

HP 3PAR StoreServ Architecture



Table of contents

HP 3PAR StoreServ hardware architecture overview.....	4
HP 3PAR ASIC.....	4
Full-mesh controller backplane.....	5
Active/Active vs. Mesh-Active.....	6
System-wide striping.....	7
Controller node architecture.....	7
Drive enclosures.....	10
Highly virtualized storage operating system.....	10
Fine-grained approach to virtualization.....	10
Multiple layers of abstraction.....	10
Logical disks.....	11
Common provisioning groups.....	12
Virtual volumes.....	12
VLUNs and LUN masking.....	13
Metadata handling.....	13
System-wide sparing.....	14
Flash-optimized innovations.....	14
Workload-centric Storage Personas.....	15
HP 3PAR Block Persona.....	15
HP 3PAR File Persona.....	15
Multi-tenant architecture benefits.....	16
Tier-1 resiliency to support multi-tenancy.....	16
Hardware and software fault tolerance.....	16
Advanced fault isolation and RAID protection.....	17
Controller node redundancy.....	17
Data integrity checking.....	18
Memory fencing.....	18
Persistent technologies.....	19
HP 3PAR Persistent Checksum.....	19
HP 3PAR Persistent Cache.....	19
HP 3PAR Persistent Ports.....	20
HP 3PAR Remote Copy.....	20
HP 3PAR Virtual Domains.....	21
Data encryption.....	21

Maintaining high and predictable performance levels	22
Load balancing	22
Mixed-workload support.....	22
Storage quality of service	23
Adaptive Flash Cache.....	23
Express Writes.....	23
Performance benefits of system-wide striping	24
Bandwidth and communication	24
Data transfer paths.....	25
Sharing and offloading of cached data	26
Pre-fetching.....	26
Write caching	26
Fast RAID 5.....	27
HP 3PAR RAID MP (Fast RAID 6)	27
Data compaction	27
HP 3PAR ASIC for inline data reduction.....	28
HP 3PAR Thin Deduplication and Thin Clones.....	28
HP 3PAR Thin Provisioning	29
HP 3PAR Thin Persistence.....	29
HP 3PAR Thin Conversion	29
HP 3PAR Thin Copy Reclamation	29
HP 3PAR Virtual Copy	30
Autonomic storage management.....	30
Ease of use.....	30
Data optimization.....	31
Autonomic rebalancing	32
Proactive support.....	32
HP 3PAR Storage Federation.....	33
HP 3PAR Peer Motion	34
HP 3PAR Online Import	34
HP 3PAR Peer Persistence.....	35
Summary.....	36
For more information	37

Modern, Tier-1 Storage for the New Style of IT

HP 3PAR StoreServ Storage provides a single product family to meet your primary storage needs—regardless of whether you are a midsize enterprise experiencing rapid growth, a large enterprise looking to support IT as a service (ITaaS), or a global service provider building a hybrid or private cloud.

HP 3PAR StoreServ Storage supports true convergence of block, file, and object access while offering the performance and flexibility that you need to accelerate new application deployments and support server virtualization, the cloud, ITaaS, and your future technology initiatives. It's a storage platform that allows you to spend less time on management, gives you technically advanced features for less money, and eliminates trade-offs that require you to sacrifice critical capabilities such as performance and scalability. With HP 3PAR StoreServ Storage, you can serve unpredictable and mixed workloads, support unstructured and structured data growth, and meet block, file, and object access storage needs from a single capacity store.

Figure 1. Next Gen HP 3PAR StoreServ Portfolio



The all-flash HP 3PAR StoreServ models—20850, 7450c—offer massive performance scalability from hundreds of thousands of IOPS to millions of IOPS all within a sub-millisecond response time. These all-flash systems uniquely offer performance without any compromises, be it the enterprise class tier 1 data services (such as 3-data-center replication) or resiliency using their scalable (>2-nodes) architecture. These systems are well suited for massive workloads consolidation, and should be the first choice for customers building their next generation (all-flash) data centers.

The converged flash HP 3PAR StoreServ models—20800, 7440c, 7400c, 7200c—offer flash performance with the additional flexibility to add spinning media (HDD) as a tier. In an all-flash configuration, these systems also deliver hundreds of thousands to millions of IOPS all within a sub-millisecond response time. 20800 and 7440c are hyper-scalable models, supporting multi-PiB raw capacity scalability. 7400c and 7200c balance performance and scalability and offer a very economic starting point for customers looking to take advantage of the HP 3PAR architecture.

It is important to remember that all HP 3PAR StoreServ models are built on a single architecture, run the exact same HP 3PAR Operating System and offer a common set of enterprise data services, be it the 7200c or the 20850. They can natively replicate and federate amongst each other without the need for any external replication or virtualization appliance.

This white paper describes the architectural elements of the HP 3PAR StoreServ Storage family that deliver Tier-1 resiliency across midrange, all-flash, and tiered storage arrays, as well as by supporting a single operating system and rich set of data services across the portfolio.

HP 3PAR StoreServ hardware architecture overview

Each HP 3PAR StoreServ Storage system features a high-speed, full-mesh passive backplane that joins multiple controller nodes (the high-performance data movement engines of the HP 3PAR StoreServ Architecture) to form a cache-coherent, Mesh-Active cluster. This low-latency interconnect allows for tight coordination among the controller nodes and a simplified software model.

In every HP 3PAR StoreServ Storage system, each controller node has a dedicated link to each of the other nodes that operates at 4 GB/s in each direction. In an HP 3PAR StoreServ 20800 Storage system, a total of 28 of these links form the array's full-mesh backplane. In addition, each controller node may have one or more paths to hosts—either directly or over a storage area network (SAN). The clustering of controller nodes enables the system to present hosts with a single, highly available, high-performance storage system. This means that servers can access volumes over any host-connected port—even if the physical storage for the data is connected to a different controller node. This is achieved through an extremely low-latency data transfer across the high-speed, full-mesh backplane.

The HP 3PAR StoreServ Architecture can be scaled from 1.2 TiB to 6 PiB of raw capacity, making the system deployable as a small, remote office system, or a very large centralized system. Until now, enterprise customers were often required to purchase and manage at least two distinct primary architectures to span their range of cost and scalability requirements. HP 3PAR StoreServ Storage is the ideal platform for virtualization and cloud computing environments. The high performance and scalability of the HP 3PAR StoreServ Architecture is well suited for large or high-growth projects, consolidation of mission-critical information, demanding performance-based applications, and data lifecycle management. High availability is also built into the HP 3PAR StoreServ Architecture through full hardware redundancy. Controller node pairs are connected to dual-ported drive enclosures owned by that pair. Unlike other approaches, the system offers both hardware and software fault tolerance by running a separate instance of the HP 3PAR Operating System on each controller node, thus facilitating the availability of customer data. With this design, software and firmware failures—a significant cause of unplanned downtime in other architectures—are greatly reduced.

HP 3PAR ASIC

The HP 3PAR ASICs in each controller node feature a uniquely efficient, silicon-based zero-detection and deduplication mechanism that gives HP 3PAR StoreServ Storage systems the power to perform inline deduplication and remove allocated but unused space without impacting performance. The HP 3PAR ASICs also deliver mixed-workload support to alleviate performance concerns and cut traditional array costs.

Transaction- and throughput-intensive workloads run on the same storage resources without contention, thereby cutting array purchases in half. This is particularly valuable in virtual server environments, where HP 3PAR StoreServ Storage boosts virtual machine density so you can cut physical server purchases.

The HP 3PAR StoreServ 20000 systems use the fifth and latest generation of the HP 3PAR ASIC, the HP 3PAR Thin Express ASIC. The HP 3PAR Thin Express ASIC is engineered and designed for solid-state performance. The ASIC enables the new 20000 series to deliver over 5X improvement in system bandwidth and faster XOR operations. It offloads CPU and evenly spreads I/O workload across 8 Node Active—Mesh scale—out architecture, ensuring lower latency. The HP 3PAR Thin Express ASIC comes with a new powerful data deduplication engine which powers inline deduplication for block workloads without compromising performance or scale.

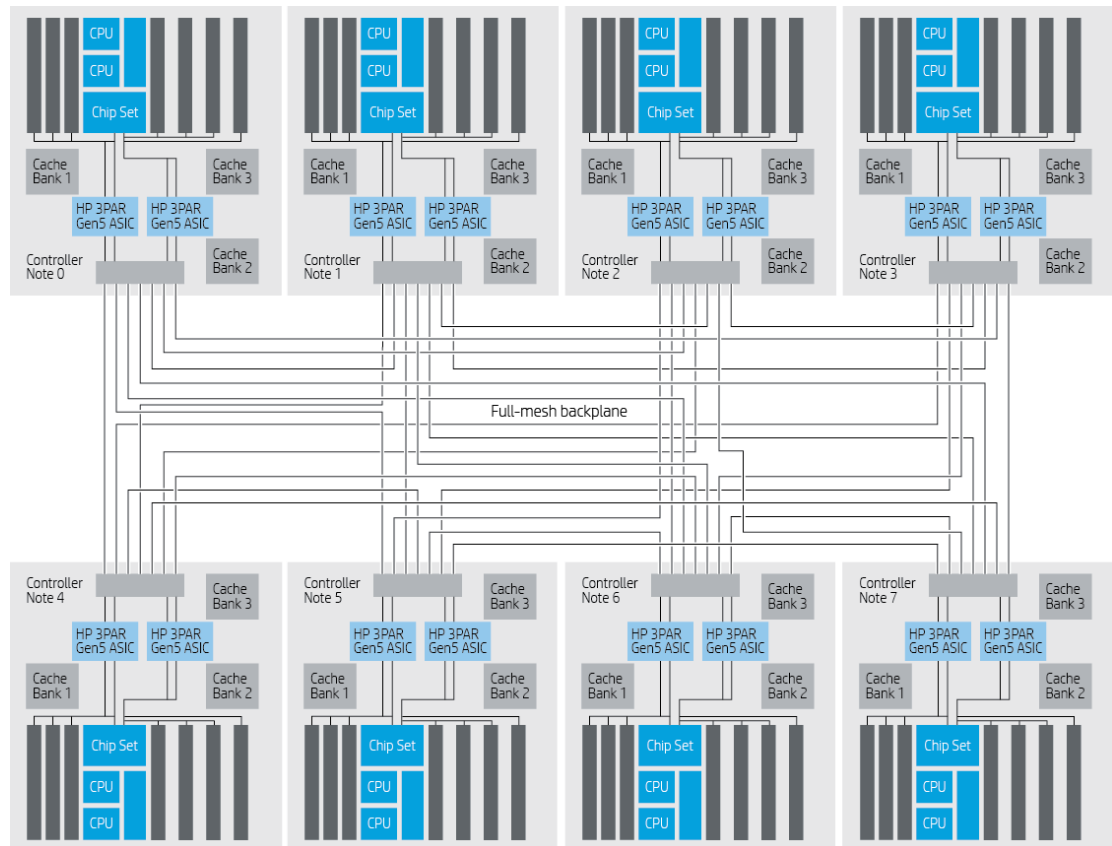
The Thin Express ASIC also enables Persistence Checksum that delivers T10-PI (Protection Information) for end-to-end data protection (against media and transmission errors) with no impact to applications or host operating systems.

