Technical white paper

Protecting server data with HP StoreOnce Catalyst and HP Data Protector

HP StoreOnce Backup systems—HP StoreOnce Catalyst—HP Data Protector

Table of contents

Introduction

In today's business environment, customers rely on the most efficient, high performing, and reliable backup systems to protect critical business information. Customers need to protect increasing levels of data while keeping costs under control. HP StoreOnce Backup systems provide a disk-based data protection platform while addressing data growth by applying HP StoreOnce deduplication software for efficient, longer term data retention.

The HP StoreOnce 6000 Backup system, the latest deduplication appliance in the HP StoreOnce product line, provides a unique combination of features, including industry-leading performance (up to 100 TB/hr), high-availability, and high-capacity making the HP StoreOnce 6000 Backup system the industry leader in the enterprise deduplication sector.

HP StoreOnce Catalyst software was developed to dramatically improve the performance, function, and integration of backup applications such as HP Data Protector. HP StoreOnce Catalyst delivers deduplication on an appliance server, media server, or dedicated appliance. Since it uses the same deduplication algorithm globally, data can be moved between Catalyst Stores on different platforms without rehydration. HP StoreOnce Catalyst allows better utilization of advanced, disk-based storage solutions while increasing efficiency and performance.

This document describes the benefits of using HP StoreOnce Backup systems combined with HP StoreOnce Catalyst software and HP Data Protector to back up important enterprise data. This document also recommends backup and recovery implementations.

The following are key recommendations for backing up data to an HP StoreOnce Backup system utilizing HP StoreOnce Catalyst software:

- **To improve backup storage utilization:** back up one server at a time (sequentially) for quickly rising deduplication ratios that are maintained over time.
- **To achieve better deduplication ratio than usual:** if the backup process can accommodate it, use a unique HP StoreOnce Catalyst store for each data type or same type of operating system.
- **For the trade-off between backup impact and ease of recovery:** configure a weekly full with daily incremental or initial full with daily incremental backup schedule to reduce the amount of end-to-end data and decrease the time required to run daily backups. HP StoreOnce Backup systems data deduplication will be utilized for full and incremental backups.
- **For efficient and cost effective movement of backup data offsite:** use the HP StoreOnce Catalyst remote copy feature to seamlessly replicate all servers to an appliance in a remote facility for simpler recovery in the event of a disaster.
- **To increase backup speed without affecting deduplication ratios:** use a larger HP Data Protector backup block size for faster backup throughput performance with little or no effect on HP StoreOnce data deduplication ratios.

HP StoreOnce Backup systems are a disk-based backup system that deliver leading price-performance and deduplicate enterprise backup data. The HP StoreOnce Backup system can consolidate the storage of multiple backups onto a single, rack mountable device while improving reliability compared to backup to tape by reducing potential errors caused by media handling. For business environments with remote offices, or a disaster recovery site, the HP StoreOnce Backup system can be used to replicate data to an offsite location.

HP StoreOnce Backup systems are ideal for mission-critical application backup data for small to large data centers running key business applications. Proper backup configurations with a data protection application to the HP StoreOnce Backup system provide the shortest backup times and most efficient use of capacity. HP StoreOnce Backup systems integrate into current IT environments and offer the flexibility of VTL and NAS targets, as well as Catalyst stores.

Technology overview

HP StoreOnce Backup systems—key features and benefits

- **HP StoreOnce deduplication, store more data on disk**

  HP StoreOnce deduplication reduces the disk space required to store backup data sets without impacting backup performance. Retaining more backup data on disk for longer periods of time enables greater data accessibility for rapid restore of lost or corrupt files and reduces downtime.

  Deduplication ratios are strongly influenced by two factors—data change rate and backup data retention periods. Low data change rates and data retained for longer periods of time yield higher deduplication ratios.

- **Optimized replication**

  HP StoreOnce deduplication is the technology enabler for HP StoreOnce replication which allows fully automated replication without rehydration. This optimized replication works over low bandwidth links to a disaster recovery (DR) site, giving Remote Office/Branch Office (ROBO) and small data centers a cost effective DR solution for the first time.

- **Rapid restore of data for dependable, worry-free data protection**

  HP StoreOnce Backup systems offer immediate access to backups for rapid restores. HP StoreOnce deduplication allows more data to be stored closer to the data center for longer periods of time which offers immediate access for rapid restores.

- **Automate, simplify, and improve the backup process**

  HP StoreOnce Backup systems automate the backup processes allowing reduced time spent managing data protection. Implementing hands-free, unattended daily backup is especially valuable for environments with limited IT resources, such as remote or branch offices.

