

Title: Tape in the Cloud
Magnetic Tapes Changing Role in support of Public and Private Clouds

By Greg Schulz

Founder and Senior Analyst, the StorageIO Group
Author “The Green and Virtual Data Center” (CRC)



November 28, 2009

This Industry Trends and Perspectives Paper is Compliments of:
FUJIFILM
www.tapepower.com

FujiFilm TapePower - Tape in the Cloud

Introduction

Just what are cloud computing, cloud storage and other forms of public and private IT capabilities? How and where do they fit in with your current environment? Buzzwords include: enabling agility, flexibility, transparency, adaptability, doing more with less while reducing costs, boosting utilization and productivity.

IT Trends and Issues

Cloud computing and cloud storage are tiers of traditional IT resources with different performance, availability, capacity and economic service attributes that complement existing resources.

For some, clouds are all the rage, the answer to everything; at the other extremes there are cynics who see little to no value at all. However, the majority of respondents in on-going poll (refer to <http://storageioblog.com/?p=665>) want to understand more about where and when to use clouds as an enabling co-existing technology, technique or paradigm.

What Are IT Clouds?

Let's take a quick look at what IT clouds are, where they fit, what to look out for and how tape co-exists with clouds and can be an enabling technology. Like many IT professionals, you may already be using or leveraging cloud-based computing or storage techniques either as a product or service without realizing it.

Common cloud-based functions or services include:

- Remote or off-site backup, replication, vaulting or data copy
- Remote or off-site storage on-line, near-line or off-line
- Email and messaging services including social media networking (e.g. Web 2.0)
- Archive, fixed content and other reference or look-up data
- Website, blog, video, audio, photo and other rich media content hosting
- Application hosting (e.g. salesforce.com, concur expense, social media)
- Virtual server or virtual machine (VM) hosting (e.g. Amazon EC2, Acadia)
- General on-line storage or photo sharing (Google, HP Snapfish)

Does this mean that if backup, business continuance (BC) or disaster recovery (DR) or archive data is sent off-site to a storage or hosting facility that it has been sent to the cloud? Some say no unless the data were transmitted electronically to on-line disk at a service provider location leveraging programmatic interfaces and other cloudware. That might also be a product or services based definition. However, in theory, the concept is not that far off as clouds, in addition to being a product or service, are also a management philosophy or paradigm to do more with what you have without negatively impact service delivery.

Clarifying Cloud Confusion

Confusion exists as there are many different types of clouds including public and private, products and services, some that use familiar interfaces or protocols with others using different technologies. For example some MSP, SaaS or SSP provide internet based storage services presented as Network Attached Storage (NAS) using NFS or Windows CFS, file transfer protocol (FTP), Web HTTP, or via a programmatic interface among others.

FujiFilm TapePower - Tape in the Cloud

An example of a programmatic interface is a backup client installed on your computer or server that collects local data, perhaps deduplication and or compressing along with encrypting the information, before transmitting to a hosting site.

Other examples of cloud services include:

- Managed Service/Storage Provider (MSP)
- Storage as a Service (SaaS)
- Storage Service Provider (SSP)
- Archive as a Service (AaaS)
- Backup as a Service (BaaS)

Is Hosting or Co-Location a Cloud?
 External, both public and private, clouds can be housed at a hosting or co-location facility as well as at traditional outsourcing or managed service providers. Thus a hosting site may or may not be a cloud, and a cloud can leverage hosting services however does not require them.

Does this mean that data put into a public cloud is accessible or viewable by anyone? Generally, the answer is no, the public nature means that the service is offered to many different subscribers by a provider. For example (Figure 1) your data may be placed on the storage system and accessed via the same physical or virtual server used by other subscribers; however, unless placed into a public folder, your data should be logically isolated. For some service providers, value added capabilities at premium fees can be to physically isolate data for security, compliance or other purposes.