Full-mesh controller backplane

Backplane interconnects within servers have evolved dramatically over the years. Most, if not all, server and storage array architectures have traditionally employed simple bus-based backplanes for high-speed processor, memory, and I/O communication. Parallel to the growth of SMP-based servers, significant investments were also made to switch architectures, which have been applied to one or two enterprise storage arrays.

The move from buses to switches was intended to address latency issues across the growing number of devices on the backplane (more processors, larger memory, and I/O systems). Third-generation full-mesh interconnects first appeared in the late 1990s in enterprise servers. However, HP 3PAR StoreServ Storage represents the first storage platform that applies this interconnect, which has been incorporated into HP 3PAR StoreServ Storage to reduce latencies and address scalability requirements. Figure 2 shows the full-mesh backplane of an HP 3PAR StoreServ 20000 Storage system.

Figure 2. Full-mesh backplane of an HP 3PAR StoreServ 20000 Storage system



The HP 3PAR StoreServ full-mesh backplane is a passive circuit board that contains slots for up to eight controller nodes, depending on the model. As noted earlier, each controller node slot is connected to every other controller node slot by a high-speed link (4 GB/s in each direction, or 8 GB/s total), forming a full-mesh interconnect between all controller nodes in the cluster—something that HP refers to as a Mesh-Active design. These interconnects deliver low-latency, high-bandwidth communication and data movement between controller nodes through dedicated point-to-point links and a low overhead protocol that features rapid inter-node messaging and acknowledgment. It's important to note that, while the value of these interconnects is high, the cost of providing them is relatively low. Because these interconnects are passive and consists of static connections embedded within a printed circuit board, this innovation does not represent a large cost within the overall system and only one is needed. In addition, a completely separate full-mesh network of RS-232 serial links provides a redundant low-speed channel of communication for exchanging control information between the nodes.

The HP 3PAR StoreServ 20000 features an eight node capable backplane that supports two to eight controller nodes. HP 3PAR StoreServ 7000 Storage systems feature either a dual-node or quad-node-capable systems that is essentially an equivalent of what was used in erstwhile enterprise class arrays which offer the same high-speed links between nodes.

Active/Active vs. Mesh-Active

Most traditional array architectures fall into one of two categories: monolithic or modular. In a monolithic architecture, being able to start with smaller, more affordable configurations (i.e., scaling down) presents challenges. Active processing elements not only have to be implemented redundantly, but they are also segmented and dedicated to distinct functions such as host management, caching, and RAID/drive management. For example, the smallest monolithic system may have a minimum of six processing elements (one for each of three functions, which are then doubled for redundancy of each function). In this design—with its emphasis on optimized internal interconnectivity—users gain the Active/Active processing advantages of a central global cache (e.g., LUNs can be coherently exported from multiple ports). However, these architectures typically involve higher costs relative to modular architectures.

In traditional modular architectures, users are able to start with smaller and more cost-efficient configurations. The number of processing elements is reduced to just two, because each element is multifunction in design—handling host, cache, and drive management processes. The trade-off for this cost-effectiveness is the cost or complexity of scalability. Because only two nodes are supported in most designs, scale can only be realized by replacing nodes with more powerful node versions or by purchasing and managing more arrays. Another trade-off is that dual-node modular architectures, while providing failover capabilities, typically do not offer truly Active/Active implementations where individual LUNs can be simultaneously and coherently processed by both controllers.

The HP 3PAR StoreServ Architecture was designed to provide cost-effective single-system scalability through a cache-coherent, multi-node clustered implementation. This architecture begins with a multifunction node design and, like a modular array, requires just two initial controller nodes for redundancy. However, unlike traditional modular arrays, enhanced direct interconnects are provided between the controllers to facilitate Mesh-Active processing. Unlike legacy Active/Active controller architectures—where each LUN (or volume) is active on only a single controller—this Mesh-Active design allows each LUN to be active on every controller in the system, thus forming a mesh. This design delivers robust, load-balanced performance and greater headroom for cost-effective scalability, overcoming the trade-offs typically associated with modular and monolithic storage arrays.

System-wide striping

Through a Mesh-Active design and system-wide striping, the HP 3PAR StoreServ Architecture can provide the best of traditional modular and monolithic architectures in addition to massive load balancing.

The HP Mesh-Active design not only allows all volumes to be active on all controllers, but also promotes system-wide striping that autonomically provisions and seamlessly stripes volumes across all system resources to deliver high, predictable levels of performance. System-wide striping of data provides high and predictable levels of service for all workload types through the massively parallel and fine-grained striping of data across all internal resources (disks, ports, loops, cache, processors, etc.). As a result, as the use of the system grows—or in the event of a component failure—service conditions remain high and predictable. Unlike application-centric approaches to storage, HP 3PAR StoreServ Storage provides autonomic rebalancing that enables the system to evenly balance and use all available physical resources. This is particularly important with hardware upgrades since existing data should be rebalanced and stripped across new available resources. On HP 3PAR StoreServ Storage, this is done without service disruption or preplanning.

For flash-based media, fine-grained virtualization combined with system-wide striping drives uniform I/O patterns by spreading wear evenly across the entire system. Should there be a media failure, system-wide sparing also helps guard against performance degradation by enabling a many-to-many rebuild, resulting in faster rebuilds. Because HP 3PAR StoreServ Storage autonomically manages this system-wide load balancing, no extra time or complexity is required to create or maintain a more efficiently configured system.

A detailed discussion of resource allocation, including the system's virtualized tri-layer mapping methodology, is provided in the section "[Highly virtualized storage operating system](#)".

Controller node architecture

An important element of the HP 3PAR StoreServ Architecture is the controller node, a proprietary and powerful data movement engine that is designed for mixed workloads. As noted earlier, a single system, depending on the model, is modularly configured as a cluster of two to eight controller nodes. This modular approach provides flexibility, a cost-effective entry footprint, and affordable upgrade paths for increasing performance, capacity, connectivity, and availability as needs change. In addition, the minimum dual-controller configuration means that the system can withstand an entire controller node failure without impacting data availability. Controller nodes can be added in pairs to the cluster non-disruptively, and each node is completely hot-pluggable to enable online serviceability.

Unlike legacy architectures that process I/O commands and move data using the same processor complex, the HP 3PAR StoreServ Storage controller node architecture separates the processing of control commands from data movement, which helps ensure that CPU memory bandwidth is available for control processing and is not used for bulk data transfer. This innovation eliminates the performance bottlenecks of existing platforms that use a single processing element to serve competing workloads, for example online transaction processing (OLTP) and data warehousing workloads.

The HP 3PAR ASIC within each controller node performs parity calculations (for RAID 5 and RAID MP/Fast RAID 6) on the data cache. The zero-detection mechanism built into the ASIC allows a hardware-assisted fat-to-thin volume conversion in conjunction with HP 3PAR Thin Conversion software that enables users to take "fat" provisioned volumes on legacy storage and convert them to "thin" provisioned volumes on the system, inline and non-disruptively. This zero-detection capability also removes streams of zeroes present in I/O prior to writing data to the back-end storage system in order to reduce capacity requirements and prolong SSD life span. The HP 3PAR ASIC is also a crucial element of the system's ability to perform inline, block-level Thin Deduplication with Express Indexing (see the "[Thin Deduplication with Express Indexing](#)" section for more details).

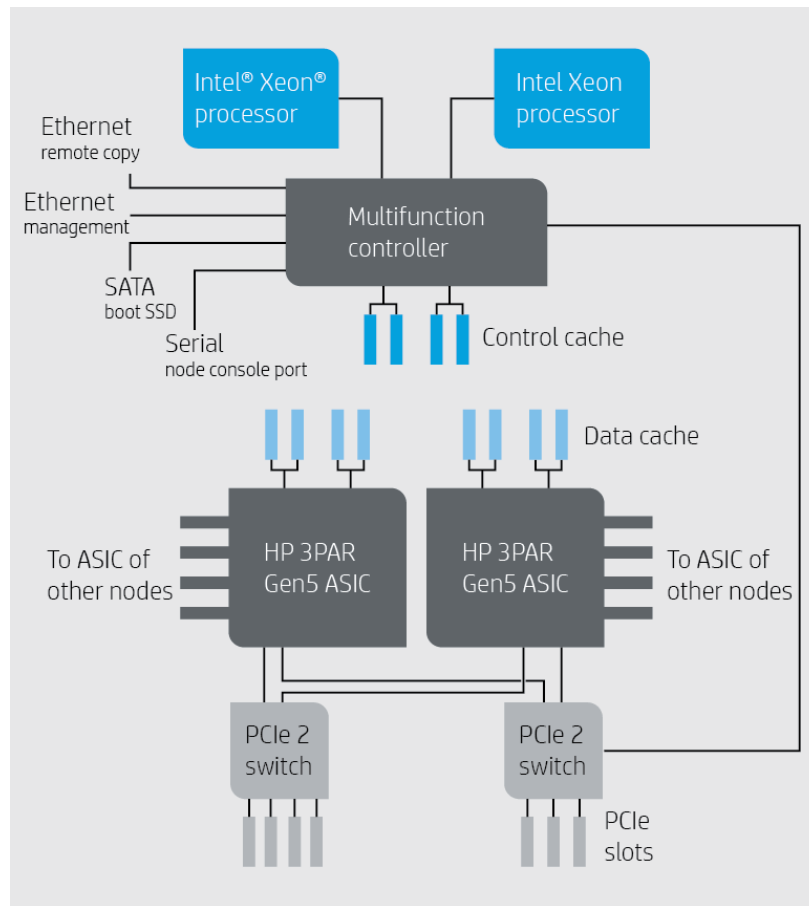
HP 3PAR StoreServ 20800 offers scalability in its truest sense not only with support for 1,920 drives and 6 PiB raw capacity but also with a port count of up to 160 FC ports for host connectivity and federation. Each of these ports is connected directly on the I/O bus, so all ports can achieve full bandwidth up to the limit of the I/O bus bandwidths that they share. Furthermore, each HP 3PAR StoreServ 20000 series controller node can have a minimum of the following adapters:

- (20) Quad-port 16 Gbps Fibre Channel adapter
- (10) Dual-port 10 Gbps iSCSI or Fibre Channel over Ethernet (FCoE) converged network adapter
- (6) 10 Gbps Ethernet adapter for file and object access services using File Persona

With up to a total of 160 host ports available on an 8-controller configuration, HP 3PAR StoreServ 20000 series systems offer abundant multiprotocol connectivity. For back-end connectivity, each node can have up to (3) Quad-port 12 Gbps SAS Adapter.

