Data Protection in a Hybrid Cloud Environment

Author: Russ Fellows

May 2018
Executive Summary

Data protection is a topic of concern for IT managers as well as for a growing number of company executives and managers who rely on data in order to operate their businesses. With the increased prevalence of breaches and unauthorized access to data, data protection is an area of increased importance within companies of all sizes. Creating copies or backing up data has been a common practice since modern computing emerged. Although tools and technologies available have changed significantly, the core tenants of data protection have remained relatively constant over time. This includes maintaining multiple copies of data on different types of storage devices, with at least one copy residing off-line and off-site.

Corporate executives and IT professionals are now embracing an IT strategy that is based on using hybrid clouds. By enabling the use of on-premises capabilities for critical applications while leveraging external public cloud resources to expand their capabilities, companies can operate more efficiently while responding to business needs more quickly. One aspect of IT operations that has experienced success in utilizing cloud resources is data protection and disaster recovery.

In this paper, Demartek, together with Evaluator Group, evaluate an HPE solution that leverages a hybrid cloud data protection strategy that combines multiple HPE products together with a public cloud service to retain data protection copies. This solution is compared to alternative tools that are also able to utilize public cloud services to retain protection points. The results are then compared for their on-going costs, network bandwidth required, and how quickly data protection and recovery operations are completed.

Key Lab Insight Findings

- HPE StoreOnce with HPE Cloud Bank Storage vs. ISV backups to the cloud produced:
  - 3X – 5X lower costs over a 5-year period (storage tier dependent)
  - 30% faster backup times (utilizing the same network connection)
- HPE Recovery Manager Central (RMC) protection with HPE Cloud Bank Storage delivered:
  - 10X faster backup, taking 90% less time

Enterprise Data Protection

The concept of data protection has evolved from being concerned primarily with backing up systems and applications to a more holistic approach that now encompasses disaster recovery and business continuity. This has further evolved from an approach that was promulgated on in-house capabilities to a more nuanced strategy that now encompasses many aspects of external cloud resources and services.

A holistic strategy around data protection is one that includes maintaining backup copies in multiple locations, along with other technologies such as replication to protect critical applications. In a previous technical evaluation, Demartek together with Evaluator Group looked at a holistic data protection strategy...
centered around HPE technologies, including HPE StoreOnce backup appliances along with multiple replication technologies implemented with HPE 3PAR all-flash arrays.

In this Lab Insight, Evaluator Group and Demartek extend the analysis of an HPE-centric solution to focus on how HPE StoreOnce backup appliances with HPE Cloud Bank Storage can lower costs while increasing availability and BC/DR options for enterprise IT operations with the use of external cloud services.

**Hybrid Cloud Storage in the Enterprise**

The way in which enterprises are using cloud storage has not changed significantly over the past several years, however the number of companies actively utilizing cloud storage resources and their reliance on the cloud has increased. Evaluator Group performed an extensive analysis of Hybrid Cloud utilization in the enterprise in 2017.¹

This study looked at multiple aspects of Hybrid Cloud usage, including cloud storage. Two questions and responses were particularly relevant to cloud storage as it pertains to enterprise data protection. The first asked about intentions for cloud use, with nearly 60% of respondents answering that they are already using cloud storage and intend to increase their use in the coming year. Another 30% of users indicated they are just beginning to utilize cloud storage.

Given that enterprise hybrid cloud environments often begin with disaster recovery as a primary use case, it is not surprising to see disaster recovery and business continuance capabilities rising to the top of the list. Additionally, it is worth noting that many respondents indicate they now use hybrid cloud storage for additional use cases, such as storage tiering and archiving to supplement their data protection.

In another question, survey respondents were asked to characterize their hybrid cloud storage environments by selecting all applicable use cases.

Results are shown below in Figure 1, with answers listed in descending order.

¹ Evaluator Group Enterprise Hybrid Cloud -A Work in Progress 2017

[www.evaluatorgroup.com/enterprise-hybrid-cloud](http://www.evaluatorgroup.com/enterprise-hybrid-cloud)
As expected, there are issues and concerns around the use of cloud storage in general and in particular as it relates to data protection and BC/DR. Security, performance and availability are the concerns most commonly cited along with the costs associated with cloud storage. Network performance and bandwidth are critical factors for cloud-based protection and are important factors in determining the ability to effectively backup data to the cloud. In some instances, sufficient network capacity may be unavailable or expensive. Moreover, efficient network utilization is a critical component of any cloud disaster recovery or data protection strategy.

