



Research Report

Moving System z Data to the Cloud: A State-of-the-Art Solution for Scalable Capacity, Back-up/Recovery and Archiving

Executive Summary

Cloud computing has changed the way that enterprises look at their computing infrastructure. Enterprises now realize that cloud computing can be used to help drive innovation and grow the business, but enterprises also recognize that they need better ways to manage and protect data that is being deployed across hybrid cloud infrastructures.

IBM's System z is a key participant in modern hybrid cloud architecture because it processes critical business data that drives enterprise applications. Further, it provides the scalability, resilience and security needed to support mission critical applications. For over 40 years, mainframe managers have been virtualizing mainframe resources, achieving extremely high levels of utilization on these enterprise server platforms. *Storage, on the other hand, has been slower to evolve.*

Information technology (IT) executives are now actively considering a move from storing data on traditional Direct Access Storage Devices (DASD) to cloud-based storage. Using cloud-based storage, IT organizations can accelerate business responsiveness using the cloud to efficiently store, access, retrieve and recover z/OS data. New storage cloud innovations can help lower operating costs, further improve security and reduce administrative overhead associated with managing and scaling System z storage.

From a disaster recovery (DR) perspective, cloud-based solutions for DR and backup can enable the recovery of z/OS data faster than using traditional solutions such as tape back-up (especially if the data is off line in a physical vault). Further, cloud-based solutions can help improve reliability and decrease recovery times. Cloud storage solutions also help prevent “over-purchasing” of storage – enabling IT organizations to capitalize on “pay-as-you-go” pricing to grow and scale storage capacity quickly and cost-effectively.

In this *Research Report*, *Clabby Analytics* takes a closer look at how cloud-based storage solutions can significantly reduce the costs associated with data availability and protection both on-premise Network Attached Storage (NAS) and off-premise (in the cloud). We provide one example of such a solution: CA Technologies' *CA Cloud Storage for System z*.

Market Background

In pre-cloud days, storage was all about adding capacity through additional hardware acquisition. Businesses added new storage arrays for new applications. Departments purchased their own dedicated storage for increased control over departmental data. Storage vendors oversold and customers overbought, gearing up for expected growth with ready access to additional capacity to ensure scalability and performance. Disaster recovery plans called for mirroring data in multiple locations -- doubling and sometimes

Moving Mainframe Data to the Cloud: A State-of-the-Art Solution for Scalable Capacity, Back-up/Recovery, and Archiving

tripling the cost of storage. A lengthy storage acquisition process prompted businesses to buy based on projected growth rather than actual growth, so newly acquired storage was underutilized in many cases. Paying hundreds of dollars per square foot and even more for power and cooling caused operational costs to soar. As departments made individual storage acquisition decisions, soon the enterprise was rife with different types of storage purchased from a range of vendors, creating huge management challenges while increasing administrative costs.

How Storage Management Innovation Has Driven Down Storage Costs

Several technology advances have contributed to driving down storage costs including thin provisioning (applications consume only the capacity they are using); storage tiering (data is stored on different classes of storage {disk, SSD, tape} based on price/performance requirements; deduplication (eliminates redundant blocks/files of data); and compression (eliminates redundancy within blocks or files). Further, the movement to “software defined” storage (which enables several of the features above to be implemented in software and thus be shared among a range of hardware devices) has also improved efficiency and utilization of storage. But equally as significant: the advent of cloud computing, and the emergence of cloud-based storage and new storage delivery models have also helped to drive storage costs downward.

Cloud-based online file storage was first introduced in 2006. Today, public cloud providers including Amazon Web Services (AWS), Microsoft Azure, Google Cloud, IBM SoftLayer and others offer cloud-based online file storage. These storage vendors have implemented aggressive “pay-as-you-go/pay-as-you-grow” pricing models with no minimum fee. In addition, low-cost cloud archive storage services are available for data archiving, disaster recovery and online backup, optimized for data that is infrequently accessed and for which retrieval times of several hours are suitable. One vendor, for example can provide this capability for as little as \$0.01 per gigabyte per month, a significant savings compared to on-premise solutions.