Figure 3 shows the controller node design of the HP 3PAR StoreServ 20000 Storage system.

Figure 3. Controller node design of the HP 3PAR StoreServ 20000 Storage system

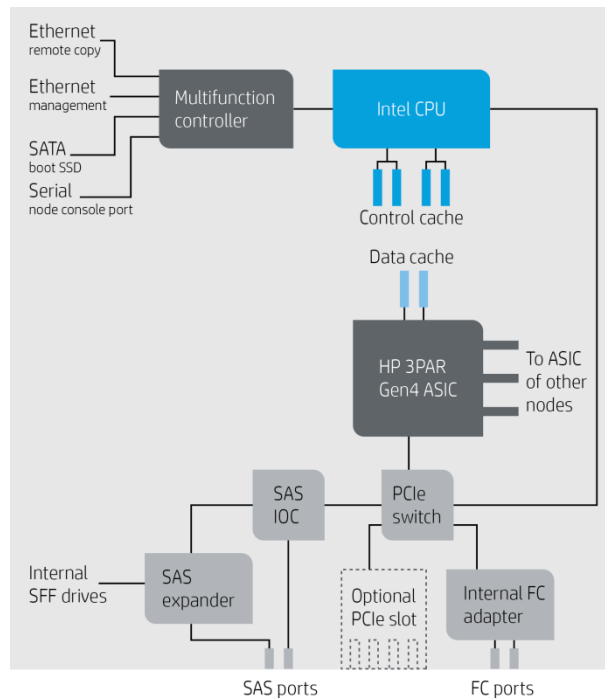


Each HP 3PAR StoreServ 7000-series controller node has two built-in 8 Gbps Fibre Channel ports and one PCIe expansion slot. This slot can hold one of the following adapters:

- Quad-port 8 GB/s Fibre Channel adapter
- Dual-port 16 GB/s Fibre Channel adapter
- Dual-port 10 Gbps iSCSI or Fibre Channel over Ethernet (FCoE) converged network adapter
- Dual port 10 Gb/s Ethernet adapter for file and object access services using HP 3PAR File Persona
- Quad port 1 Gb/s Ethernet adapter for file and object access services using HP 3PAR File Persona

With up to 24 ports available on a quad-controller configuration, HP 3PAR StoreServ 7000-series systems offer abundant multiprotocol connectivity. For back-end connectivity, each node has two built-in 2 x 4-lane 6 Gbps SAS ports. Figure 4 shows the controller node design of the HP 3PAR StoreServ 7000 Storage system.

Figure 4. Controller node design of the HP 3PAR StoreServ 7000 Storage system



Across all 7000-series models, this controller node design extensively leverages commodity parts with industry-standard interfaces to achieve low costs and keep pace with industry advances and innovations. At the same time, the HP 3PAR ASICs add crucial bandwidth and direct communication pathways without limiting the ability to use industry-standard parts for other components. Processor specifications by HP 3PAR StoreServ model are shown in Table 1.

Table 1. Processor specifications by HP 3PAR StoreServ Storage system model

Model	CPUs	Controller nodes	Total on-node cache	Total cache including optional Flash Cache
7200c	Up to 2 Intel® six-core processors	2	Up to 40 GB	768 GB
7400c	Up to 4 Intel six-core processors	2 or 4	Up to 96 GB	1596 GB
7440c	Up to 4 Intel eight-core processors	2 or 4	Up to 192 GB	3192 GB
7450c	Up to 4 Intel eight-core processors	2 or 4	Up to 192 GB	Not Applicable
20800	Up to 16 Intel six-core processors	2, 4, 6, or 8	Up to 1.8 TB	33.8 TB
20850	Up to 16 Intel eight-core processors	2, 4, 6, or 8	Up to 3.6 TB	Not Applicable

Drive enclosures

Another key element of the HP 3PAR StoreServ Storage system is the drive enclosure or drive chassis, which serves as the capacity building block within the system. This section looks in more detail at the different drive enclosures supported by HP 3PAR StoreServ Storage systems.

A single unit of HP 3PAR StoreServ 20800 system can support up to 1920 drives (with 1024 of these drives being SSDs), totaling to 6 PiB of raw capacity. HP 3PAR StoreServ 20850 Storage systems support up to 1024 SSD totaling to a raw capacity of 3.9 PiB of raw capacity.

HP 3PAR StoreServ 20000 Storage systems support two different form factors. 2U 12 drives LFF and 2U 24 drives SFF. All together offering a support of total of 1024 SFF drives and 512 LFF drives with HP 3PAR StoreServ 20850 Storage system and 1920 SFF or 960 LFF drives with HP 3PAR StoreServ 20800 Storage system.

The HP 3PAR StoreServ 7000 Storage systems offer ample capacity expansion options in a dense form factor, accommodating 24 small form factor (SFF) HDDs or SSDs in a dual-controller storage base system. A quad-controller HP 3PAR StoreServ 7400c, 7440c, or 7450c Storage system can hold a combination of up to 48 SFF HDDs and SSDs (SSDs only in the HP 3PAR StoreServ 7450c Storage system).

Additional capacity expansion options for HP 3PAR StoreServ 7000 Storage systems are available in the form of a 2U, 2.5-inch SFF drive enclosure that can hold 24 SFF HDDs or SSDs or a 4U, 3.5-inch large form factor (LFF) drive enclosure that can hold 24 LFF HDDs or SSDs (no HDD support for HP 3PAR StoreServ 7450c All-Flash, supports SSDs only). The HP 3PAR StoreServ 7200c Storage system can accommodate nine additional drive enclosures and a maximum of 240 HDDs with up to 120 SSDs, while the HP 3PAR StoreServ 7400c Storage system can accommodate 22 additional drive enclosures and a maximum of 576 HDDs with up to 240 SSDs. The HP 3PAR StoreServ 7440c Storage system can scale up to 38 additional drive enclosures and a maximum of 960 HDDs with up to 240 SSDs. The HP 3PAR StoreServ 7450c can scale up to 18 additional enclosures and 240 SSDs. A StoreServ 7000 Storage system can be configured with both SFF and LFF drives.

Highly virtualized storage operating system

HP 3PAR StoreServ Storage uses the same highly virtualized storage operating system across all models—including high-end, midrange, hybrid, and all-flash arrays. To help ensure performance and improve the utilization of physical resources, the HP 3PAR Operating System employs a tri-level mapping methodology with three layers of abstraction that is similar to the virtual memory architectures of the most robust enterprise operating systems on the market today.

Fine-grained approach to virtualization

The tri-level mapping methodology imposed by the HP 3PAR Operating System relies on a fine-grained virtualization approach that divides each physical disk into granular allocation units referred to as chunklets, each of which can be independently assigned and dynamically reassigned to different logical disks that are used to create virtual volumes. The first layer of abstraction employed by the OS breaks media devices into 1 GB chunklets to enable higher utilization and avoid stranded capacity. This fine-grained virtualization unit also enables mixed RAID levels on the same physical drive, thereby eliminating dedicated RAID groups and seamlessly supporting new media technologies such as SSDs.

Multiple layers of abstraction

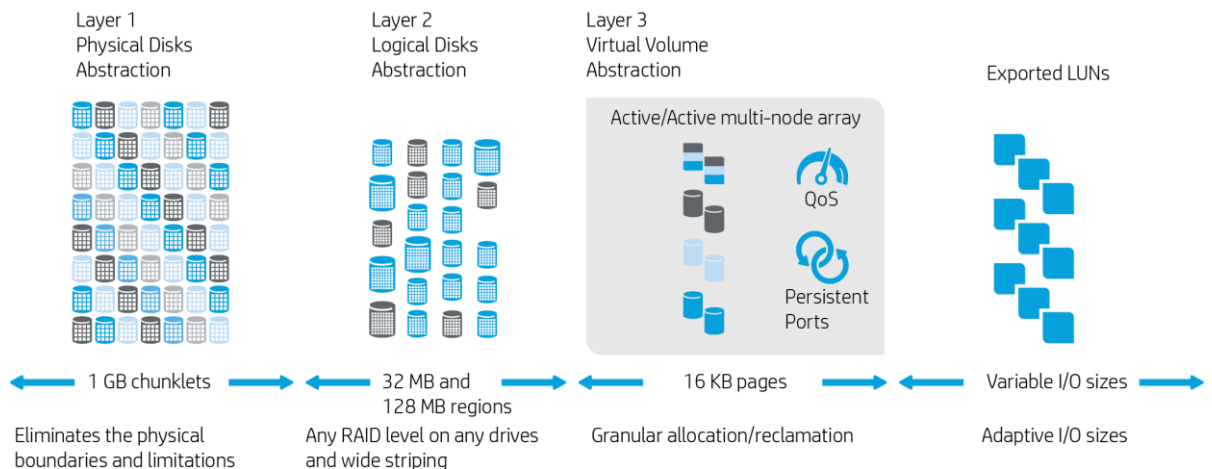
As shown in Figure 5, the physical disk abstraction layer breaks physical drives of any size into a pool of uniform-sized, 1 GB virtual chunklets. The fine-grained nature of these chunklets eliminates underutilization of precious storage assets.