  HP StoreOnce systems can back up multiple servers via a standard Ethernet or Fibre Channel network simultaneously to a disk-based solution at peak speeds of up to 100 TB per hour instead of sequentially to a tape drive or autoloader, meaning that substantially reduced backup windows are possible.

  HP StoreOnce systems can be managed and configured by using the built-in Web browser based graphical user interface (GUI). For larger deployments of replicating HP StoreOnce appliances, the HP StoreOnce Enterprise Manager can monitor multiple backup systems across geographies. HP StoreOnce systems are self-managing backup appliances that require little, if any, routine maintenance. Unlike other disk-based storage devices, HP StoreOnce systems do not require virus protection or additional disk storage.

HP StoreOnce Catalyst—seamless data movement across the enterprise

HP StoreOnce Catalyst brings the HP StoreOnce vision of a single, integrated enterprise-wide deduplication algorithm a step closer. It allows the seamless movement of deduplicated data across the enterprise to other HP StoreOnce Catalyst systems without rehydration. This means that benefits can be seen from:

- **Simplified management of data movement from a single pane of glass:** Tighter integration with the backup application to centrally manage file replication across the enterprise.
- **Seamless control across complex environments:** Supporting a range of flexible configurations that enable the concurrent movement of data from one site to multiple sites, and the ability to cascade data around the enterprise (sometimes referred to as multihop).
- **Enhance performance:** HP StoreOnce Catalyst has the ability to perform deduplication at any DP client with the DP media agent installed. This “federated” deduplication only transfers new data across LAN or WAN links to the StoreOnce appliance and delivers faster throughput thus reducing backup windows.

Note:
Actual performance is dependent upon configuration data set type, compression levels, number of data streams, number of devices emulated and number of concurrent tasks, such as housekeeping or replication.

HP StoreOnce Backup systems in small to large data centers

Advantages of using HP StoreOnce Backup systems and HP Data Protector

- The HP StoreOnce deduplication engine can run on a client with the HP Data Protector media agent, an HP Data Protector backup server that can back up data from other servers, or an HP StoreOnce Backup system, providing flexibility in deduplication strategy
  - HP StoreOnce Backup systems target-side deduplication can reduce workload on the backup server
  - HP Data Protector configured with HP StoreOnce Catalyst using source-side/server-side deduplication, allows deduplication to occur at the client. This reduces amount of data packets sent across network which reduces network load on IT infrastructure.
- Easier setup to protect, manage, and access data
- Individual HP StoreOnce Catalyst stores can be configured in HP Data Protector as separate storage devices for different purposes such as applications, host operating systems, duplicate copies and archival
- HP Data Protector backup architecture can be scaled out easily by adding Media servers, and HP StoreOnce Catalyst stores
- Server side deduplication reduces workload on IT infrastructure and reduces CPU and memory requirements on the HP StoreOnce Backup system allowing for faster throughput performance
- HP StoreOnce Backup systems common deduplication engine allows replication to remote sites with limited bandwidth and without having to be rehydrated
- With federated data deduplication, HP StoreOnce Backup system allows more enterprise backup data to be retained on disk for longer periods
- Improves functionality, performance and total cost of ownership while migrating enterprise backup data protection environments from disparate small systems into scalable HP StoreOnce Backup systems

HP StoreOnce Backup systems for enterprise data backup using HP Data Protector

An important part of server administration is maintaining a consistent set of backup data, which should be available for recovery. When data is lost due to user error, system failure, or catastrophic site failure, there is a need for complete server recovery along with application data recovery. HP Data Protector provides data protection for multiple, heterogeneous servers and operating systems. The backup data can be consolidated to a single HP StoreOnce Backup system leveraging 10 Gb Ethernet and 8 Gb Fibre Channel speed. HP StoreOnce Backup systems integrated with a well-planned data protection strategy include regular backups to maintain a consistent set of data for recovery purposes.

