DELLEMC

LONG TERM RETENTION TO THE CLOUD WITH DELL EMC CLOUDBOOST

Total Cost of Ownership and the Cloud

ABSTRACT

Tape has been the traditional backup medium for long-term retention. With the proliferation of cloud storage as an inexpensive storage tier, this traditional paradigm is quickly shifting. This white paper discusses the benefits of the Dell EMC® CloudBoost™ solution and a total cost of ownership comparison between Dell EMC ECS and public cloud options.

November, 2016

DELLEMC

WHITE PAPER

TABLE OF CONTENTS

EXECUTIVE SUMMARY

Organizations are finding themselves stuck in a quandary, with the need to maintain data for the entirety of its potentially lengthy lifecycle clashing with the simple truth of IT economics: do more with less. Fortunately, with advances in Storage as a Service (SaaS), customers are finally able to assume a proactive approach towards unbridled data growth by migrating content that otherwise would have been retained on tape for long-term retention to the public/private cloud.

Traditionally organizations have utilized tape backups as an integral part of their Business Continuity (BC) plan. As a cost-savings measure, tape infrastructure has persisted as the status quo that is "just good enough for backup", since adopting an alternative strategy would force:

- A restructuring of the current backup infrastructure

- Expensive retraining of IT admins on new backup workflows

- Bringing a new bill to management

With the advent of cost-effective disk-based backup with deduplication technology and the means for multi-site replication at nearzero recovery point objectives (RPO) and near instantaneous recovery time objectives (RTO), a compelling use case exists to store data with longer life-cycles on disk. The story continues to unfold with the rise of object-storage and its accompanying flat structure allowing for ease of management, scalability and replication across geographically distributed locations.

With the growth of options and a compelling use case for each, organizations are left wondering where they should be placing their data to most effectively meet today's business requirements, not to mention future industry mandates or upcoming government regulations. With traditional tape backups, organizations do not have to restructure or make changes to the established modus operandi. Disk-based backups already in place are serving the purpose of keeping data highly available for business continuity, but their access volume/frequency and economics may not be the best solution for long-term retention. Object-based storage in a public or private cloud is a viable solution, but there may be apprehension about bandwidth constraints, adapting to new TCO requirements, and the ability to store primary backups in the cloud without the benefit of local copies. With the rapid maturation of both of these technologies, organizations can now look to a new approach for addressing data requiring long term retention; namely, a hybrid deployment leveraging existing disk-based backups partnered with the durability, elasticity and economics of the cloud.

INTRODUCTION

This paper intends to explore the potential benefits of replacing tape-based backups for long-term retention with a hybrid model of on-site availability coupled with the flexibility of cloning data to the cloud using the Dell EMC CloudBoost solution. Other benefits explored are utilizing the CloudBoost solution for long-term retention with object storage, and backup service TCO comparisons between public cloud providers and Dell EMC Elastic Cloud Storage (ECS).

WHY CLOUDBOOST?

THE POWER OF CHOICE

The CloudBoost solution enables organizations to utilize existing NetWorker workflows to clone backups to a chosen public, private or hybrid cloud implementation for long term retention. The CloudBoost solution acts as a NetWorker storage node configured as an Advanced File Type Device (AFTD), allowing for NetWorker clone operations to the CloudBoost solution, after which the data is deduplicated, encrypted and sent to the object store. The CloudBoost solution is deployed as an Open Virtual Appliance (OVA) template within vSphere while cloud profiles and system monitoring are managed from a convenient web interface, Dell EMC Cloud Portal, which allows admins to manage one or many CloudBoosts from a single pane of glass.

The CloudBoost solution supports a number of public and private cloud providers, includes various mechanisms to ensure the security of data in-flight and at-rest, and addresses performance bottlenecks, data integrity and cloud abstraction.

CLOUD ABSTRACTION

A split-plane architecture decouples the metadata (control plane) from the data (data plane), enabling software-defined backups that meet business continuity requirements as well as the ability to select from a number of public and private cloud target options.