Complete access to every chunklet eliminates large pockets of inaccessible storage. This fine-grained structure enhances performance for all applications as well, regardless of their capacity requirements. For example, while a small application might only allocate a small amount of physical capacity, this capacity will be virtualized and striped across dozens or even hundreds of drives. With this approach, even a small application can leverage the performance resources of the entire system without provisioning excess capacity.

The second layer of abstraction takes the 1 GB chunklets created from abstracting physical disk capacity and creates logical disks (LDs) striped across the system’s physical drives and implementing specified RAID levels. Multiple chunklet RAID sets from different PDs are striped together to form an LD. All chunklets belonging to a given LD will be from the same drive type. LDs can consist of all NL, FC, or SSD chunklets. There are no mixed-type LDs, with the exception of Fast Class (Fibre Channel or SAS) LDs, where the LD may consist of mixed 10K and 15K drive chunklets. The association between chunklets and LDs allows LDs to be created with template properties based on RAID characteristics and the location of chunklets across the system. LDs can be tailored to meet a variety of cost, capacity, performance, and availability characteristics. In addition, the first- and second-level mappings taken together serve to parallelize work massively across physical drives and their Fibre Channel or SAS connections. LDs are divided into “regions,” which are 128 MB of contiguous logical space. Virtual volumes (VVs) are composed of these LD regions, with VV space allocated across these regions.

The third layer of abstraction maps LDs to VVs, with all or portions of multiple underlying LDs mapped to the VV. VVs are the virtual capacity representations that are ultimately exported to hosts and applications as virtual LUNs (VLUNs) over Fibre Channel, iSCSI, or FCoE target ports. A single VV can be coherently exported through as many or as few ports as desired. This layer of abstraction uses a table-based association—a mapping table with a granularity of 32 MB per region and an exception table with a granularity of 16 KB per page—as opposed to an algorithmic association. With this approach, a very small portion of a VV associated with a particular LD can be quickly and non-disruptively migrated to a different LD for performance or other policy-based reasons, whereas other architectures require migration of the entire VV. This layer of abstraction also implements many high-level features such as snapshots, caching, pre-fetching, and remote replication.

Figure 5. Virtualization with a tri-level mapping methodology that provides three layers of abstraction



One-stop allocation, the general method employed by IT users for volume administration, requires minimal planning on the part of storage administrators. By an administrator simply specifying virtual volume name, RAID level, and size, the HP 3PAR Operating System autonomously provisions LDs at the moment that an application requires capacity. This process is also known as “just-in-time” provisioning. Contrast this to traditional architectures where the storage administrator must assign physical disks to RAID sets when the array is installed, which can be difficult to change later on and makes it difficult to respond to changing requirements.

The three-layer abstraction imposed by the HP 3PAR Operating System can effectively utilize any underlying media type. This means that HP 3PAR StoreServ Storage is able to make the most efficient use of SSDs through massive load balancing across all drives to enable ultra-high performance and prolong flash-based media lifespan.

Logical disks

There are three types of logical disks (LDs):

- User (USR) LDs provide user storage space to fully provisioned VVs.
- Snapshot data (SD) LDs provide the storage space for snapshots (or virtual copies), thinly provisioned (TPVV) and thinly deduplicated (TDVV) virtual volumes.
- Snapshot administration (SA) LDs provide the storage space for metadata used for snapshot and TPVV and TDVV administration.

As mentioned earlier, RAID functionality is implemented at the LD level, with each LD mapped to chunklets in order to implement RAID 1+0 (mirroring + striping), RAID 5+0 (RAID 5 distributed parity + striping), or RAID MP (multiple distributed parity, with striping).

The HP 3PAR Operating System will automatically create LDs with the desired availability and size characteristics. In addition, several parameters can be used to control the layout of an LD to achieve these different characteristics:

- **Set size:** The set size of the LD is the number of drives that contain redundant data. For example, a RAID 5 LD may have a set size of 4 (3 data + 1 parity), or a RAID MP LD may have a set size of 16 (14 data + 2 parity). For a RAID 1 LD, the set size is the number of mirrors (usually 2). The chunklets used within a set are typically chosen from drives on different enclosures. This helps ensure that a failure of an entire loop (or enclosure) will not result in data becoming unavailable until the drive enclosure is repaired. It also helps ensure better peak aggregate performance because data can be accessed in parallel on different loops.
- **Step size:** The step size is the number of bytes that are stored contiguously on a single physical drive.
- **Row size:** The row size determines the level of additional striping across more drives. For example, a RAID 5 LD with a row size of 2 and set size of 4 is effectively striped across 8 drives.
- **Number of rows:** The number of rows determines the overall size of the LD given a level of striping. For example, an LD with 3 rows, with each row having 6 chunklets' worth of usable data (+2 parity), will have a usable size of 18 GB (1 GB/chunklet x 6 chunklets/row x 3 rows).

An LD also has an "owner" and a "backup owner." The owner is the controller node that, under normal circumstances, performs all operations on the LD. The chunklets used to create LDs and always belong to physical drives that have an active path to that controller. If the owner fails the secondary path, the drives become active and the backup owner takes over ownership of the LD. The owner sends sufficient log information to the backup owner so that the backup owner can take over without loss of data.

Common provisioning groups

A common provisioning group (CPG) creates a virtual pool of LDs that allows VVs to share the CPG's resources and allocate space on demand. You can create fully provisioned VVs and TPVVs that draw space from the CPG's logical disk pool.

CPGs enable fine-grained, shared access to pooled logical capacity. Instead of pre-dedicating logical disks to volumes, a CPG allows multiple volumes to share the buffer pool of LDs. For example, when a TPVV is running low on user space, the system automatically assigns more capacity to the TPVV by mapping new regions from LDs in the CPG to the TPVV. As a result, any large pockets of unused but allocated space are eliminated. Fully provisioned VVs cannot create user space automatically, and the system allocates a fixed amount of user space for the volume.

Virtual volumes

There are two kinds of VVs: "base volumes" and "snapshot volumes." A base volume can be considered to be the "original" VV and is either a fully provisioned virtual volume or a thinly provisioned virtual volume. In other words, it directly maps all the user-visible data. A snapshot volume is created using HP 3PAR Virtual Copy software. When a snapshot is first created, all of its data is mapped indirectly to the parent volume's data. When a block is written to the parent, the original block is copied from the parent to the snapshot data space and the snapshot points to this data space instead. Similarly, when a block is written in the snapshot, the data is written in the snapshot data space and the snapshot points to this data space. These snapshots are copy-on-write (COW) snapshots.

VVs have three types of space:

- The **user space** represents the user-visible size of the VV (i.e., the size of the SCSI LUN seen by a host) and contains the data of the base VV.
- The **snapshot data space** is used to store modified data associated with snapshots. The granularity of snapshot data mapping is 16 KB pages.
- The **snapshot admin space** is used to save the metadata (including the exception table) for snapshots.

Each of the three space types is mapped to LDs. One or more controller nodes may own these LDs; thus, VVs can be striped across multiple nodes for additional load balancing and performance.

The size limit for an individual virtual volume is 16 TB. A VV is classified by its provisioning type, which can be one of the below types:

- **Fully provisioned VV** (FPVV): Has either no snapshot space or deprecated, statically allocated snapshot space.
- **Thinly provisioned VV** (TPVV): TPVV has space for the base volume allocated from the associated CPG and snapshot space allocated from the associated snapshot CPG (if any).

On creation, a TPVV the size of the VV is specified, but no storage is allocated. Storage is allocated on demand in the snapshot data area as required by the host operation being performed. The snapshot admin area contains the metadata indexes that point to the user data in the SD area. Because the SA metadata needs to be accessed to locate the user data, the indexes are cached in policy memory to reduce the performance impact of the lookups.

TPVVs associated with a common CPG share the same LDs and draw space from that pool as needed, allocating space on demand in small increments for each controller node. As the volumes that draw space from the CPG require additional storage, the HP 3PAR Operating System automatically extends existing LDs or creates new LDs until the CPG reaches the user-defined growth limit, which restricts the CPG's maximum size.

- **Thinly deduped VV** (TDVV): TDVVs behave similarly to TPVV volumes with the fundamental difference that TDVVs within the same CPG will share common pages of data. The data shared is determined via the inline deduplication mechanism described later in this paper. TDVVs are supported only on CPGs that use SSDs as a tier of storage.
- **Commonly provisioned VV** (CPVV): The space for this VV is fully provisioned from the associated CPG, and the snapshot space is allocated from the associated snapshot CPG.

VLUNs and LUN masking

VVs are only visible to a host once the VVs are exported as VLUNs.

VVs can be exported in three ways:

- To specific hosts (set of Worldwide Names or WWNs)—the VV is visible to the specified WWNs, regardless of which port(s) those WWNs appear on. This is a convenient way to export VVs to known hosts.
- To any host on a specific port—this is useful when the hosts (or the WWNs) are not known prior to exporting, or in situations where the WWN of a host cannot be trusted (host WWNs can be spoofed).
- To specific hosts on a specific port.

Metadata handling

For metadata, HP 3PAR StoreServ implements a mechanism of fast lookup tables that store location pointers to accelerate data access. This process relies on tri-layer addresses translation mechanism akin to Virtual Memory lookup tables. The VV metadata is also known as Snapshot Admin space, which contains bitmaps indicating which pages of the shared admin and data areas are used along with the exception tables associated with each volume.

The exception tables provide the offset into the shared data area where the shared data is located. There are three levels (L1, L2, and L3) of exception tables for each Thin Volume. The bits of the I/O request's logical block address (LBA) are used as indexes into the L1, L2, and L3 exception tables to determine the page being accessed.

Metadata management is shared across all cluster resources from CPU memory (control cache) by looking at L1/L2/L3 exception tables and copying to cluster memory. LRU (Last Recently Used) algorithms will eventually flush these pages to SA LDs.

When HP 3PAR Adaptive Flash Cache is configured, all Snapshot Admin space associated with any TPVV (or snapshot) on the array will always be also cached to Flash Cache.

For inline deduplication, where fast lookups are extremely important, HP 3PAR StoreServ implements an innovative mechanism to detect duplicate pages called Express Indexing. The technology uses the computed hash signature as an index to determine whether a match already exists using the three level translation mechanisms described earlier.