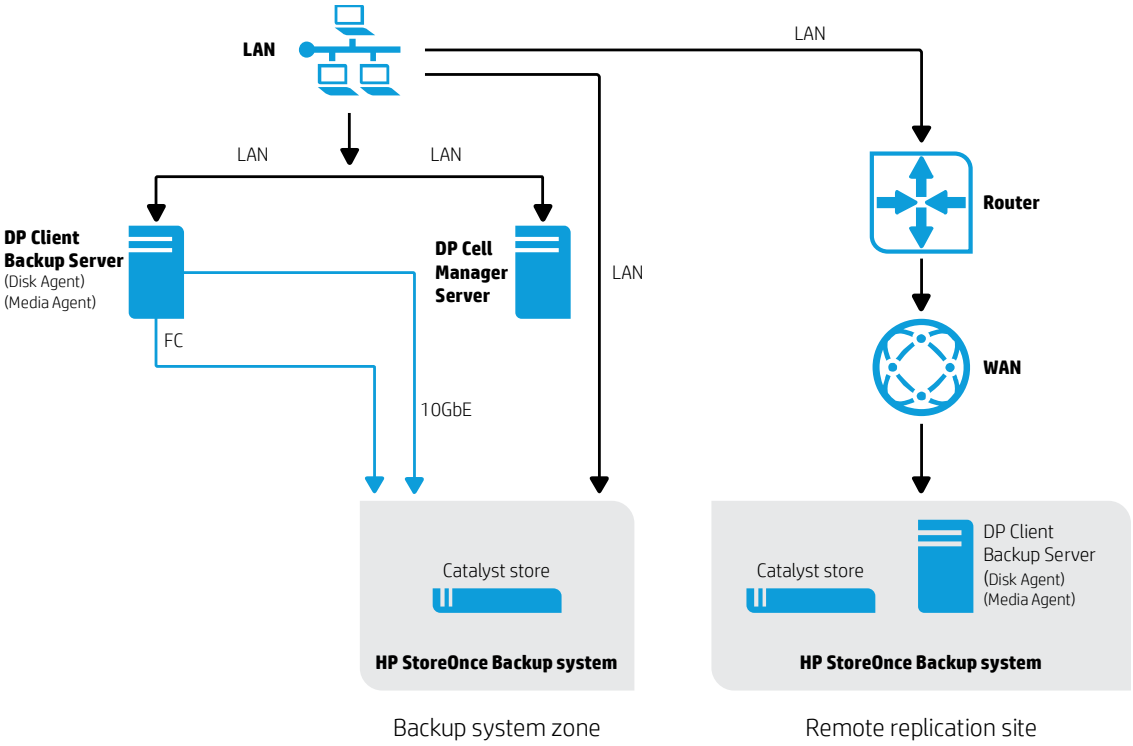
HP Data Protector backup environment

Table 1. HP Data Protector backup architecture

Component	Description
Cell Manager Server	An HP Data Protector cell is a set of systems with a common backup policy existing on the same LAN/SAN. The Cell Manager is the main system that is the central point for managing this network environment. It contains the HP Data Protector internal database (IDB) and runs core HP Data Protector software and session managers. The IDB keeps track of backed up files and the cell configuration.
DP Client	A host system becomes a HP Data Protector client when one or more of the HP Data Protector software components are installed on the system. Client systems with disks that need to be backed up must have an appropriate Data Protector Disk Agent component installed. The Disk Agent enables you to back up data from the client disk or restore it. Client systems that are connected to a backup device must have a media agent component installed. This software manages backup devices and media.

Figures 1 Illustrates an HP Data Protector backup and recovery topology that supports HP StoreOnce Backup systems.

Figure 1. HP Data Protector with an HP StoreOnce Backup system



Capacity planning

The required backup storage capacity for server backups depends on the following:

- Size and number of virtual servers
- Backup retention policy (recovery points needed)
- Type of backups (full, incremental, differential)
- Frequency of backups
- Data rate of change
- The deduplication ratio achieved by the HP StoreOnce Backup system

HP StoreOnce Backup systems do not deduplicate across Catalyst stores. Each Catalyst store is an independent deduplication domain. To achieve a better deduplication ratio a unique Catalyst store should be created specifically for similar data types. For larger environments, multiple Catalyst stores may work best with backups of similar type of operating system data stored in the same HP StoreOnce Backup system.

Note:
The rate of change of a data refers to the amount of data that would be contained in an incremental backup as a percentage of a full backup. A 100 GB full backup with a subsequent 5 GB incremental backup before the next full backup would be a five percent rate of change.

In performing these tests HP used as standard customer representative data set with realistic structure and content.

An administrator may desire to have two weeks of backup stored on the HP StoreOnce Backup system for quick recovery access.