**HPE Enterprise Data Protection**

HPE’s data protection portfolio includes multiple products and technologies, including:

- Array-based replication capabilities for primary storage (HPE 3PAR, Nimble and MSA)
- Backup target (HPE StoreOnce purpose-built backup appliance portfolio)
- Cloud-based backup retention (HPE StoreOnce with HPE Cloud Bank Storage)
- Application-based data protection (HPE Recovery Manager Central software)
HPE’s data protection portfolio is designed to enable flexibility, with each aspect of protection able to work independently while offering benefits when combined with other HPE products through integration. An example is Express Protect, a feature of HPE Recovery Manager Central (RMC) that is available when HPE primary storage is used with RMC to create protected data copies on a StoreOnce backup target.

HPE RMC uses snapshots from primary storage as a protection point and then creates a protected copy of the snapshot on a StoreOnce secondary target. By utilizing snapshots as the basis for data protection, operations occur quickly, utilizing native storage snapshot management while also operating efficiently. Using the same snapshot-based techniques that enable rapid data protection when operating in reverse, Express Restore is able to restore systems and applications to an operational state quickly.

Reducing the amount of data moved is one of the key factors in Express Protect and Express Restore (another RMC feature) efficiency. By transferring less data, network bandwidth requirements are reduced while also reducing the amount of data retained. These factors help reduce cloud storage and data transfer costs along with network utilization, thereby addressing the most common issues identified by IT users with cloud storage.

HPE Cloud Bank Storage acts to extend HPE StoreOnce capacity and retention into the cloud. By supporting multiple cloud storage providers, Cloud Bank Storage enables organizations to maintain off-site protected data copies without the need for expensive secondary data centers or cloud-hosted appliances.

### Data Security

As discussed previously, the most important aspects for the use of cloud storage for data protection include costs and security of data along with network access and performance. The leading Cloud Storage providers (AWS, Azure and Google) provide data encryption and key management capabilities. Additionally, StoreOnce supports the use of both local and external key management, such as with HPE Enterprise Key Manager or other standards-based key managers. This enables multiple levels of data encryption for securing data, both locally and off-site, including within Cloud Service providers.

HPE StoreOnce supports local key management for managing keys used for data encryption. Additionally, StoreOnce supports external key management, enabling it to access an external Secure Key Manager. Thus, StoreOnce provides simple integrated key management and a more comprehensive solution that utilizes a common key manager across the enterprise.

With respect to encryption of data, it is possible to utilize two levels of data encryption for data stored in an HPE Cloud Bank Storage repository. Data can be encrypted on the StoreOnce system, which may then be sent in its encrypted form to the Cloud Bank Storage repository; which itself may be encrypted with a different set of keys managed via the specific Cloud Storage provider. With these options, corporate data security, compliance and governance requirements can likely be met by utilizing proper key management techniques and best practices.
Data Retention and Recovery Costs

Cloud storage is viewed as having a single cost (such as $ / GB / month), but in reality, charges are incurred from many aspects. Cloud providers typically charge for several factors including capacity, the data tier utilized, replication, availability and other operational options. For this project, cloud storage cost analysis and testing were performed using Amazon S3 services, specifically utilizing S3, S3-IA and Glacier. A complete breakdown of all charges incurred during testing is provided in the appendices.

Typically, cloud providers do not charge for data uploaded into their cloud, but they do charge for data downloaded. Additionally, depending on the cloud provider, there are often costs for operations involved with data transfer and management. Thus, charges will be incurred for creating and retaining backup data in the cloud, whether using HPE Cloud Bank with StoreOnce systems or another method that stores data in the cloud. One of the primary factors influencing costs is how efficiently data reduction is performed, which impacts the amount of data retained and transferred to the cloud.

Data recovery costs can mount quickly, due to charges both for the amount of data transferred as well as for operations required to perform data recovery. These factors, coupled with limitations on network bandwidth, make recovering data from the cloud less desirable than recovering from an on-site copy of data when possible.