Introducing Low-cost Cloud Storage for the Mainframe

Until recently, there has been little integration between the mainframe and cloud storage environments. But the benefits of cloud computing have become too compelling to ignore. Deploying cloud-based solutions can result in significant reductions in capital expenses (CapEx), as well as operating expenses (OpEx). In fact, Forrester Research calculates up to a 74% savings resulting from the use of cloud-based file storage as compared to on-premise storage. Meanwhile, Gartner reports that 80% of CIO’s believe the mainframe is the essential to their cloud infrastructure.

So logically, if mainframes are essential members of enterprise hybrid clouds – and big money can be saved using cloud storage – then mainframes should be using cloud storage.

The Mainframe and the Cloud Storage Model

Wouldn’t it be great if mainframe data could easily be stored in the cloud – thus providing storage capacity as needed, while offering a secure and cost-effective back-up, recovery and archiving solution? And wouldn’t it be even better if storing mainframe data in the

Moving Mainframe Data to the Cloud: A State-of-the-Art Solution for Scalable Capacity, Back-up/Recovery, and Archiving

cloud cost less than on-premise DASD solutions and on-premise/off-premise tape solutions? The good news is that products are available today that can bridge the gap between mainframe legacy systems which house the majority of “systems of record” (financials, manufacturing, CRM, HR), and the cloud—with its cost, security, scalability and reliability advantages. By combining software on System z with data from System z, a public cloud gateway (such as Riverbed SteelStore) and a public cloud platform (Amazon Web Services, IBM SoftLayer, Microsoft Azure or the Google Cloud) or a file-based storage subsystem (such as NetApp or Data Domain), mainframe users can significantly reduce storage expenses, while improving data security, scalability and availability.

By acquiring storage “as needed” through the cloud, taking advantage of storage features including deduplication, compression and encryption, as well as using cloud or file-based storage backup and recovery, data center storage costs can drop to pennies per GB and data protection costs can be reduced by up to 80%

The benefits of using cloud solutions for mainframe data include:

- 100% tape free (meaning no need for tape on site or for managing the placement and retrieval of physical tape off site in a secured vault);
- Flexible configuration options are available to support both small and large businesses with different capacity requirements
- By moving archival and back-up data to an offsite cloud storage environment, more storage capacity is available on-premise for mission-critical data
- The process of storing data on local NAS in the cloud has a low impact on mainframe CPU overhead;
- By leveraging industry leading NAS devices an extra layer of user-controlled security can be added to mainframe data prior to backing up or archiving to a cloud provider
- Cloud/NAS data storage represents a cost effective way to scale storage, allowing enterprises to take advantage of “pay-as-you go/pay-as-you-grow” pricing;
- Reduces the need for geo-replication and the need to transport physical media (tape) off site; and finally,
- Lower Total Cost of Ownership (TCO) when using a tapeless approach combined with cloud technologies

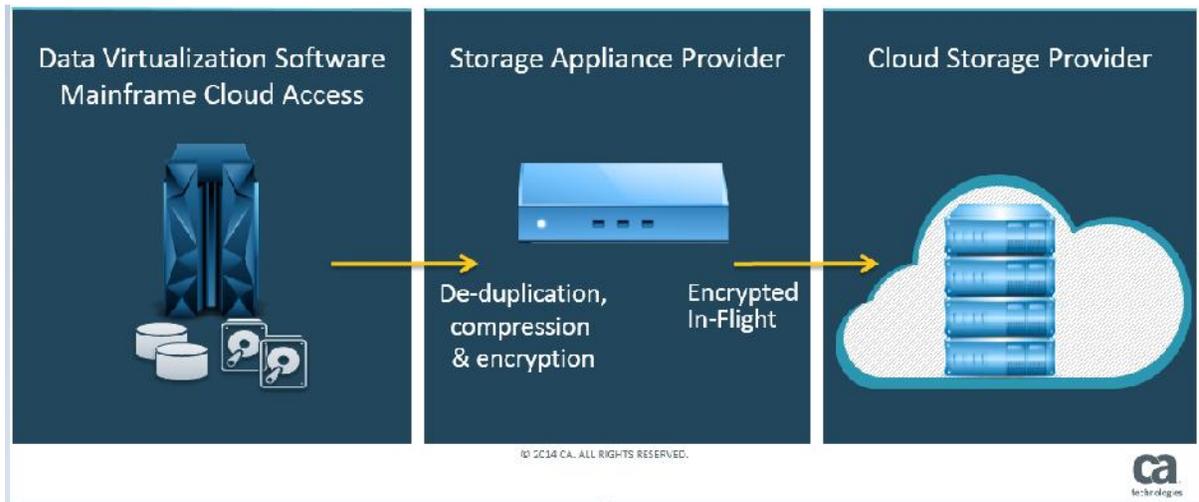
A Closer Look at CA Technologies’ CA Cloud Storage for System z

CA Cloud Storage for System z has been designed to simplify the management of back-up, recovery and archival storage, while improving storage availability and security – and significantly lowering storage costs.