Additionally, the CloudBoost solution is purpose-built for cloud agnosticism, enabling one to grow with the incumbent object storage ecosystem, or a future supported vendor's implementation. Cloud agnosticism mitigates vendor lock-in, particularly in a market progressing so rapidly, where it may be necessary to switch providers for any number of reasons.

PERFORMANCE BOOST

The separation of data and metadata also leads to performance benefits, as metadata is cached on the CloudBoost for faster lookups while also removing metadata bottlenecks that would restrict faster data movement. Performance is improved as well with inline, variable length source-side deduplication ensuring that duplicate data is never sent over the wire; or conversely, no data is ever sent twice. This is a powerful property since only new and unique data is sent over the wire. To deal with the high latency of object store APIs, CloudBoost segments files into variable sized chunks averaging 256 KB, a size that strikes the optimal balance between maximizing the deduplication ratio (with small chunks) and minimizing metadata overhead (with large chunks).

SECURITY MATTERS

The CloudBoost solution is installed in the customer's local data center and encrypts all traffic, utilizing SSL for transmission to the object store and leveraging signed URLs with a brief expiration value, eliminating the common exploit in which transactions can be replayed. Each chunk is encrypted with its own independent key, utilizing the American Encryption Standard (AES) 256-bit encryption algorithm, a cypher that has no known exploit today.

ENSURING DATA INTEGRITY

To ensure that data has been properly written, CloudBoost Server verifies each chunk after it's written by the CloudBoost agent to validate that the hash for each chunk matches the hash provided by the agent.

WHY TCO MATTERS

When dealing with TCO and the multitude of public cloud offerings, it's important to consider other factors outside of storage costs alone.

EXISTING INFRASTRUCTURE COSTS

An organization's investment in existing backup infrastructure must be considered in determining TCO. A comprehensive understanding of how cloud may benefit your organization must take into account what is already in place and, more importantly, what is actively being used to fulfill organizational requirements as part of your business continuity strategy.

COMPLIANCE AS A COST

Depending on the type of organization, compliance is often a primary concern, not only in terms of the monetary cost of fines levied for being out of compliance, but also the damage to the organization's reputation. Large organizations are often concerned about securing their backups in a multi-tenant public cloud offering even if the multi-tenancy reflects only their lines of business (LOB).

EXIT COSTS

Vendor lock-in is a real concern when dealing with public cloud storage, so having the freedom to control where data is stored and retaining ultimate control over movement of that data can be the deal breaker for many organizations. Mission-critical data movement, particularly if the organization is an acquisitive one, is key to benefiting from new economies of scale in storing longterm retention data. This keeps the tipping point clear and ensures that whatever the new price-point it takes to reconsider a new cloud vendor, the organization can exercise it easily.

WAN COSTS AND CONSIDERATIONS

A dedicated WAN link may be needed in order to accommodate the original influx of backup data that must be sent over the wire daily data demands are likely to be lower than this first burst of "new" data. It must be determined how much time is needed to transmit new baseline backups to the cloud, comprised of daily rates plus some percentage of restore operations (likely less than 10%). WAN acceleration techniques may be utilized to mitigate some transmission delays, but the critical factor is the distance that data must travel.

RECOVERY COSTS

Backups stored in public cloud storage have a cost associated with them for restore operations. An aggressive RTO may also be a primary factor in considering overall TCO. An important consideration is how much downtime and resulting expense the organization is willing to tolerate while waiting for data restores from the public cloud.