In order to accurately capture and model data protection and recovery costs, Evaluator Group developed analysis tools to model costs based on data backup operations and validate that predicted costs matched those actually incurred. These tools include programmatic analysis of AWS XML usage data combined with Excel templates for calculating costs based on usage.

Testing and the developed modeling tools revealed that charges for a single recovery operation can easily surpass the cost to retain an application’s data for a month or longer. Moreover, we found recovering data from the cloud should only be considered when other options are not available.

Evaluation of HPE with Cloud Storage

The evaluation consisted of establishing primary and secondary sites, along with the use of an off-site public cloud service for retaining backup data. For all testing, Amazon S3 was utilized as the primary interface for retained data, both for HPE Cloud Bank Storage as well as for the alternative backup application tested. For all tests, the same internet connection was utilized as well as the same Amazon S3 location in order to minimize any variances. In this way, cost, performance and other differences could be attributed to the backup technology utilized rather than the cloud service or network. In two instances, an enterprise backup application was utilized for backup and recovery operations. In the third test case, HPE’s RMC software was used for data protection and recovery.

Below in Figure 2 is a diagram of the test environment used for testing.
Evaluation Scenarios

All three of the evaluation scenarios utilized cloud storage as the off-site location for backup data. The three scenarios tested were as follows:

- **Scenario 1**: Direct to cloud with ISV Backup Application
- **Scenario 2**: HPE StoreOnce and HPE Cloud Bank Storage with ISV Backup Application
- **Scenario 3**: HPE StoreOnce and HPE Cloud Bank Storage with HPE RMC Backup Application

**Scenario 1: Direct to Cloud with ISV Backup Application**

The first scenario utilized a leading independent software vendor (ISV) backup application to backup data directly to the cloud using the ISV’s Amazon S3 integration with deduplication enabled. In this scenario, data was retained in the cloud and recovery operations occurred from the cloud. This method may be coupled with additional on-site data stores for retaining backups or utilize the cloud directly.

**Scenario 2: HPE StoreOnce and Cloud Bank Storage with ISV Backup Application**

This scenario again utilized an ISV enterprise backup application. First, data was backed up to a local StoreOnce appliance and then a secondary copy was created as part of the backup policy in the HPE Cloud Bank store residing on an Amazon S3 instance.
Scenario 3: HPE StoreOnce and Cloud Bank Storage with RMC Backup Application

The third scenario was similar to the prior case; instead of utilizing a third-party backup application, HPE RMC was used to perform data protection and recovery operations. With this test case, Express Protect and Express Restore technologies were utilized for data movement between the primary HPE 3PAR storage system and the secondary StoreOnce system. Finally, a policy within RMC was used to create a secondary copy in HPE Cloud Bank Storage on Amazon S3 storage.

Evaluation Findings

Cost Savings

The evaluation found that HPE StoreOnce with HPE Cloud Bank Storage provided significant cost advantages compared to utilizing cloud storage directly with Amazon S3, even when the ISV’s built-in deduplication was used.

- **Up to 5X Lower Cost** — Using Cloud Bank Storage vs. Direct Amazon S3
  - 3X lower if data transferred to Glacier yearly
  - 5X lower if data moved to S3-IA yearly
- **Results for 1 PB of primary data being protected**
  - ISV direct backup to Amazon S3
    - $5.87M with S3-IA and $3.26M with Glacier archival
  - ISV backup with Cloud Bank Storage
    - $1.24M with S3-IA and $1.22M with Glacier

Time to Protect Data Savings

- **30% Faster Backup** — When using Cloud Bank Storage with ISV backup application
  - For 100 TB of primary data, using 100 Mb/s uplink
    - 696 hours (29 days) using ISV to Amazon S3 direct
    - 485 hours (20.25 days) using ISV to HPE StoreOnce Cloud Bank Storage in S3
- **10X Faster Backup** — When using HPE Cloud Bank Storage with RMC
  - For 100 TB of primary data, using 100 Mb/s uplink
    - 696 hours (29 days) using ISV to Amazon S3 direct
    - 66 hours (2.75 days) using RMC to Amazon S3 direct