When new a write I/O request comes in, the Logical Block Address (LBA) is used as an index into three different page tables as per a regular TPVV. However, instead of allocating a new page, the hash signature of the incoming data page is computed by the HP 3PAR ASIC and compared to the signatures of the TDVV data already stored in the CPG.

If a match is found then the L3 page table entry will be set to point to the existing copy of the data page. Only if no match is found, a new page is allocated.

System-wide sparing

The HP 3PAR Operating System has a logical volume manager that handles volume abstraction at the VV layer and also handles sparing. This logical volume manager reserves a certain number of chunklets as spare chunklets depending on the sparing algorithm and system configuration. Unlike many competitive arrays that reserve dedicated spare drives that then sit idle, system-wide sparing with HP 3PAR StoreServ Storage means that spare chunklets are distributed across all drives. This provides additional protection and enables a balanced load that extends the SSD lifespan by providing even wearing. It also protects against performance degradation by enabling a “many-to-many” rebuild in the event of a failure.

Flash-optimized innovations

HP 3PAR StoreServ Mesh-Active Architecture, system-wide striping, fine grain virtualization, advanced metadata handling and system-wide sparing are just some of the pillars of HP 3PAR StoreServ architecture that enhance flash-based media. Flash-based media can deliver many times the performance of conventional spinning HDDs and it can do so at very low, sub-millisecond latency. However, it is important to understand that these advantages can only be realized by an architecture that has optimized its entire I/O path to be performance-centric. If the storage controllers that sit between servers and back-end flash devices can't keep up with the performance of the flash drives, they become performance bottlenecks.

To work with flash-based media in the most performance-optimized manner, the HP 3PAR StoreServ Architecture includes features designed to handle it in a substantially different way than spinning media. It also exploits every possible opportunity to extend flash-based media lifespan by reducing factors that contribute to media wear.

This flash-optimized architecture relies on several new and unique HP 3PAR StoreServ Storage innovations that accelerate performance and extend flash-based media lifespan:

- **Thin Deduplication with Express Indexing:** The system's Thin Deduplication software feature uses a hashing engine capability built into the HP 3PAR ASICs in combination with a unique Express Indexing feature to deduplicate data inline and with a high degree of granularity. Hardware-accelerated Thin Deduplication delivers a level of capacity efficiency that is superior to other approaches without monopolizing CPU resources and degrading performance, thereby delivering the only primary storage deduplication solution in the industry that is truly enterprise-class. ASIC-assisted, block-level deduplication takes place inline, which provides multiple benefits, including increasing capacity efficiency, protecting system performance, and extending flash media lifespan.
- **Adaptive Flash Cache:** This performance acceleration feature can as much as double IOPS and reduce latency by using SSD capacity to extend the system's cache. As a built-in system feature, it can be enabled on both all-flash arrays and systems with a dedicated SSD tier.
- **Express Writes:** Protocol optimization within the HP 3PAR OS enables HP 3PAR StoreServ to deliver lower CPU interrupts per I/O transaction, which results in higher IOPS and reduced latency for writes. Express Writes does not require any changes on the initiator side and all supported HBAs and Hosts will benefit of this optimization.
- **Adaptive Read and Write:** This feature matches host I/O size reads and writes to flash media at a granular level to avoid excess writes that cause unnecessary wear to flash media. Adaptive reads and writes also significantly reduce latency and enhance back-end performance to enable more applications to be consolidated.
- **Autonomic Cache Offload:** Automatically adjusts the frequency at which data is offloaded from cache to flash media based on utilization rates. This reduces cache bottlenecks by automatically changing the frequency at which data is offloaded from cache to flash media—based on utilization rate and without requiring any user intervention. This helps achieve consistently high performance levels as you scale workloads to hundreds of thousands of IOPS.
- **Multi-tenant I/O processing:** Multi-tenant I/O processing enables performance improvement for mixed workloads or virtual desktop infrastructure (VDI) deployments by breaking large I/O into smaller chunks so that small read requests don't get held up or stuck behind larger I/O requests, which also helps ensure reduced latency expected of flash-based media.
- **Adaptive Sparing:** Using patented Adaptive Sparing technology, HP has collaborated with SSD suppliers to extend usable capacity per drive by up to 20 percent. This is achieved by reducing capacity typically reserved by media suppliers for wear management and then using that space more efficiently. At a system level, increasing usable drive capacity also helps spread writes more broadly to extend SSD endurance.

For complete details of how the HP 3PAR Architecture is flash optimized both at the hardware and software layers, refer to the white paper [HP 3PAR StoreServ Storage: optimized for flash](#).

Workload-centric Storage Personas

HP 3PAR StoreServ Storage is expressed by workload-centric Storage Personas. A persona is *the aspect of one's character that is presented to or perceived by others*. Storage Personas are thus comprised of data access protocols and data services for the presentation of storage to hosts and clients. Specifically, HP 3PAR StoreServ features a Block Persona and a File Persona that are engineered into the core of the HP 3PAR OS and system architecture, and are managed seamlessly together via the HP 3PAR StoreServ Management Console and scriptable HP 3PAR CLI. Through these Storage Personas, HP 3PAR StoreServ provides truly converged block, file, and object access to simultaneously support an expanse of workloads while allowing the best storage approach to be employed for a given workload.

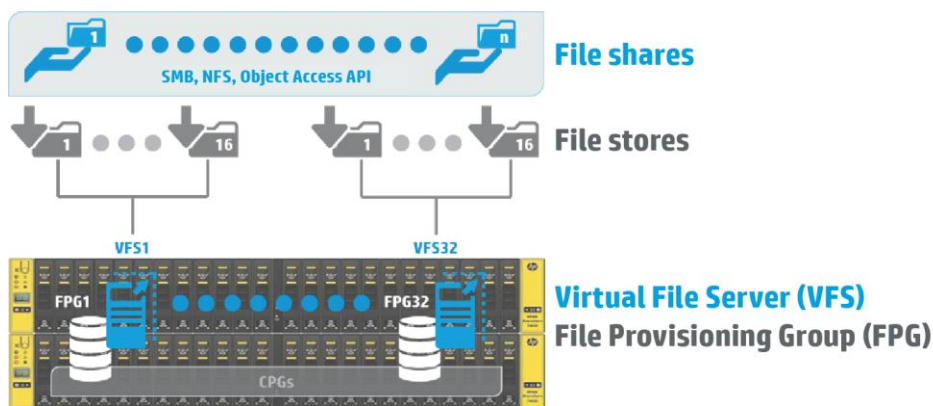
HP 3PAR Block Persona

HP 3PAR StoreServ Storage is expressed by the Block Persona in the form of block volumes to server hosts via Fibre Channel, iSCSI, and FCoE. The block data services are as mentioned throughout this white paper.

HP 3PAR File Persona

The HP 3PAR File Persona can be enabled on pairs of Converged Controller nodes configured with either 2-port 10 Gb or 4-port 1 Gb Ethernet adapters. The HP 3PAR File Persona is designed for client workloads such as home directory consolidation, group/department shares, corporate shares, and custom cloud applications by presenting file shares via SMB (CIFS) and NFS as well as object shares via the Object Access API to client devices. File data services include: User Authentication Services; capacity and user/group Quota Management; File Store Snapshots with user-driven file restore; and Antivirus Scan Services for integration with third-party antivirus software.

Figure 6. Logical view of HP 3PAR File Persona managed objects



As illustrated in Figure 6, the HP 3PAR File Persona has four types of managed objects. File Provisioning Groups, which are instances of the HP intellectual property Adaptive File System, control how files are stored and retrieved. Each File Provisioning Group is transparently constructed from one or multiple Virtual Volumes. Virtual File Servers (which present virtual IP addresses), participate in User Authentication Services, and can have properties for such things as user/group Quota Management. File Stores are the slice of a Virtual File Server and File Provisioning Group at which File Store Snapshots are taken, capacity Quota Management can be performed, and Antivirus Scan Services policies customized. File Shares are presented to clients via SMB, NFS, and the Object Access API, and are where share permissions are set. File Stores and File Provisioning Groups are typically only explicitly managed for advanced operations.

Multi-tenant architecture benefits

With HP 3PAR StoreServ Storage, you can securely partition resources within a shared infrastructure in order to pool physical storage resources for lower storage costs without compromising security or performance.

The HP 3PAR StoreServ Storage platform was built from the ground up to deliver multi-tenant resiliency that supports massive consolidation with ultra-high performance. The multi-controller scalability and extreme flexibility built into HP 3PAR StoreServ Storage makes deploying and maintaining separate storage silos to deliver different QoS levels a thing of the past. Unlike application-centric approaches to storage, one-click autonomic rebalancing on HP 3PAR StoreServ Storage enables you to enhance QoS levels without service disruption, pre-planning, or the need to purchase separate arrays to support different service levels. To support multiple tenants and workloads, HP 3PAR StoreServ Storage provides secure administrative segregation of users, hosts, and application data by using virtual machine technology. The following sections provide insight into the architectural elements that support each of these core capabilities.

Tier-1 resiliency to support multi-tenancy

HP 3PAR StoreServ Storage is designed to support massive consolidation by supporting mixed workloads and secure administrative segregation of users, hosts, and application data. Multi-tenancy allows IT organizations to deliver higher performance levels, greater availability, and next-generation functionality to multiple user groups and applications from a single storage system.

Today's IT realities—including virtualization, cloud computing, and ITaaS—demand the ability to deliver predictable service levels in an inherently unpredictable world, and make system resiliency the single most important requirement for multi-tenancy. Traditionally, Tier-1 storage has been characterized by hardware redundancy, advanced replication capabilities, and massive scalability in capacity and host connectivity. However, in order to enable the consolidation of multiple tenants onto a single system, hardware and software fault tolerance, as well the ability to predictably prevent downtime and handle failures in a way that is non-disruptive to users and applications, become critical. The HP 3PAR StoreServ Architecture supports multi-tenancy by allowing you to consolidate with confidence and achieve higher service levels for more users and applications with less infrastructure.

Hardware and software fault tolerance

HP 3PAR StoreServ Storage delivers Tier-1 resiliency with an architecture designed to eliminate any single point of failure (hardware or software) in the system. To mitigate single points of failure at the hardware layer, the system is designed with redundant components, including redundant power domains. In fact, to raise the bar with the fault tolerance mechanism, HP 3PAR StoreServ 20000 Storage system is configured with redundant boot drives (internal to the controller nodes) with two self-encrypting boot drives that work in redundant mode.