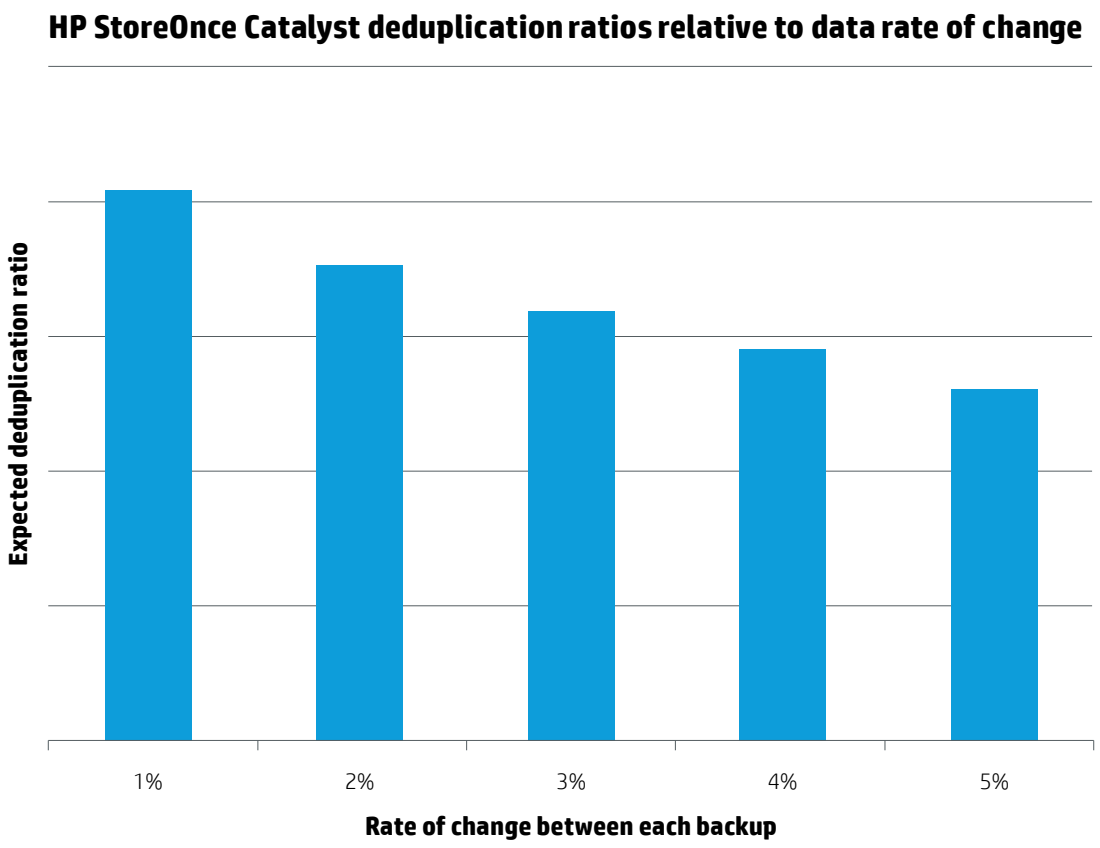
HP Data Protector offers three different deduplication options for transferring data to HP StoreOnce Catalyst stores.

- **Source side deduplication:** Deduplication takes place within the backup server. The HP StoreOnce deduplication code is embedded within the HP Data Protector Media Agent. Use source side deduplication to back up data which is located on the same server. This is also referred to as Catalyst Low Bandwidth.
- **Server side deduplication:** Deduplication takes place within the backup server as in source side deduplication. Server side deduplication can be used for backing up data from other servers with HP Data Protector disk agents as well as data held itself. Data is transferred via network connection to the backup server.
- **Target side deduplication:** Deduplication takes place within the HP StoreOnce Backup system. The deduplication code is embedded within the HP StoreOnce Backup system.

HP Data Protector rate of data change effect on deduplication ratios

Figure 2 shows the data rate of change effect on deduplication ratios when backing up to a HP StoreOnce Catalyst target. The source side deduplication option was used in this test.

Figure 2. HP StoreOnce Catalyst rate of change effect on HP StoreOnce deduplication ratios



HP Data Protector block size effect on throughput and deduplication ratio

HP Data Protector supports varying block sizes for backing up server data. Larger block sizes usually result in better throughput. HP recommends keeping the block size set at 256 and above for better throughput.

Figure 3 illustrates how HP StoreOnce Catalyst backup throughput benefits from higher block size.

Figure 3. Varying block sizes vs. Backup throughput (using low bandwidth)