HPE Data Protection Portfolio

HPE’s data protection portfolio includes multiple products and technologies that operate independently while adding benefits when utilized together. As stated, a separate evaluation looked at HPE array-based replication technologies together with HPE StoreOnce.
The products and technologies evaluated in this paper include:

- **HPE StoreOnce** (D2D backup target) for scalable, on-site backup retention
- **HPE Cloud Bank Storage** – extends HPE StoreOnce capacity to external object storage
  - Scality Object storage for long-term backup retention in a private cloud
  - AWS and Azure for long-term backup retention in the public cloud
- **HPE Recovery Manager Central**
  - Application-directed protection to HPE StoreOnce (VMware, Oracle, SQL Server, SAP HANA & Exchange)
  - HPE 3PAR direct snapshot backup to HPE StoreOnce via Express Protect

**HPE StoreOnce with HPE Cloud Bank Storage**

Utilizing StoreOnce with Cloud Bank Storage enables IT users to replace tape libraries and media management with retention options for both on-site private clouds and off-site public clouds. While D2D backup targets such as StoreOnce may provide a more reliable, higher performance option than tape, these appliances often lack data transport flexibility and have higher retention costs than tape solutions.

Object-based storage systems were designed to store large amounts of data cost effectively, which helps explain their use as a long-term retention target, both on-site and in the cloud. With Cloud Bank Storage, extended retention is possible in both private and public clouds via a direct object interface to support object stores. Typical use cases include:

1. **Extended Retention for DR**
   - HPE Cloud Bank Storage is a tier of storage supported via the HPE StoreOnce appliance
   - HPE Cloud Bank Storage supports multiple public and private storage targets simultaneously
     - Amazon S3, MS Azure, Scality, etc.
   - No separate meta data storage required
2. **Offline Archive Capability**
   - Ability to detach HPE Cloud Bank Storage from HPE StoreOnce with license
   - Migrate data to a different storage tier (e.g. Glacier) for deep archive
   - Ability to attach HPE Cloud Bank Storage to same or different HPE StoreOnce without license
   - Migrate data back to active archive for access
Summary

Data protection and recovery continue to be critical to IT operations due to regulatory requirements, data governance obligations and to protect companies from data loss or other threats posed by security breaches. Cloud storage services are now widely accepted as a viable location for storing off-site data copies to protect against both natural and man-made disasters. While using cloud storage has advantages, both costs and network bandwidth may be an impediment to utilizing this resource.

HPE Cloud Bank Storage provides hybrid cloud capabilities by enabling organizations to maintain backup copies on external, object-based storage hosted in a private cloud and in either Azure or AWS public cloud services. Our evaluation demonstrated that three aspects of Cloud Bank emerged as differentiating factors:

1. HPE Cloud Bank Storage utilizes native cloud object storage interfaces, eliminating the need for cloud-hosted appliances or other interfaces.

2. HPE StoreOnce with HPE Cloud Bank Storage can reduce the amount of backup data retained, thereby lowering retention costs and the amount of time to protect data.

3. All meta-data is contained within the cloud repository, enabling recovery at another location even when all data and systems at a primary site are unavailable.

The capability to utilize both private and public cloud resources is an important factor for StoreOnce with HPE Cloud Bank Storage. Some alternative solutions require cloud-hosted applications, which adds cost and complexity to a hybrid cloud protection strategy. Additionally, some alternatives do not maintain backup data within the repository, which complicates recovery when a disaster or other data loss occurs.

While the ability to utilize private and public cloud storage is significant, the primary reason that StoreOnce with Cloud Bank was more cost-effective and able to store backups more quickly is due to its superior data reduction capability compared to the alternative. Testing also found that additional benefits were realized by using HPE RMC data protection software, either in place of, or as an enhancement to other data protection software. Due to the high efficiency of Express Protect and Restore, backups can be created rapidly and with little or no impact on running applications. This enables IT professionals and application owners the ability to create multiple protection points daily for critical applications or systems.

A hybrid cloud strategy maximizes data on-premises for fast and efficient recovery when required while also ensuring a protected copy of data is maintained off-site. The use of StoreOnce with Cloud Bank Storage was shown to be a more cost-effective solution than an alternative solution using an enterprise backup application backing up directly to a cloud object repository. Additionally, HPE RMC software enabled rapid data protection for sensitive applications.