00	20,921.59	0.0046	1403	259.9	1,662.90	31270.8
Nov-20	2018				310	35	4	40	185	225.00	21,146.59	0.0046	1426	259.9	1,685.90	32956.7
Dec-20	2018				315	35	4	40	185	225.00	21,371.59	0.0046	1449	259.9	1,708.90	34665.6
	2020 Total									5,736.99					18,988.80	
Jan-21	2019		1,527.15	1,539.84	320	36	4	40	185	3,291.99	24,663.58	0.0046	1472	259.9	1,731.90	36397.5
Feb-21	2019				325	37	4	40	185	225.00	24,888.58	0.0046	1495	259.9	1,754.90	38152.4
Mar-21	2019				330	37	4	40	185	225.00	25,113.58	0.0046	1518	259.9	1,777.90	39930.3
Apr-21	2019				335	38	4	40	185	225.00	25,338.58	0.0046	1541	259.9	1,800.90	41731.2
May-21	2019				340	38	4	40	185	225.00	25,563.58	0.0046	1564	259.9	1,823.90	43555.1
Jun-21	2019				345	39	4	40	185	225.00	25,788.58	0.0046	1587	259.9	1,846.90	45402
Jul-21	2019				350	39	4	40	185	225.00	26,013.58	0.0046	1610	259.9	1,869.90	47271.9
Aug-21	2019				355	40	4	40	185	225.00	26,238.58	0.0046	1633	259.9	1,892.90	49164.8
Sep-21	2019				360	40	4	40	185	225.00	26,463.58	0.0046	1656	259.9	1,915.90	51080.7
Oct-21	2019				365	41	5	50	185	235.00	26,698.58	0.0046	1679	259.9	1,938.90	53019.6
Nov-21	2019				370	42	5	50	185	235.00	26,933.58	0.0046	1702	259.9	1,961.90	54981.5
Dec-21	2019				375	42	5	50	185	235.00	27,168.58	0.0046	1725	259.9	1,984.90	56966.4
	2021 Total									5,796.99					22,300.80	
Jan-22	2020		1,527.15	1,539.84	380	43	5	50	185	3,301.99	30,470.57	0.0046	1748	259.9	2,007.90	58974.3
Feb-22	2020				385	43	5	50	185	235.00	30,705.57	0.0046	1771	259.9	2,030.90	61005.2
Mar-22	2020				390	44	5	50	185	235.00	30,940.57	0.0046	1794	259.9	2,053.90	63059.1
Apr-22	2020				395	44	5	50	185	235.00	31,175.57	0.0046	1817	259.9	2,076.90	65136
May-22	2020				400	45	5	50	185	235.00	31,410.57	0.0046	1840	259.9	2,099.90	67235.9
Jun-22	2020				405	45	5	50	185	235.00	31,645.57	0.0046	1863	259.9	2,122.90	69358.8
Jul-22	2020				410	46	5	50	185	235.00	31,880.57	0.0046	1886	259.9	2,145.90	71504.7
Aug-22	2020				415	47	5	50	185	235.00	32,115.57	0.0046	1909	259.9	2,168.90	73673.6
Sep-22	2020				420	47	5	50	185	235.00	32,350.57	0.0046	1932	259.9	2,191.90	75865.5
Oct-22	2020				425	48	5	50	185	235.00	32,585.57	0.0046	1955	259.9	2,214.90	78080.4
Nov-22	2020				430	48	5	50	185	235.00	32,820.57	0.0046	1978	259.9	2,237.90	80318.3
Dec-22	2020				435	49	5	50	185	235.00	33,055.57	0.0046	2001	259.9	2,260.90	82579.2
	2022 Total									5,886.99					25,612.80	
Jan-23	2021		1,527.15	1,539.84	440	49	5	50	185	3,301.99	36,357.56	0.0046	2024	259.9	2,283.90	84863.1
Feb-23	2021				445	50	5	50	185	235.00	36,592.56	0.0046	2047	259.9	2,306.90	87170
Mar-23	2021				450	50	5	50	185	235.00	36,827.56	0.0046	2070	259.9	2,329.90	89499.9
Apr-23	2021				455	51	6	60	185	245.00	37,072.56	0.0046	2093	259.9	2,352.90	91852.8
May-23	2021				460	52	6	60	185	245.00	37,317.56	0.0046	2116	259.9	2,375.90	94228.7
Jun-23	2021				465	52	6	60	185	245.00	37,562.56	0.0046	2139	259.9	2,398.90	96627.6
Jul-23	2021				470	53	6	60	185	245.00	37,807.56	0.0046	2162	259.9	2,421.90	99049.5
Aug-23	2021				475	53	6	60	185	245.00	38,052.56	0.0046	2185	259.9	2,444.90	101494.4
Sep-23	2021				480	54	6	60	185	245.00	38,297.56	0.0046	2208	259.9	2,467.90	103962.3
Oct-23	2021				485	54	6	60	185	245.00	38,542.56	0.0046	2231	259.9	2,490.90	106453.2
Nov-23	2021				490	55	6	60	185	245.00	38,787.56	0.0046	2254	259.9	2,513.90	108967.1
Dec-23	2021				495	55	6	60	185	245.00	39,032.56	0.0046	2277	259.9	2,536.90	111504
	2023 Total									5,976.99					28,924.80	
	Grand Total									39,032.56					111,504.00	