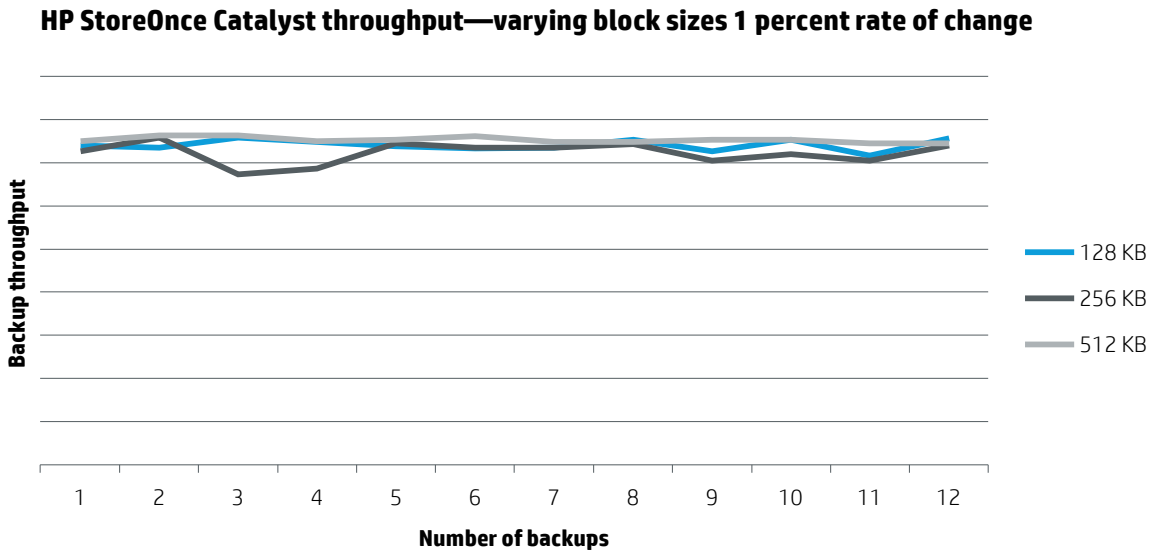
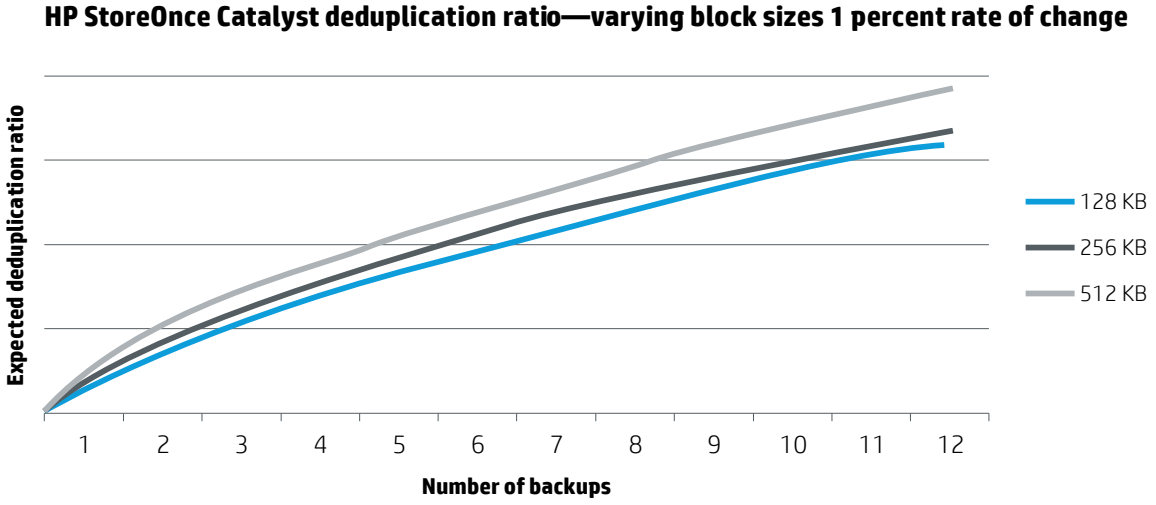


Figure 4 illustrates HP StoreOnce Catalyst deduplication ratios when using different backup block sizes (using low bandwidth).

Figure 4. Varying block sizes vs. deduplication ratios (using low bandwidth)



Other factors affecting throughput

- **Concurrent operations:** HP Data Protector concurrent operations allow multiple backup jobs to run in parallel to a single backup target. Use concurrent operations to increase HP StoreOnce device throughput and reduce the backup window. Using HP Data Protector concurrent operations leads to interleaved data and may decrease deduplication ratios.
- **Running verify operations during backup can affect backup performance:** The verify operation ensures the integrity of backup data after the backup is completed. Schedule the verify operation to run at a convenient time outside of backup window.

Other factors affecting deduplication ratio

- **Backup policies:** A business need to retain data for longer a period of time improves the chance that common data will already exist in storage, resulting in greater storage savings and better deduplication ratios.
- **Compression/Encryption:** Software compression and encryption prevents optimal deduplication from taking place since data that is already compressed and encrypted cannot be efficiently deduplicated.
- **Data types:** To improve deduplication ratio, configure individual HP StoreOnce Catalyst store devices as separate storage devices in HP Data Protector for individual server/type of operating systems.