CA Cloud Storage for System z is a cloud-enabled enterprise storage solution designed to help organizations eliminate costly purchases and reduce the TCO for storing and managing z/OS data. This solution stores data in the cloud and/or in NAS that appears as a tape subsystem. Using this solution, the mainframe operates as if the data is copied to tape – when, instead, data is copied to low-cost cloud storage or on locally available NAS storage (see Figure 1, next page).

Moving Mainframe Data to the Cloud: A State-of-the-Art Solution for Scalable Capacity, Back-up/Recovery, and Archiving

Figure 1- CA Cloud Storage for System z



Source: CA Technologies, November 2014

As a result of using a cloud storage approach, IT buyers can expect to achieve TCO savings that can approach 50% as compared with tape subsystems and associated management/transportation/real estate footprint expenses.

CA Cloud Storage for System z is comprised of 3 key components:

1. *Cloud Storage for System z* is a patent pending cloud storage solution providing for virtual tape data backup from the mainframe to the cloud using tape virtualization and cloud enabling software. In addition, the solution provides a gateway to open systems from System z. Using a combination of z/OS and Linux on System z, virtual tape images are offloaded to a NAS or storage cloud through a TCP/IP channel-to-channel (CTC) connection, ensuring data is backed up with minimal impact on the mainframe CPU.
2. A *Storage Appliance* that acts as a storage gateway or NAS storage subsystem accepts data from CA Cloud Storage for System z and (depending on the capabilities of the device) can perform deduplication, file compression and encryption. These devices typically have an intelligent local cache for immediate access to recent and frequently accessed data, eliminating the need to wait for restores from traditional archives while also replicating data to the cloud or NAS device. Riverbed SteelStore, EMC Data Domain, and NetApp are supported today while others will be added in the future.
3. A *Cloud Storage Provider* offers a file storage web service for low-cost data archiving and back-up – used for infrequently accessed data with suitable retrieval times of several hours. Cloud Storage Providers offer cloud services that provide access to low cost, reliable storage for back-up and archiving, data storage “bursting” (augmenting on-premise storage with storage capacity in the cloud on an as-needed basis), and disaster recovery. The solution works with Amazon (S3, Glacier), Microsoft Azure, Google Cloud and IBM SoftLayer.

Moving Mainframe Data to the Cloud: A State-of-the-Art Solution for Scalable Capacity. Back-up/Recovery, and Archiving

The Motivations/Cost Discussion

There are four scenarios where it makes great sense to move from DASD- or internal NAS-based storage to cloud-based storage. They are:

1. If an organization sends physical tape off-site – but would like to enhance its DR practices to include tape and eliminate manual recovery;
2. If there is a need to store a lot of data long term – and if there is a need to clean-up and compress DFHSM (Data Facility Hierarchical Storage Manager) or FDR (fast, dump, restore) data;
3. If there is a need to retain data backups for long periods of time (such as to meet compliance requirements);
4. If there is a need to retain unreferenced tape for long periods of time.

In short, if your organization has a lot of data that is sent off premise, it can greatly reduce physical tape management costs; it can do a better job of compression and thus end up using less disk; it can meet compliance and long-term storage requirements for decades to come – and it can improve the DR process with faster access to locally stored data with assurance that back-up copies can be accessed quickly in the cloud.

The cost per gigabyte of stored data when leveraging CA's Cloud Storage product can be as low as a penny, provided volume discounts are applied. The low cost is also achieved because of better compression algorithms which reduce storage capacity requirements. Figure 2 illustrates some of the compression advantages seen with CA's Cloud Storage for System z.

