Technical white paper

Protecting VMware Virtual Machines with HP StoreOnce Catalyst

Using HP Data Protector

Table of contents

Introduction........................................................................................................................................................................... 2
Technology overview............................................................................................................................................................ 3
  HP StoreOnce Backup systems—key features and benefits .......................................................................................... 3
  HP StoreOnce Catalyst—seamless data movement across the enterprise ................................................................. 3
HP StoreOnce Backup systems in small to large data centers ....................................................................................... 4
  Advantages of using HP StoreOnce Backup systems for VMware VM backups ............................................................. 4
VMware backup methods with HP Data Protector ............................................................................................................ 4
  HP StoreOnce Backup systems for VMware VM backup using HP Data Protector ....................................................... 5
  VMware virtual infrastructure backup components ...................................................................................................... 5
Capacity planning ................................................................................................................................................................. 7
  Capacity planning usage models .................................................................................................................................... 8
Virtual machine disaster recovery with local HP StoreOnce Backup system and remote replication ................... 11
  HP Data Protector replication using HP StoreOnce Catalyst ..................................................................................... 11
Recovery scenarios ............................................................................................................................................................ 12
Recommendations .............................................................................................................................................................. 13
Conclusion.......................................................................................................................................................................... 13
Introduction

VMware is increasingly prevalent in today’s business atmosphere. VMware customers need efficient, high performing, and reliable backup systems that are easily integrated into the vSphere environment. With increasing virtual machine (VM) count, keeping backup infrastructure costs under control can be challenging. 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 B6200 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 B6200 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 along with VMware vStorage APIs for Data protection (VADP) and Media agent in VM to backup VMware VMs. This document also recommends backup and recovery implementations.

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

- To protect VM images: VADP in combination with HP StoreOnce Catalyst is a great way to protect VM images. HP Data Protector uses VADP to automatically discover VMs, HP StoreOnce Catalyst software deduplicates VM backup images at the backup server, and HP Data Protector copies the VM backup images to an HP StoreOnce Catalyst store.
- Media agent in VM: HP StoreOnce Catalyst stores can be presented directly to VMware VMs. Installing HP Data Protector on a VM allows the VM to use HP StoreOnce Catalyst software to deduplicate the data prior to sending to the Catalyst store. Remote office/branch office (ROBO) can take advantage of this feature by deduplicating backup data prior to sending to remote HP StoreOnce Backup targets, thus using less network bandwidth. Installing HP Data Protector in a VM enables typical backup configuration and the use of HP Data Protector features such as database backup integrations.
- 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 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.

HP StoreOnce Backup systems are a disk-based backup system that deliver leading price-performance and deduplicates VMware VM image backups. The HP StoreOnce Backup system can consolidate the backups of many virtual machine images 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 virtual machine backup images. Proper configuration of VMware VM backups with a data protection application and the VMware VADP to the HP StoreOnce Backup system provides the best backup throughput performance and data deduplication ratios. 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.

Deduplication-enabled replication
HP StoreOnce deduplication is the technology enabler for HP StoreOnce Deduplication-enabled replication which allows fully automated replication 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 backup 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 intuitively managed and configured by using the built-in Web browser’s administrative interface. For larger deployments of replicating HP StoreOnce appliances, the HP StoreOnce Replication 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 LUN provisioning.

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 multi-hop).
- Enhance performance: Distributed deduplication processing using HP StoreOnce Catalyst stores on the HP StoreOnce Backup systems and on multiple servers can optimize loading and utilization of backup hardware, network links and backup servers for faster deduplication and backup performance.