HP Data Protector schedule type characteristics with an HP StoreOnce Backup system

Many backup environments take advantage of HP Data Protector's different backup schedule types such as:

- Daily full
- Weekly full with daily incremental
- Incremental forever

Table 2 lists some characteristics of full and incremental backups to a HP StoreOnce Backup

Table 2. Comparison of different backup schedule types.

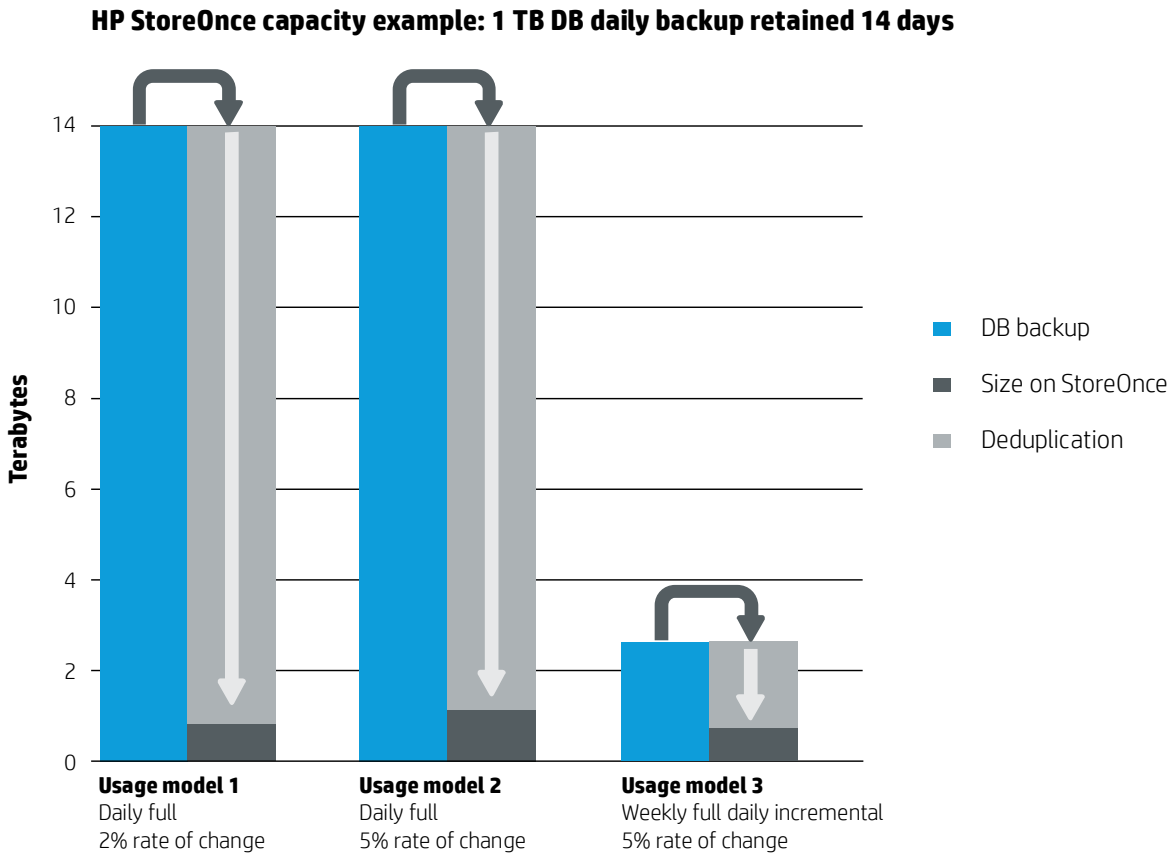
	Daily full	**Weekly full/Daily incremental**
Description	Daily backup of all files	Weekly backup of all files, daily backup of files that have changed since the last backup
Backed up data	All data	Weekly full—all data Daily incremental—changed data since the last backup
Relative backup time	Long	Short, except on days of full backup
Relative recovery time	Short	Long
Server load	High	Low, except on days of full backup
SAN/LAN bandwidth requirement	High	Low, except on days of full backup
Relative dedupe ratio on appliance	High	Low
Size on HP StoreOnce compared to non-deduplication	Very small	Very small

Figure 5 illustrates the end-to-end data compaction of full and incremental backups to a HP StoreOnce Backup system over a period of 14 days (using low bandwidth) with a 1 TB dataset.

Note:
Data compaction refers to the removal of redundant information from a backup set prior to storing on a backup device. Incremental backups, deduplication, and compression are all methods for removing redundant data from a backup set.

Each backup shows the overall size of the backup data without deduplication vs. the size of the data on the HP StoreOnce Backup system after deduplication with Catalyst.

Figure 5. Data compaction comparison of different backup schedule types (using low bandwidth).