Figure 2 – Cost and Compression Advantages with Cloud Storage for System z

Cloud Cost (GB per month)	\$.01
Deduplication Ratios	
-- Uncompressible Data (like DB2)	3
-- Workfiles	4
-- Backups	7
-- Archives	5

Source: CA Technologies

CA Technologies showed us two case studies that illustrate how low cloud storage pricing can be (but note, prices can vary depending on the volume of data, type of data, and service level requirements):

- A small holding company that regulates electric and natural gas as well as other energy- related businesses, had a corporate-wide initiative to move relevant functions to the cloud. This organization was using physical tape, manual transportation and offsite vaulting, with 150 TB of data, no replication and a peak

***Moving Mainframe Data to the Cloud: A State-of-the-Art Solution for Scalable Capacity.
Back-up/Recovery, and Archiving***

throughput of 262GB/hour at a cost of \$0.12/GB for a 5-year TCO of \$1, 116,892. With smaller data storage/transfer requirements, the small energy company could deploy 1 IFL running Linux on System z, a Riverbed Model 2030 SteelStore storage appliance with no additional shelves and cloud storage nets out to a cost of \$0.07/GB and a five year TCO of \$610,059 for a **45% cost savings of \$506,833**.

- A global research, manufacturing and distributor of pharmaceuticals needed a cost-effective long term retention solution for files and data associated with rodent drug testing in order to comply with FDA regulations. Their current solution employed physical tape, manual transportation, and offsite vaulting for 1400TB with no replication and peak throughput of about 1600 GB/hour at a cost of \$0.16GB for a 5 year TCO of \$1,460,058. CA Cloud Storage for System Z solution including 2 IFL's running Linux on System z, a Riverbed Model 3030 SteelStore storage appliance and cloud storage nets out to a cost of \$.11/GB and a five year TCO of \$954,286 **for a 35% savings of \$505,772**.

Summary Observations

There many reasons why IT executives should consider the elimination of tape-based storage in favor of local NAS or cloud-based storage:

1. *The biggest reason is cost savings* (examples found in this report show 35-45% savings in storage costs); existing storage can be repurposed also cutting costs
2. *Risk reduction:* Using NAS/cloud storage means that tapes will no longer need to be physically transported to off-site vaulted storage. The risks of loss-in-transit, misfiling and theft are thus greatly reduced. Encryption reduces risk and ensures compliance with security regulations;
3. *Online accessibility:* Data that is stored in NAS or in the cloud remains on-line and accessible. This means that when the business needs access to aged, stored data, it is more easily accessed from the NAS/cloud than it is from a physical tape library;
4. *Enhanced capacity management:* Data sent to a low-cost, off premise cloud service provider can free up storage space within existing arrays. The volume of data collected and stored has grown exponentially over the past several years (thanks in large part to the addition of video, audio, mobile, and social media data and data analytics). Finding more storage within the enterprise can help forestall the need to purchase more capacity. By some estimates, enterprises may be able to recover up to 40% of their storage capacity by taking advantage of NAS/cloud storage.
5. *Multiple Layers of Encryption:* NAS/cloud stored data can actually be double-encrypted, in some cases. CA Technologies' CA Cloud Storage for System z encrypts data during the storage process – while Amazon Web Services also encrypts data that it stores

***Moving Mainframe Data to the Cloud: A State-of-the-Art Solution for Scalable Capacity.
Back-up/Recovery, and Archiving***

6. *Lower Cost CPU Usage:* CA Cloud Storage for System z runs on lower cost IFL CPUs in the mainframe (as compared with z/OS processors) – making it more economical to use mainframe MIPS for back-up, recovery and archiving.
7. *Flexibility:* CA Cloud Storage for System z supports AWS, IBM SoftLayer, Google Cloud, Microsoft Azure, in addition to other public cloud environments.

Enterprises looking to save Big Money in storage need to recognize that the cost for NAS/ cloud storage from cloud service providers is now so low that it actually costs less to use these technologies for archival, back-up and disaster recovery in place of tape. Enterprises that have made the move are finding that they are saving significant amounts of money using NAS/cloud storage – while at the same time improving Quality-of-Service through double encryption as well as by gaining more immediate access to cloud/NAS-based data as compared with accessing physical archived tape data.

Clabby Analytics
<http://www.clabbyanalytics.com>
Telephone: 001 (207) 239-1177

© 2014 Clabby Analytics
All rights reserved
November, 2014

Clabby Analytics is an independent technology research and analysis organization. Unlike many other research firms, we advocate certain positions – and encourage our readers to find counter opinions – then balance both points-of-view in order to decide on a course of action. Other research and analysis conducted by Clabby Analytics can be found at: www.ClabbyAnalytics.com.