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 for VMware VM backups**

- The HP StoreOnce deduplication engine can run on a VADP backup server, a VM media server, a VM client or an HP StoreOnce Backup system, providing flexibility in deduplication strategy
  - HP StoreOnce Backup systems target-side deduplication can reduce workload on VADP backup server, VM, or ESXi host
  - HP Data Protector configured with HP StoreOnce Catalyst using source-side/server-side deduplication, allows deduplication to occur at VM backup client or VADP server. 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 VM backup and VM backup mirror copies
- HP Data Protector backup architecture can be scaled out easily by adding Media servers, and HP StoreOnce Catalyst stores
- HP StoreOnce Backup systems common deduplication engine allows VMs to move to remote sites with limited bandwidth and without having to be rehydrated
- With federated data deduplication, HP StoreOnce Backup system allows more VM backup data to be retained on disk for longer periods
- Improves functionality, performance and total cost of ownership while migrating VMware data protection environments from disparate small systems into scalable HP StoreOnce Backup systems

**VMware backup methods with HP Data Protector**

Different VMware VM backup and recovery methods can be used with HP StoreOnce Catalyst and HP Data Protector.

**VStorage APIs for Data Protection**

VADP is a framework which enables backup applications such as HP Data Protector to backup VMware virtual machines using an off-host backup server running the HP Data Protector media agent. VADP reduces the resources required by an ESXi host for VM backup by offloading backup responsibilities to a backup server. Integrating HP StoreOnce Catalyst with VADP reduces the space needed for VM backups and when integrated with multiple backup servers may provide increased throughput.

**Media agent in VM**

Installing the HP Data Protector media agent in each VM is an easy way to backup VMs; plus it allows use of the HP Data Protector integrations for applications such as MS SQL Server that may be running in VMs. HP StoreOnce Catalyst integration with HP Data Protector on a VM enables data deduplication at the VM thus reducing the backup throughput bandwidth required. However, ESXi host resources can be impacted when VMs are vying for resources for backup. Backing up VMs to an HP StoreOnce Catalyst store with the Media agent in VM may result in decreased throughput. Restoring data on virtual machines is the same as traditional data restore.
Figure 1 shows the throughput comparison when used backup method VADP vs. Media agent in VM with single stream.

Figure 1. HP StoreOnce Catalyst throughput performance comparisons

![Throughput Comparison Diagram](image)

HP StoreOnce Backup systems for VMware VM backup using HP Data Protector

An important part of virtual machine administration is maintaining a consistent set of virtual machine image backups available for recovery. When data is lost due to user error, system failure, or site catastrophe, there is a need for virtual machine image and possibly application data recovery. The backup data can be consolidated to a single HP StoreOnce Backup system leveraging 10 Gb Ethernet and Fibre Channel speed. HP StoreOnce Backup systems integrated with a well-planned data protection strategy that includes regular virtual machine image backups, will maintain a consistent set of virtual machine images for recovery purposes.

VMware virtual infrastructure backup components

Table 1. VMware backup architecture

<table>
<thead>
<tr>
<th>Component</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>VMware ESXi server</td>
<td>Virtual infrastructure software for portioning, consolidating and managing computing resources.</td>
</tr>
<tr>
<td>VMware Virtual Machine</td>
<td>A virtual server that emulates various servers based on different operating systems. VMs are hosted by ESXi servers.</td>
</tr>
<tr>
<td>vCenter Server</td>
<td>A Windows service or VM appliance that acts as a central administrator for VMware ESXi servers and VMs that are connected on a network.</td>
</tr>
<tr>
<td>vStorage APIs for Data Protection (VADP)</td>
<td>Enables data protection software to protect system, application and user data in your virtual machines in a simple and scalable way.</td>
</tr>
<tr>
<td>VADP Backup server</td>
<td>For complete integration with VADP and HP StoreOnce Backup Systems a physical backup server is best. The backup server uses data protection software working with VADP to backup VMs.</td>
</tr>
<tr>
<td>Media Server</td>
<td>Provides additional backup storage by allowing HP Data Protector to use the backup storage devices attached to the server.</td>
</tr>
</tbody>
</table>
Figures 2 and 3 illustrate a backup and recovery topology with a backup application that supports HP StoreOnce Backup systems and VMware VADP or a Media Agent in a VM.