Disaster recovery with local HP StoreOnce Backup system and remote replication

Most companies recognize the importance of a robust data protection strategy. Enterprise-level customers are likely to invest in local server recovery as well as site disaster recovery at a remote site using replication. Many companies, large and small, are protecting servers in remote offices where untrained IT staff are expected to manage a daily backup process—generally involving the changing of physical tapes, which is a process prone to more time consumption, resource draining, inefficient and human error.

Replicating large volumes of data over a typical WAN is expensive. However, today's products with data deduplication have made it possible to replicate data over lower bandwidth links for a more cost effective, network efficient replication solution that provides a practical disaster recovery solution and an ideal solution for centralizing the backup of remote offices.

Data deduplication shrinks the amount of backup data that needs to be replicated from the source HP StoreOnce appliance, and as a result significantly reduces replication bandwidth requirements. Once a replica of the backup dataset has been created on a remote HP StoreOnce target appliance all that is required to keep the replica identical to the source is the automatic, periodic copying and movement of the new data segments which are created during each backup. With such small amounts of data being transmitted asynchronously, lower bandwidth networks offer sufficient performance and a much lower cost solution.

Note:
Replication of data can only occur between devices within the same product family i.e., HP StoreOnce, but not VLS.

HP Data Protector replication using HP StoreOnce Catalyst Copy

One of the key features HP StoreOnce Catalyst stores provide is allowing HP Data Protector to utilize low bandwidth copy feature to replicate backup data. HP Data Protector allows backup data to function between Catalyst stores, providing HP Data Protector complete control over the backup data lifecycle. This is accomplished by using HP Data Protector's Object Copy feature. Replicating backup data between HP StoreOnce Backup systems is accomplished by properly configuring HP Data Protector Object Copy to replicate the backup data to alternate deduplication devices, one being the original backup data store and the other being the copy data store. Therefore, whenever the backup job sends data to the primary HP StoreOnce Backup system it is automatically copied to the alternate HP StoreOnce Backup system.

For configuring HP Data Protector Object Copy see the concepts guide available from the HP Data Protector Manager GUI:

HP Data Protector Manager GUI -> Help -> Guides -> Concepts Guide

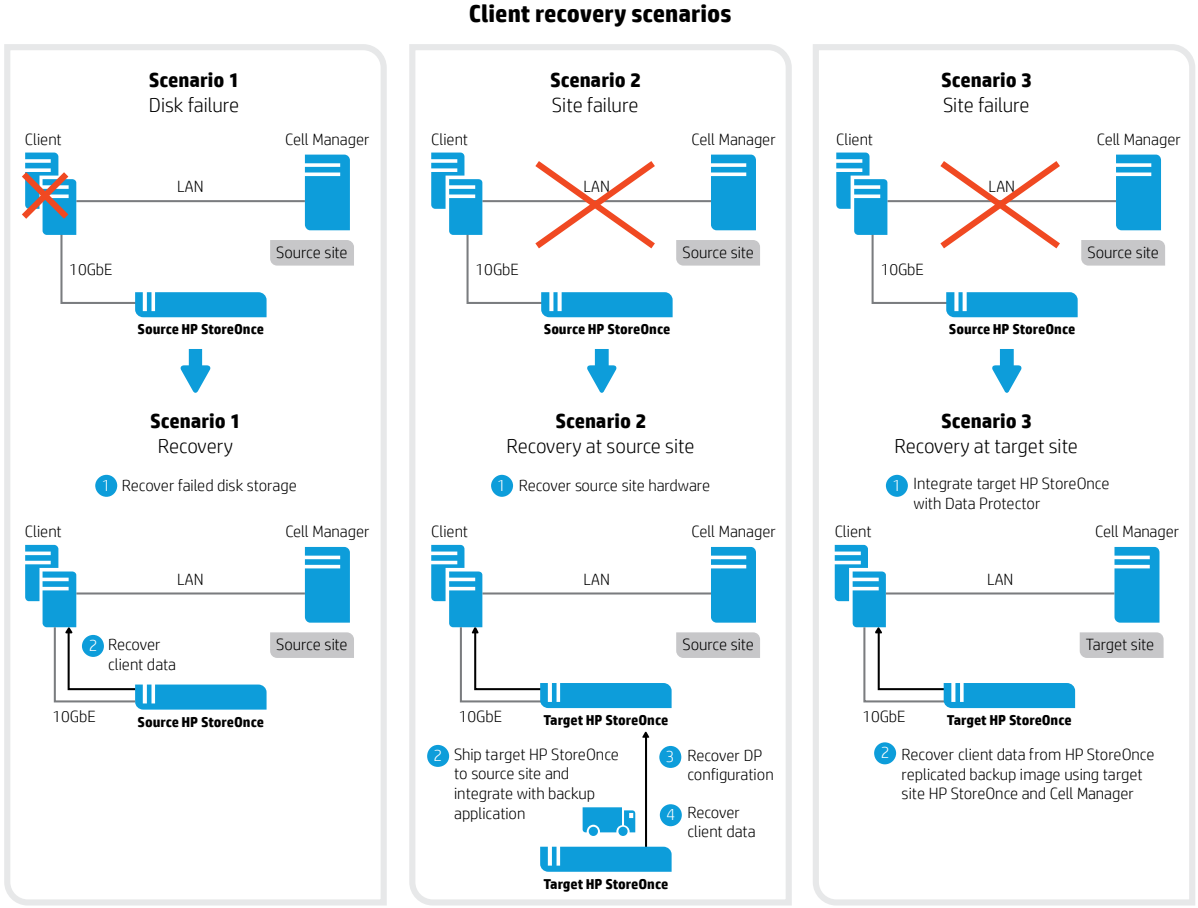
There is flexibility in doing data recovery, depending upon the situation or type of failure (see Figure 6 below). For instance:

- Clients can be recovered at the HP StoreOnce source site (original client location).
- In the event of a client source site disaster, the target site HP StoreOnce can be shipped to the source site or the backup data can be replicated back to the source site for complete client recovery.
- Clients can be recovered at the HP StoreOnce target site (remote location).

Recovery scenarios

Figure 6 illustrates disaster recovery scenarios that may occur and the recovery path available when replicating between HP StoreOnce Backup systems.

Figure 6. Recovery scenarios



Recommendations

- **Concurrent operations**
  - Increased backup speed: If backup throughput performance is the highest priority, use HP Data Protector's concurrent operations to send multiple backup jobs simultaneously to the HP StoreOnce Backup system.
  - Better deduplication ratios: If deduplication ratio is the highest priority, backup one server at a time by disabling HP Data Protector concurrent operations.
- **Backup block size**
  - Increased backup speed: Use a larger block size for improved backup throughput performance.
  - Deduplication ratios: Changing the backup block size has very little effect on HP StoreOnce deduplication ratios.
- **Daily full or weekly full with daily incremental backups**
  - Daily full backups deduplicate at a much higher rate than weekly full with daily incremental, but require more server and HP StoreOnce processing resources and SAN/LAN bandwidth.
  - End-to-end data compaction is greater for incremental backup schedules over an extended time period, which means less storage space will be used on the HP StoreOnce Backup system.
- **Disaster recovery**
  - HP StoreOnce remote replication offers a low bandwidth replication solution to and from remote sites, which is ideal for server disaster recovery.
  - HP StoreOnce Catalyst store pairs configured with HP Data Protector's Object Copy feature to duplicate the backup set provides recovery for local disk failures, complete server failure, or complete site failures by keeping server backup copies at local and remote sites.

Conclusion

Customers demand an efficient, reliable data growth management backup system environment while keeping costs under control. HP provides a variety of reliable data protection storage solutions that address such requirements. HP StoreOnce Catalyst, along with HP Data Protector, is one such solution. HP StoreOnce Backup systems offer high performance and reliability, while addressing data growth through HP StoreOnce data deduplication technology. In addition, HP Data Protector's data protection solution brings together a full generation of traditional and next generation data protection from backup, to disk to replication management, to tape under one platform. In all, HP StoreOnce Backup systems integrate easily with HP Data Protector to protect important data for mission-critical applications.

Useful links

HP StoreOnce Backup
hp.com/go/StoreOnce

HP StoreOnce Backup system user guide
bizsupport1.austin.hp.com/bc/docs/support/SupportManual/c02295179/c02295179.pdf

HP StoreOnce Backup systems Linux and UNIX® configuration guide
h20000.www2.hp.com/bc/docs/support/SupportManual/c02299831/c02299831.pdf

HP StoreOnce Backup systems best practices guide
bizsupport2.austin.hp.com/bc/docs/support/SupportManual/c02511912/c02511912.pdf

HP StoreOnce Catalyst Solution Service
Technical data sheet: h20195.www2.hp.com/V2/GetPDF.aspx/4AA4-4489ENW.pdf
Customer brochure: h20195.www2.hp.com/V2/GetPDF.aspx/4AA4-4480ENW.pdf

HP Data Protector Administrator's guide
bizsupport1.austin.hp.com/bc/docs/support/SupportManual/c00663793/c00663793.pdf

Learn more at
hp.com/go/StoreOnce

Sign up for updates
hp.com/go/getupdated

Share with colleagues

Rate this document

4AA4-8380ENW, September 2013