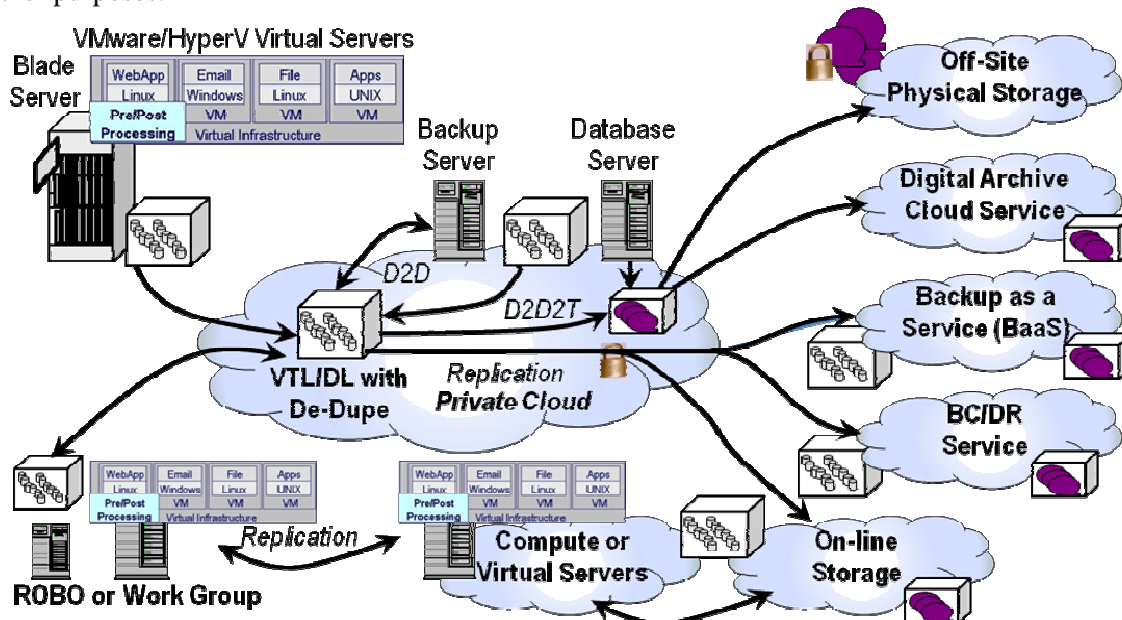


Figure 1: Public and Private Cloud Services and Solutions with tape co-existence

Do clouds have to be physically off-site hosted by a third party? No, clouds can be implemented internally known as a private cloud on-site at your premises using existing technologies.

Clouds can be a service or a product, architecture or a management paradigm, similar to previous generations such as the information utility or service oriented architectures (SOA), client server computing, among others. What this means is that some of the tenants of cloud storage and computing involve shifting how resources are used and managed thus enabling the notion of an information factory.

FujiFilm TapePower - Tape in the Cloud

Enabling and Protecting Information Factories

Using the information factory analogy means that like a physical factory, some work is done on premise and some off-site at other locations including those of subcontractors or suppliers.

In the case of information factories, the product being produced is information services with the machinery being servers, storage and I/O networking managed with software, processes, procedures and metrics. Raw materials include data, energy to power and cool the physical facility, and technologies all operating to deliver the services at a low defect or error rate while meeting or exceeding Quality of Service (QoS), performance, availability and accessibility requirements in a cost effective manner.

For some cloud compute or storage providers, the value proposition is that they can supply the service at a lower cost than if using your own capabilities. Similar to service bureaus, outsourcing, managed service or hosting facilities of the past, cloud-based services are a means of shifting or avoiding costs by moving work or data elsewhere to be processed or stored.

However, clouds should be looked at not just for their cost saving abilities as this is similar to the fallacy of evaluating storage on a cost per capacity basis while ignoring performance, availability, data integrity, ease of management and other items that can impact service delivery and expenses. Consequently, clouds should be looked at as not a replacement or competing technology or technique, rather, as a complimentary approach to existing in-house resources.

Local or Remote Clouds

Cloud resources can be local or remote, public or private depending upon the service, product or functionality along with management objectives you are pursuing for your environment.

Cloud computing and storage are simply additional tiers of servers and data repositories that may have different performance, availability, capacity or economics associated with them to meet specific business and/or application needs. That is, cloud computing and cloud storage co-exist and compliment what is currently being done with the objective of boosting quality of service, availability or customer satisfaction while supporting more data being processed, moved and stored for longer periods of time at a lower unit cost.

Tapes and IT Clouds = Co-existence

Given tape's ability to store large amounts of data in a dense, energy efficient manner at a low cost per Gbyte, it is the perfect complement to disk-based cloud services. The role of tape is changing and, with disk, it is being used more frequently as a buffer, cache or staging area supporting disk to disk (D2D) backups, snapshots, replication and archiving.

One advantage of staging data to disk is the ability to stream data to tape devices more efficiently, resulting in higher utilization of devices and media, thereby reducing data protection costs. Another advantage of staging data to disk is the ability to keep more copies of recent backups on-line for rapid individual file restore while leveraging tape for large scale or bulk recovery and portability.

Consequently, given a common value proposition of cloud-based services of reduced cost by increased utilization of resources, tape remains the perfect complement to disk for local and remote data protection.