**Figure 2.** VMware VADP with an HP StoreOnce Backup system

**Figure 3.** VMware Media Agent in VM 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 machines
- 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 virtual machine backup images. For larger environments, multiple Catalyst stores may work best with backups of disparate VM operating systems.

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 VMware ESXi host with two Windows and two Linux VMs.

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

Figure 4 shows the data rate of change effect on deduplication ratios when backing up VMware VMs using VADP. In performing these tests HP used standard Windows and Linux VMs and a customer representative dataset with realistic structure and content. For the chart below, VM data was updated between each backup until the desired rate of change was reached.

Figure 4. VMware VM rate of change effect on HP StoreOnce deduplication ratios
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 Data Protector Media agent. Use Source side deduplication to backup data which is located on the same server.

- **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 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.

Figure 5 shows the deduplication ratio comparison when using Server side deduplication vs. Target side deduplication for backing up VMware VMs to an HP StoreOnce Catalyst store.

**Figure 5.** HP StoreOnce Catalyst deduplication ratio comparisons for Server side deduplication vs. Target side deduplication

![](image)

### Capacity planning usage models

A VMware environment with 100 VMs of 75 GB each with a 14-day backup data retention requirement can have several HP StoreOnce backup usage models. Usage models change based on variables such as the following:

- **Backup schedule type**
  - Daily VADP full image backups de-duplicate well but use more server and HP StoreOnce compute and bandwidth resources during a backup.
  - Weekly full with daily incremental backups do not deduplicate as well but use less compute and bandwidth resources.
  - End-to-end data compaction for weekly full with daily incremental backup schedules is comparable to daily full backups and may be slightly better.

- **VM daily rate of change**—Lower change rates result in better deduplication ratios and require less HP StoreOnce backup storage.

- **Backup block size**—The backup block size set by the HP Data Protector application can affect HP StoreOnce deduplication ratios. A larger block size will usually result in higher deduplication ratios.

- **Sequential or simultaneous VM backups**—Multiple VM backups running simultaneously, typically have better backup throughput but affects HP StoreOnce deduplication ratios.
Figure 6 compares five usage models for VMware backups to an HP StoreOnce Backup system with the following common characteristics:

- Number of VMs: 100 VM
- VM Size: 75 GB per VM
- Backup schedule: Daily
- Retention period: 14 days

**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 VM data without deduplication vs. the size of the data on the HP StoreOnce Backup system after deduplication with Catalyst.

**Figure 6.** Data compaction comparison of different backup types impacting deduplication ratios.
Figure 7 shows how VADP VM image full backup performance throughput increased when multiple data streams were used with four VMware VMs.

**Figure 7.** Backup performance increases with multiple data streams (using low bandwidth)

Other factors affecting throughput:

- Data multiplexing: When multiple VMware VMs backup in parallel to a single device. When backing up multiple VMs simultaneously use parallel streams to increase throughput. The HP StoreOnce Backup system interleaves backup data when writing parallel streams to a single Catalyst store. Interleaved data may decrease deduplication ratio.

- Concurrency: HP Data Protector Concurrency allows multiple VM backup jobs to run in parallel to a single backup target. Use Concurrency to increase HP StoreOnce device throughput and reduce the backup window. Using HP Data Protector Concurrency leads to interleaved data and may decrease deduplication ratios.

Figure 8 illustrates how deduplication ratios decrease when multiple VADP VM image data streams are used.

**Figure 8.** Deduplication ratios decrease with multiple data streams (using low bandwidth)
Other factors affecting deduplication ratio:

- Backup policies: Retaining VM backup images for longer periods of time improves the chance that common data will already exist in storage, resulting in greater storage savings and better deduplication ratios.
- Data types: To improve deduplication ratios, configure individual HP StoreOnce Catalyst stores in HP Data Protector for individual VM operating systems.