An independent copy of HP 3PAR OS runs on every controller node, so even in the smallest configuration, with two controller nodes, the system offers resiliency also for the software stack.

HP 3PAR StoreServ Storage components such as storage nodes, cache cards, disk- and host-facing host bus adapters (HBAs), power supplies, batteries, and disks all feature N+1 and in some cases N+2 redundancy so that any of these components can fail without system interruption. The only non-redundant component in the system is a 100 percent completely passive controller node backplane that, given its passive nature, is virtually impervious to failure. Return material authorization (RMA) and mean time between critical failures (MTBCF) hardware calculations include this component and substantiate this claim.

HP 3PAR StoreServ Storage offers up to four current load-balanced power distribution units (PDUs) per rack, which provide a minimum of two separate power feeds. The system can support up to four separate data center power feeds, providing even more power resiliency and further protection against power loss as well as brownouts. Redundant power domains help ensure that as many as two disk chassis power supplies can fail without power being lost to back-end disk devices.

Each controller node in an HP 3PAR StoreServ Storage system includes a local physical drive that contains a separate instance of the HP 3PAR Operating System as well as space to save cached write data in the event of a power failure.

The controller nodes are each powered by two (1+1 redundant) power supplies and backed up by a string of two batteries. Each battery has sufficient capacity to power the controller nodes long enough to save all necessary data in memory into the local physical drive. Although many architectures use "cache batteries," these are not suitable for extended downtimes that are usually associated with natural disasters and unforeseen catastrophes.

Another common problem with many battery-powered backup systems is that it is often impossible to ensure that a battery is charged and working. To address this problem, the HP 3PAR StoreServ Storage controller nodes are each backed by a string of at least two batteries. Batteries are periodically tested by discharging one battery while the other remains charged and ready in case a power failure occurs while the battery test is in progress. The HP 3PAR Operating System keeps track of battery charge levels and limits the amount of write data that can be cached based on the ability of the batteries to power the controller nodes long enough to save the data to the local drive.

The HP 3PAR StoreServ Storage controller node battery configuration also eliminates the need for expensive batteries to power all of the system's drive chassis. Note that, because all cached write data is mirrored to another controller node, a system-wide power failure would result in saving cached write data onto the internal drives of two nodes. Because each node's dual power supplies can be connected to separate AC power cords, providing redundant AC power to the system can further reduce the possibility of an outage due to an AC power failure.

Advanced fault isolation and RAID protection

Advanced fault isolation and high reliability are built into the HP 3PAR StoreServ Storage system. The drive chassis, drive magazines, and physical drives themselves all report and isolate faults. A drive failure will not take all drives offline.

HP 3PAR StoreServ Storage constantly monitors drives via the controller nodes and enclosures, and isolates faults to individual drives, then "offlines" only the failed component.

HP 3PAR StoreServ Storage is capable of RAID 1+0 (mirrored then striped), RAID 5+0 (RAID 5 distributed parity, striped in an X+1 configuration where X can be between 2 and 8), or RAID MP (multiple distributed parity, and currently striped with either a 6+2 or 14+2 configuration). All available RAID options allow HP 3PAR StoreServ Storage to create parity sets on different drives in different drive cages with separate power domains for greater integrity protection.

Each drive enclosure is divided into two redundant cages that plug into the drive chassis midplane. The system's drive chassis components—power supplies, Fibre Channel or SAS Adapters, drive magazines, and drives—are serviceable online and are completely hot-pluggable. Redundant power supply/fan assemblies hot-plug into the rear of the midplane. Should the drive chassis midplane fail for any reason, partner cage or cages will continue to serve data for those volumes that were configured and managed as "High Availability (HA) Cage" volumes. If the "HA Cage" configuration setting is selected at volume creation, the controller node automatically manages the RAID 1+0, RAID 5+0, or RAID MP data placement to accommodate the failure of an entire cage without affecting data access.

Each drive chassis includes N+1 redundant power supplies, redundant FC-AL adapters that provide up to four independent 4 Gbps, full-bandwidth Fibre Channel ports, and redundant cut-through switches on the midplane for switched point-to-point connections. Drive magazines are hot-pluggable from the front of the system into the midplane. Each Fibre Channel or SAS drive is dual ported and accessible from redundant incoming Fibre Channel connections in an Active/Passive mode. For HP 3PAR StoreServ 7000-Series systems, each drive chassis includes redundant power supplies, two SAS interface cards (IFCs) that provide up to four independent 6 Gbps SAS ports, and redundant cut-through switches.

Controller node redundancy

The HP 3PAR Operating System instance running on each of the controller nodes is both statefully managed and self-healing, providing protection across all cache-coherent, Mesh-Active storage controller nodes should one or more processes fail and restart. Write cache is mirrored across controllers, and the system offers RAID 1+0 (mirroring + striping), RAID 5+0 (RAID 5 distributed parity + striping), and RAID MP (multiple distributed parity with striping).

In addition, controller nodes are configured in logical pairs whereby each node has a partner. The partner nodes have redundant physical connections to the subset of physical drives owned by the node pair. Within the pair, nodes mirror their write cache to each other and each serves as the backup node for the LDs owned by the partner node.

If a controller node were to fail, data availability would be unaffected. In the event of a node failure, that node's partner takes over the LDs for the failed node. It then immediately flushes data in the write cache on other nodes in the array that belongs on the LDs it has taken over.

Furthermore, under certain circumstances, the system is capable of withstanding a second Node failure (however rare) without affecting data availability. After the Node failover recovery process for the initial Node failure is complete, a second Controller Node from the remaining Node pairs can fail without causing system downtime.

Data integrity checking

In addition to hardware fault tolerance, all HP 3PAR StoreServ Storage systems offer automated end-to-end error checking during the data frames' journey through the HP 3PAR StoreServ Storage array to the disk devices to help ensure data integrity in support of Tier-1 resilience. In addition to this HP 3PAR Thin Express ASIC comes with the Persistent Checksum feature that ensures end-to-end data protection, from host HBA to physical drives.

Embedded CRC (Cyclical Redundancy Checking) checking includes, but is not exclusive to, the following layers within all HP 3PAR StoreServ Storage systems:

- CRC/parity checks on all internal CPU and serial buses
- Control cache ECC checks
- Data cache ECC checks
- PCIe I2C bus CRC/parity checks
- HP 3PAR ASIC connection CRC/parity checks
- Protocol (Fibre Channel/iSCSI/FCoE) CRC checks at the frame level (hardware accelerated via the HP 3PAR ASICs)
- Disk devices CRC checks at the block level, occurring once the data has landed and throughout the lifecycle of the data once it's stored to disk

CRC error checking is also extended to replicate data with HP 3PAR Remote Copy software, which help ensure that potential cascaded data issues do not occur. HP 3PAR StoreServ Storage replication includes a link pre-integration test to verify the stability of Remote Copy replication links in advance of adding these links within the HP 3PAR StoreServ Storage system for use with HP 3PAR Remote Copy over an IP network (RCIP).

All drives in the HP 3PAR StoreServ 20000 and StoreServ 7000 storage systems are formatted with 520-byte blocks in order to provide space to store a CRC Logical Block Guard, as defined by the T10 Data Integrity Feature (T10-DIF) for each block. This value is computed by the HP 3PAR HBA before writing each block and is checked when a block is read. SATA does not support 520-byte blocks, so on Enterprise SATA drives, data blocks are logically grouped with an extra block to store the CRC values. The CRC Logical Block Guard used by the T10-DIF is automatically calculated by the HBAs to validate data stored on drives without additional CPU overhead.

HP 3PAR StoreServ Storage continuously runs a background "pd scrubber" process to scan all blocks of the physical drives in the system. This is done to detect any potential issues at the device block layer and trigger RAID rebuilds down to 512-byte granularity if necessary. This is particularly important when it comes to flash media because it allows the system to proactively detect and correct any low-level CRC and bit errors.

Furthermore, Self-Monitoring, Analysis and Reporting Technology (SMART) predictive failures mean that any disk device crossing certain SMART thresholds would cause the storage controller nodes to mark a drive as "predictive failure," identifying it for replacement before it actually fails.

When inline deduplication is used, the controller node ASICs will perform a bit-by-bit comparison before any new write is marked as a duplicate. This helps ensure data integrity by introducing this critical check into the deduplication process to support mission-critical environments.

HP 3PAR StoreServ Storage systems also issue logical error status block (LESB) alerts if a frame arriving in the storage interface has CRC errors beyond a certain threshold. This indicates that a cable or component between the host and storage device needs replacing or cleaning.

Memory fencing

HP 3PAR StoreServ Storage is able to correct single-bit (correctable) errors and detect double-bit (uncorrectable) errors.

It achieves this by using a thread (memory patrol) that continuously scans the memory and keeps track of correctable errors at a 16 KB page granularity. If, during the scan, the thread detects uncorrectable errors, areas of the memory are fenced and put onto a "do not use" list. The system will raise a service alert when the threshold of correctable errors is reached and/or memory is fenced such that replacing the DIMM is recommended.

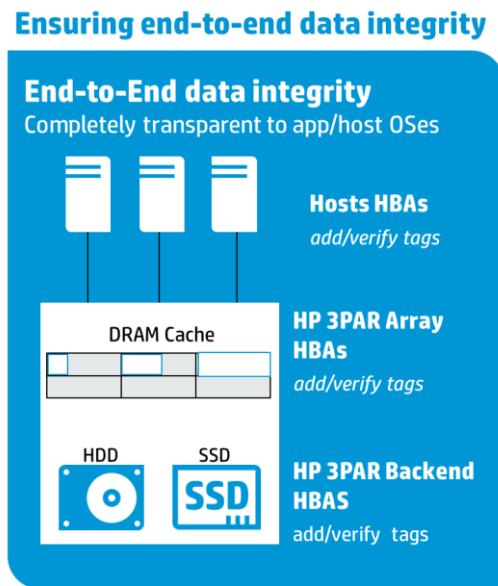
Persistent technologies

No one has time for downtime, which is why modern Tier-1 resiliency requires that data access and service levels be maintained during failure recovery, maintenance, and software upgrades. Tier-1 resiliency demands that failures not only be prevented, but that the system can recover quickly in the event that something goes wrong. Not only is HP 3PAR StoreServ Storage designed to be non-disruptively scalable and upgradable, but the system also has several advanced features to prevent unnecessary downtime and to maintain availability and performance levels during planned as well as unplanned outage events. These features are collectively known as persistent technologies.