FujiFilm TapePower - Tape in the Cloud

There is more network bandwidth available than ever before to allow data to be moved electronically between different locations. The bad news is that there is also more data than ever before to be moved in a shorter amount of time. Granted, there are bandwidth and data footprint optimization techniques available, but one of the challenges of moving data to a cloud-based provider is the initial migration or copy.

The solution is for cloud providers to offer bulk removable media data movement for initial copy as well as for large scale recovery operations. For example, a local virtual tape library could, instead of replicating a large amount of information which takes time even with deduped data and optimized networks, be used to leverage path to tape (PTT) or copy to tape functions.

Once the tapes are created, they would be physically transported to a cloud service provider and restored or ingested into their systems. From that point, data would be synchronized to the current point in time to comply with RTO and RPO requirements. Individual files or smaller restorations could be performed over the network, however, for large scale restoration either for BC/DR or reconstituting archive data for a project, data could be returned to primary site via removable tape media with any subsequent changes replicated and synchronized electronically.

Another use for tape in a cloud environment is in the background, similar to a traditional site where data resides on-line on disk but to free up space, information is migrated to tape. Likewise, for data protection, an extra copy of data is spun off to tape for safe keeping, not all that different than what you might be doing today, leveraging your own in-house or external backup as a service solution.

Some other things that can be done includes leveraging disk can be used as a buffer or cache by staging to disk so that tape drives can be used more efficiently as tapes are better utilized in streaming modes.

Footprint reduction techniques should used on tape and disk - compress and/or dedupe data to enable more backup copies to be kept on-line for faster restore, and leverage tape for master copy or to facilitate rapid large scale restoration as needed. Tape can be onsite or data can be electronically transmitted to a cloud or managed service provider who then off-loads a copy to tape, perhaps with a copy also kept on-line.

Public vs. Private Cloud

A public cloud has shared resources or services with a value proposition of being hosted off-site at a lower cost vs. legacy IT resources. Cost savings are achieved by increasing utilization of shared resource capacity. Private clouds can be on-site or off premise using existing along with new technologies being operated in a more efficient manner, both from a capacity or space utilization along with availability, security and performance manner.

Tapes Future in the Clouds

Not to worry, it is not storm clouds in store for magnetic tape, or at least in the foreseeable future

Despite continued hype about tapes demise, the venerable data storage medium is finding new roles as an ultra low-cost long term medium for data preservation and retention.

Where and how tape is used or located can expect to continue to change along with being paired with complementary technology including disk to disk (D2D) backup, VTLs with replication and dedupe.

FujiFilm TapePower - Tape in the Cloud

Look before you leap!

It is important to understand not only what type of cloud computing or storage service, product, architecture or management paradigm you may be looking at along but why you are considering it. The why may be that cost is a driver, yet if performance, availability, capacity or other service characters are negatively impacted, “cheaper” may turn out to be more expensive.

In the case of moving primary data into a public or private cloud, make sure that the information is protected under your current BC and DR plans. Look for providers that offer service level agreements (SLAs) with commitments to availability and accessibility along with data integrity, security and privacy. What this means is that any data you send to that service will be retrievable and accessible when you need it or within committed response times and the data will be safe and secure.

At the bottom line, if you are putting data into any cloud, have a backup or a copy elsewhere. Likewise, if you have local or even remote data, look at using a cloud or managed service provider as a means of parking another copy of backups or archives. After all, any information worth keeping should have multiple copies on different media in various venues.

Where to read and learn more:

Clouds and Data Loss: Time for CDP (Commonsense Data Protection)

<http://storageioblog.com/?p=704>

Clouds are like Electricity: Dont be Scared

<http://storageioblog.com/?p=657>

About the author

Greg Schulz is founder of Server and StorageIO, an IT industry analyst consultancy firm and author of the books *The Green and Virtual Data Center* (CRC) and *Resilient Storage Network* (Elsevier). Learn more at www.serverandstorageio.com or on twitter @storageio.

All trademarks are the property of their respective companies and owners. The Server and StorageIO (StorageIO) Group makes no expressed or implied warranties in this document relating to the use or operation of the products and techniques described herein. StorageIO in no event shall be liable for any indirect, inconsequential, special, incidental or other damages arising out of or associated with any aspect of this document, its use, reliance upon the information, recommendations, or inadvertent errors contained herein. Information, opinions and recommendations made by StorageIO are based upon public information believed to be accurate, reliable, and subject to change. This industry trends and perspective white paper is compliments of FujiFilm TapePower www.tapepower.com