TCO IN ACTION

It is often assumed that at least the more popular public cloud storage offerings are not only elastic, but cost efficient and agile as well. While this may be true in some cases, it's worth highlighting how real world numbers prove just where the elasticity breaks down. The chart below highlights the TCO of storing half a terabyte of data in three AWS regions with an Enterprise Support level contract and 5% of total capacity recovered from S3 buckets and 1% of total capacity retrieved from Glacier cold storage. All of the AWS-S3/Glacier costs are presented in Figure 1 below alongside the costs associated with a three site Elastic Cloud Storage (ECS) environment.
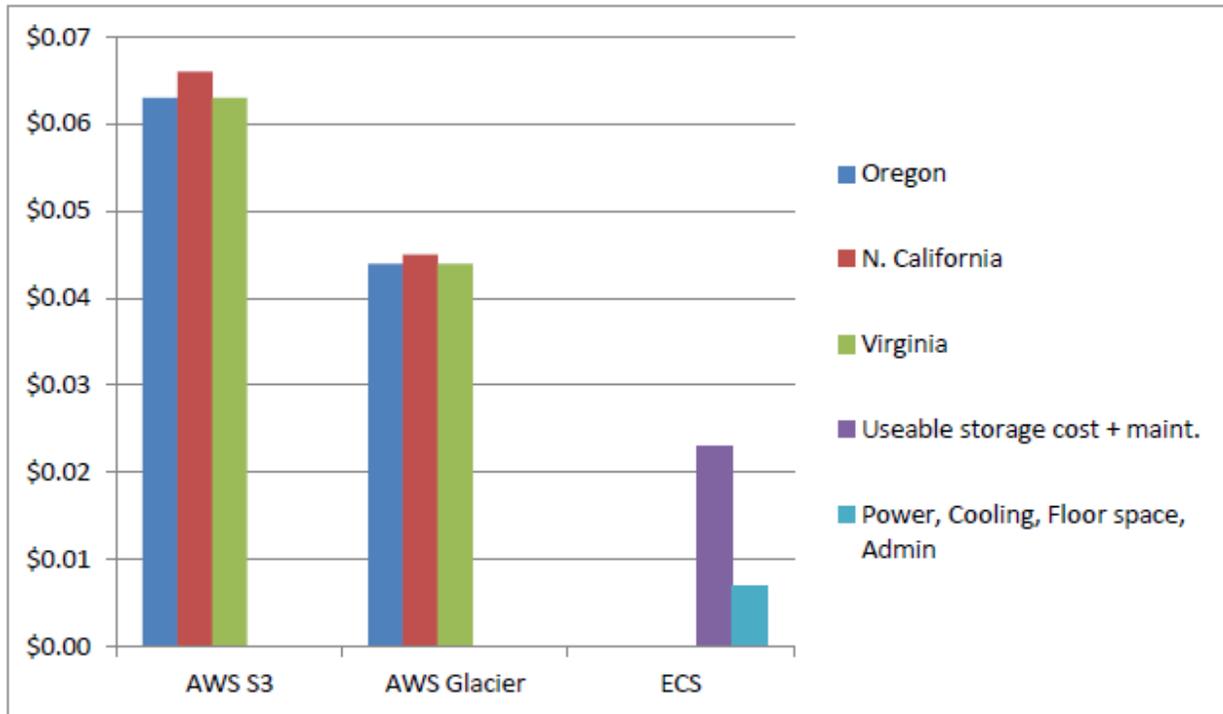

Figure 1: ECS costs are lower as compared to two AWS-based solutions

Not only are the financial benefits for maintaining a private or hybrid cloud infrastructure with ECS compelling, but data stored in ECS can be accessed at the speed of your network, under your positive control when you need it, while objects in AWS Glacier have storage retrieval times that vary between 1-3 hours, not to mention the accumulating costs of AWS S3 Restores.

CONCLUSION

The topic of backing up and retaining data for the long term is not as simple as it used to be. Organizations have been catapulted into a new and constantly changing array of choices. All are seemingly viable, making it increasingly difficult to determine the best solution as part of a business continuity strategy. With the use of tape comes higher RTO and none of the benefits of disk backup, including stream-lined replication across data centers and deduplication of data. The public cloud presents its own set of challenges including how to avoid vendor lock-in and how best to meet compliance and governance concerns against the backdrop of multitenancy and stringent SLAs. The winning solution is one that has a foot in both worlds: a hybrid solution with the flexibility of cloud accompanied by the efficiency of disk, with the CloudBoost solution providing long term retention in the cloud and ECS providing the object storage for safety and security.

ADDITIONAL RESOURCES

CloudBoost Overview – Lightboard Video
CloudBoost Customer Demo – Narrated Demo
Backup to the Cloud with Dell EMC DPS & CloudBoost – Best Practices – White Paper
Dell EMC Data Protection Solutions for Public Cloud Environments - A Technology Preview – White Paper