**Virtual machine disaster recovery with local HP StoreOnce Backup system and remote replication**

Most companies recognize the importance of a robust data protection strategy. Enterprise-level VMware customers are likely to invest in local VM recovery, as well as, site disaster recovery at a remote site using replication. Many companies, large and small, are protecting VMs 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 quantities of VM images over a typical WAN is expensive. However, today’s products with data deduplication have made it possible to replicate VM images 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 VMware VM image 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 VMware VM image 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**

One of the key features HP StoreOnce Catalyst stores provide is allowing Data Protector to utilize low bandwidth Mirror feature to replicate backup data. Data Protector allows mirror backup data to function between Catalyst stores, allowing Data Protector complete control over the backup data lifecycle. This is accomplished by using HP Data Protector’s “Mirror” feature to control the lifecycle of backup data. Replicating backup data between HP StoreOnce Backup systems is accomplished by properly configuring a mirror to replicate the backup data to alternate deduplication devices, one being the original backup data and the other being mirrored data. Therefore, whenever the backup job sends data to the primary HP StoreOnce Backup system it gets mirrored to the alternate HP StoreOnce Backup system.

For configuring mirrored devices in HP Data Protector see the Concepts guide available from the HP Data Protector Manager GUI:

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

There is flexibility in doing VM recovery, depending upon the situation or type of failure (see Figure 12 below). For instance:

- A VM can be recovered at the HP StoreOnce source site (original client location).
- In the event of an ESXi host 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 VM recovery.
- A VM can be recovered at the HP StoreOnce target site (remote location).
Recovery scenarios

Figure 9 illustrates disaster recovery scenarios that may occur and the recovery path available when replicating between HP StoreOnce Backup systems.

**Figure 9. Recovery scenarios**

### VM recovery scenarios

#### Scenario 1  
**Disk failure**

1. **Recovery**
   
   - Recover failed disk storage
   - Recover VM image

#### Scenario 2  
**Site failure**

1. **Recovery**
   
   - Recover source site hardware
   - Recover VM image

2. **Recovery at target site**
   
   - Integrate target HP StoreOnce with backup application
   - Recover VM from HP StoreOnce replicated backup image using target site HP StoreOnce


**Recommendations**

- **Source/Server side deduplication vs. Target side deduplication**
  - Reduce network load: Source/Server side deduplication reduces the amount of data packets sent across a network because deduplication occurs at the VM backup client or VADP server.
  - Reduce workload: Target side deduplication takes place within the HP StoreOnce backup system, thus reducing backup workload on VADP backup servers, VMs, ESXi hosts.

- **VADP vs. Media agent in VM**
  - VADP: Provides a good method for backing up VM images by installing the backup agent on a backup server which utilizes auto-discovery to protect all VM images. Backup software is not required on the ESXi host or VM.
  - Media agent in VM: For a more traditional and simpler backup method, just as a physical server, the media agent can be installed in a VM for direct access to HP StoreOnce Catalyst store.

- **Data multiplexing**
  - Increased backup speed: If backup performance is the high priority, use Data Protector concurrent streams to backup multiple VMs simultaneously to the same HP StoreOnce Catalyst store.
  - Better deduplication ratios: If deduplication ratio is the high priority, disable Data Protector concurrent streams so VM backups run sequentially to an HP StoreOnce store.

- **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 VMware disaster recovery.
  - HP StoreOnce Catalyst store pairs can be configured with the HP Data Protector Mirror feature to replicate data between Catalyst pairs. This provides recovery for local disk failures, complete VM failures or complete site failures by keeping VMware VM image copies at local and remote sites.

**Conclusion**

VMware 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. Combining HP StoreOnce Backup systems with leading virtualization products provides a comprehensive data protection solution for VMware environments.
Useful links
Protecting VMware Virtual Machines with HP StoreOnce D2D systems and the VMware vStorage APIs for Data Protection
HP StoreOnce Backup
HP StoreOnce Backup system user guide
HP StoreOnce Backup systems Linux and UNIX configuration guide
HP StoreOnce Backup systems best practices guide
HP StoreOnce Catalyst Solution Service
Technical Datasheet
Customer Brochure
HP Data Protector Administrator’s guide

Learn more at
hp.com/go/StoreOnce