HP 3PAR Persistent Checksum

This feature, Persistent Checksum¹ addresses media and transmission errors that can be caused by any component in the I/O stack and provides additional protection above CRC transmissions for Fibre Channel. Persistent Checksum is application independent and offers elaborate host OS support. When using supported HBAs, T10-DIF² tags are added and verified on HP 3PAR StoreServ Storage system. When using unsupported HBAs T10-DIF tags are added and verified on the array target ports and back-end HBAs, making the data secure all the way from the hosts right to the drives.

Figure 7. HP 3PAR Persistent Checksum



HP 3PAR Persistent Cache

HP 3PAR Persistent Cache is a resiliency feature built into the HP 3PAR Operating System that allows graceful handling of an unplanned controller failure or planned maintenance of a controller node. This feature eliminates the substantial performance penalties associated with traditional modular arrays and the cache “write-through” mode they have to enter under certain conditions. HP 3PAR StoreServ Storage can maintain high and predictable service levels even in the event of a cache or controller node failure by avoiding cache write-through mode via HP 3PAR Persistent Cache.

Under normal operation on an HP 3PAR StoreServ Storage system, each controller has a partner controller in which the controller pair has ownership of certain logical disks. As mentioned earlier, LDs are the second layer of abstraction in the system’s approach to virtualization of physical resources and is also where the QoS parameters are implemented (drive type, RAID, HA, etc.). Ultimately, LDs from each node pair are grouped together to form VVs. In the rare event of a controller failure or planned controller maintenance, HP 3PAR Persistent Cache preserves write caching by dynamically remirroring cache of the surviving partner controller node to the other controller nodes in the system.

For example, in a quad controller configuration (where Node 0 and Node 1 form a node pair and Node 2 and Node 3 form a second node pair), each node pair might own 100 LDs with each node within the pair fulfilling the role of the primary node for 50 of those LDs. If Node 2 fails, the system will transfer ownership of its 50 LDs to Node 3, and Node 0 and Node 1 will now be the backup (and thereby the cache mirroring partner) for the 100 LDs that Node 3 is now responsible for. The mirroring of write data coming into Node 3 for those 100 LDs will be evenly distributed across Node 0 and Node 1.

¹ HP 3PAR Persistent Checksum feature is available; supported on HP 3PAR StoreServ 20000 Storage systems.

² T10 DIF: Industry standard end-to-end data protection via T10 Data Integrity Field (DIF).

HP 3PAR Persistent Ports

Another persistent technology, HP 3PAR Persistent Ports allows for a non-disruptive environment (from the host multipathing point of view) where host-based multipathing software is not required to maintain server connectivity in the event of a node or link outage on any SAN fabric. This applies to firmware upgrades, node failures, and node ports that are taken offline either administratively, or as the result of a hardware failure in the SAN fabric that results in the storage array losing physical connectivity to the fabric.

From a host standpoint, connections to HP 3PAR StoreServ Storage systems continue uninterrupted with all I/O being routed through a different port on the HP 3PAR StoreServ Storage array. This helps you achieve an uninterrupted service level for applications running on HP 3PAR StoreServ Storage systems.

HP 3PAR Persistent Port functionality works for Fibre Channel, FCoE, and iSCSI transport layers and provides transparent and uninterrupted failover in response to the following events:

- HP 3PAR OS firmware upgrade
- HP 3PAR node maintenance or failure
- HP 3PAR array “loss sync” to the FC fabric or IP network
- Array host ports being taken offline administratively
- Port laser loss for any reason (applies to FC only)

For more information, see the [HP 3PAR Persistent Ports white paper](#).

HP 3PAR Remote Copy

HP 3PAR Remote Copy software brings a rich set of features that can be used to design disaster-tolerant solutions that cost-effectively address disaster recovery challenges. HP 3PAR Remote Copy is a uniquely easy, efficient, and flexible replication technology that allows you to protect and share data from any application.

Implemented over a native IP network (through the built-in Gigabit Ethernet interface available on all nodes) and native Fibre Channel, users may flexibly choose one of four different data replication modes—**Asynchronous Streaming** or **Asynchronous Periodic** (for asynchronous replication), **Synchronous**, or **Synchronous Long Distance**—to design a solution that meets their solution requirements for recovery point objectives (RPOs) and recovery time objectives (RTOs).

With all four of these modes, HP 3PAR Remote Copy software allows you to mirror data between HP 3PAR StoreServ Storage systems of any model or size, eliminating the incompatibilities and complexities associated with trying to mirror data between traditional vendors’ midrange and enterprise array technologies. Source and target volumes may also be flexibly and uniquely configured to meet users’ needs (e.g., different RAID levels, the use of FPVVs vs. TPVVs, or different drive types). HP 3PAR Remote Copy is “thin aware” in that it is able to replicate both thin and thick volumes by using TPVV target volumes to provide the same cost savings associated with thin-provisioned source volumes created with HP 3PAR Thin Provisioning software.

Synchronous mode solutions are suitable for data centers within metro distances and have a recovery point objective of zero seconds. When combined with HP 3PAR Peer Persistence these solutions also enables recovery time objective of zero seconds.

For asynchronous replication solutions, network bandwidth is efficiently utilized with **Asynchronous Periodic** mode. Changed data within an HP 3PAR Remote Copy Volume Group is transferred only once—no matter how many times it may have changed—between synchronization intervals. Additionally, efficiencies in the initial copy creation of the target volumes that do not require replication of “zero” data across the replication network (regardless of target volume type, thick or thin) result in a faster initial synchronization and better network utilization. With support for replication over long distances, HP 3PAR Remote Copy **Asynchronous Streaming** assures data protection against disasters by providing an RPO in seconds, whereas HP 3PAR Remote Copy Asynchronous Periodic mode ensures RPO in minutes over extended distances. Asynchronous Streaming is suitable for customers that want near zero RPO without the additional host latency associated with the synchronous mode of replication.

Synchronous Long Distance mode delivers a disaster recovery solution across long distances with a potential for zero data loss RPO and an RTO of minutes. This is achieved with a replication configuration that uses three sites (**3DC**) to simultaneously replicate a virtual volume from the primary array in synchronous mode to an HP 3PAR StoreServ Storage array located at a synchronous site (within a metropolitan area) and in asynchronous periodic mode to an HP 3PAR StoreServ Storage array located at an asynchronous site (across a long distance). In addition to the HP 3PAR Remote Copy connections from the primary array to the two backup arrays, a passive asynchronous periodic link is configured from the synchronous array to the disaster recovery array. Under the Synchronous Long Distance mode algorithm, the synchronous site intelligently tracks the delta set of I/Os that have been acknowledged to the host but which have not yet been replicated to the asynchronous site. In the event that a disaster takes the primary storage array down, the user has the flexibility to recover either from the synchronous site or the asynchronous site.

For more information, see the [HP 3PAR Remote Copy white paper](#).

HP 3PAR Virtual Domains

HP 3PAR Virtual Domains software is an extension of HP 3PAR virtualization technologies that delivers secure segregation of virtual private arrays (VPAs) for different user groups, departments, and applications while preserving the benefits delivered by the massive parallelism architected into the HP 3PAR StoreServ platform.

By providing secure administrative segregation of users and hosts within a consolidated, massively parallel HP 3PAR StoreServ Storage system, HP 3PAR Virtual Domains allows individual user groups and applications to affordably achieve greater storage service levels (performance, availability, and functionality) than previously possible.

HP 3PAR Virtual Domains is completely virtual and represents no physical reservation of resources. To use HP 3PAR Virtual Domains, a master administrator first creates a virtual domain, and then assigns logically defined entities to it. These include one or more host definitions based on Worldwide Name (WWN) groupings, one or more provisioning policies (RAID and disk type), and one or more system administrators (who are also granted role-based privileges by the master administrator).

Depending on the level of access, users can create, export, and copy standard volumes or thin-provisioned volumes. HP 3PAR Virtual Domains is ideal for enterprises or service providers looking to leverage the benefits of consolidation and deploy a purpose-built infrastructure for their private or public cloud.

Data encryption

Data is perhaps the most important asset for organizations in today's digital age. Companies are looking to protect data against theft and misuse while meeting compliance requirements. The HP 3PAR StoreServ Storage complies with the standards set forth by the National Institute of Standards and Technology (NIST) and FIPS 140-2 (Federal Information Processing Standard) and features Data at Rest (DAR) encryption that helps protect valuable data through self-encrypting drive (SED) technology. SED drives are HDDs and SSDs with a circuit (ASIC) built into the drive's controller chipset that automatically encrypts and decrypts all data being written to and read from the media.

HP 3PAR StoreServ Storage supports Full Disk Encryption (FDE) based on the Advanced Encryption Standard (AES) 256 industry standard. The encryption is part of a hash code that is stored internally on physical media. All encryption and decryption is handled at the drive level and needs no other external mechanism.

Authentication keys are set by the user and can be changed at any time. The Local Key Manager (LKM) included with the HP 3PAR StoreServ Storage encryption license is used to manage all drive encryption keys within the array and provides a simple management interface. In the event of a drive failure or the theft of a drive, a proper key sequence needs to be entered to gain access to the data stored within the drive. When an SED drive is no longer powered on, the drive goes into a locked state and requires an authentication key to unlock the drive when power is restored. In the event of a drive failure or theft, a proper key sequence needs to be entered to gain access to the data stored within the drive. Without the key, access to the data on the SED is not possible.

HP also offers the enhanced encryption support on the HP 3PAR StoreServ Storage systems by offering FIPS 140-2 compliant SED drives that provides the ability to use an external Enterprise Secure Key Manager (ESKM). ESKM is deployed whenever customers use encrypted storage or communication methods to protect their sensitive information. Herein, they store and serve keys to unlock the data stored on FIPS 140-2 compliant drives within the HP 3PAR StoreServ Storage systems with strong access controls and security.

FIPS 140-2 compliance provides the customer the satisfaction of knowing their data is securely stored on the HP 3PAR StoreServ array. Key Management on the array with either LKM or ESKM coupled with FIPS drives, offers customers a safe environment in which to securely store their data.

For more information, see the [HP 3PAR StoreServ Data-At-Rest Encryption white paper](#).

Maintaining high and predictable performance levels

The ability of HP 3PAR StoreServ Storage to maintain high and predictable performance in multi-tenant environments is made possible through architectural innovations that eliminate resource contention, support mixed workloads, and enhance caching algorithms to accelerate performance and reduce latency.

Load balancing

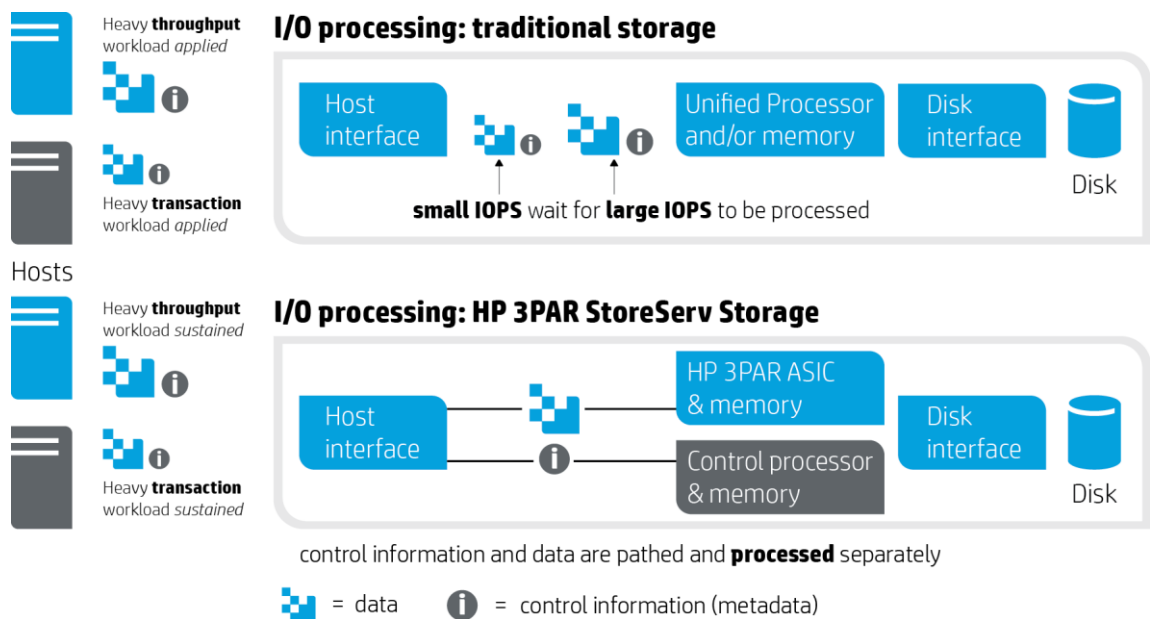
Purpose-built for virtual and cloud data centers, the HP 3PAR StoreServ Architecture is unlike legacy controller architectures in that the Mesh-Active system design allows each volume to be active on any controller in the system via a high-speed, full-mesh interconnection that joins multiple controller nodes to form a cache-coherent Active/Active cluster. As a result, the system delivers symmetrical load balancing and utilization of all controllers with seamless performance scalability by adding more controllers to the mesh.

Mixed-workload support

Unlike legacy architectures that process I/O commands and move data using the same processor complex, the HP 3PAR StoreServ Storage controller node design separates the processing of SCSI control commands from data movement. This allows transaction-intensive and throughput-intensive workloads to run on the same storage resources without contention, thereby supporting massive consolidation and multi-tenancy. This means that, for example, the system can easily handle an OLTP application and an extremely bandwidth-consuming data warehousing application concurrently with ease.

This capability is made possible by the HP 3PAR ASIC, which offloads data processing from the control processor, where metadata is processed. By pathing and processing data and metadata separately, transaction-intensive workloads are not held up behind throughput-intensive workloads. As a result, the HP 3PAR StoreServ Storage platform, as compared to the ASIC-less architectures of traditional storage vendors—including many of today’s all-flash arrays—delivers excellent performance consistently, even in mixed-workload scenarios. Figure 8 illustrates the benefits of mixed-workload support.

Figure 8. HP 3PAR StoreServ Storage with mixed-workload support



Control operations are handled as follows:

- With the HP 3PAR StoreServ 20000 Storage system, control operations are processed by up to 16 high-performance Intel Quad-Core processors (for an 8-node HP 3PAR StoreServ 20800 Storage system).
- With the HP 3PAR StoreServ 7450c and 7440c Storage systems, control operations are handled by up to four Intel 8-core processors.
- In the case of the HP 3PAR StoreServ 7400c Storage system, control operations are handled by up to four Intel Hexa-core processors.
- For the HP 3PAR StoreServ 7200c Storage system, control operations are handled by up to two Intel Hexa-core processors.

Data movement is handled as follows:

- For the HP 3PAR StoreServ 20000 Storage system, all data movement is handled by the specially designed HP 3PAR ASICs (two per controller node).
- For the HP 3PAR StoreServ 7200c, 7400c, 7440c, and 7450c Storage systems, all data movement is handled by the HP 3PAR ASICs (one per controller node).

Storage quality of service

Quality of service (QoS) is an essential component for delivering modern, highly scalable multi-tenant storage architectures. The use of QoS moves advanced storage systems away from the legacy approach of delivering I/O requests with “best effort” in mind and tackles the problem of “noisy neighbors” by delivering predictable tiered service levels and managing “burst I/O” regardless of other users in a shared system. Mature QoS solutions meet the requirements of controlling service metrics such as throughput, bandwidth, and latency without requiring the system administrator to manually balance physical resources. These capabilities eliminate the last barrier to consolidation by delivering assured QoS levels without having to physically partition resources or maintain discreet storage silos.

HP 3PAR Priority Optimization software enables service levels for applications and workloads as business requirements dictate, enabling administrators to provision storage performance in a manner similar to provisioning storage capacity. This allows the creation of differing service levels to protect mission-critical applications in enterprise environments by assigning a minimum goal for I/O per second and bandwidth, and by assigning a latency goal so that performance for a specific tenant or application is assured. It is also possible to assign maximum performance limits on workloads with lower service-level requirements to make sure that high-priority applications receive the resources they need to meet service levels.

Priority Optimization feature and industries only “Latency goal” feature that enables to set SLAs as low as 500 μ s, also makes it possible to configure service-level objectives in terms of KB/s and I/O bandwidth on a virtual volume set (VVset) or between different virtual domains.³ All host I/Os on the VVset are monitored and measured against the defined service-level objective. HP 3PAR Priority Optimization control is implemented within the HP 3PAR StoreServ Storage system and can be modified in real time. No host agents are required, and physical partitioning of resources within the storage array is not necessary.

For more information on HP 3PAR Priority Optimization, please see the [HP 3PAR Priority Optimization white paper](#).

Adaptive Flash Cache

HP 3PAR Adaptive Flash Cache is a built-in functionality of the HP 3PAR StoreServ that allows SSDs to act as Level-2 read cache holding random read data that has aged out of DRAM read cache. Adaptive flash cache automatically reduces application response time for read intensive I/O workloads and improves write throughput in mixed-workload environment. Adaptive Flash Cache effectively increases the amount of random read data cached on high-speed media on a node. By doing this, Adaptive Flash Cache can increase the overall random read IOPS an array can deliver by “unloading” the back end of the array and also result in increased write throughput. Using SSDs as Level-2 read cache to hold random read data that has aged out of DRAM cache is a cost-effective way of keeping more random read data on very fast media to improve overall system performance. An HP 3PAR StoreServ 20800 system can be configured with up to 32 TB of Flash Cache.

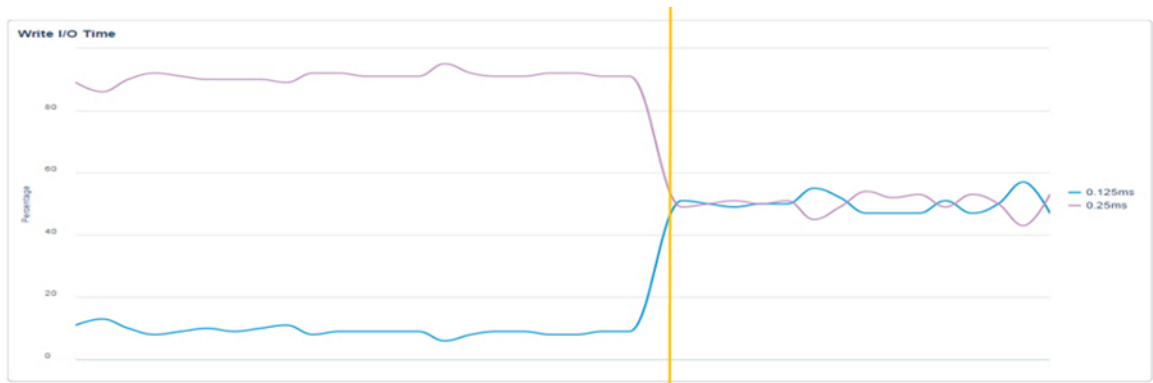
For more information on HP 3PAR Adaptive Flash Cache, please see the [HP 3PAR Adaptive Flash Cache White Paper](#).

Express Writes

HP 3PAR Express Writes represent a series of optimizations aimed to improve host write latency. This is achieved via the HP 3PAR OS by optimizing SCSI transactions and thus improving the number of interrupts per transaction, which results in improved array CPU utilization and lower host write latency. Express Writes technology enables HP 3PAR StoreServ array to achieve synchronous write latency below 200 microseconds. Depending on the workload, hosts may see an overall decrease in write latency of up to 20 percent, which may result in an increase in IOPS and throughput.

The picture on page 24 is from a system with express writes disabled and then enabled (after the line). With Express Writes disabled 90 percent of write I/O was committed in 0.250 microseconds or more and 10 percent was committed in 0.125 microseconds or less. With Express Writes Enabled 50 percent of writer I/O is committed in 0.125 microseconds or less.

³ A VVset may contain a single volume or multiple volumes. A virtual volume may also belong to multiple virtual volume sets allowing users to create hierarchical rules.



Performance benefits of system-wide striping

In a traditional storage array, small volumes either suffer from poor performance by using few drives or waste expensive resources by using more drives than required for capacity in order to obtain sufficient performance. On HP 3PAR StoreServ Storage systems, even modest-sized volumes will be widely striped using chunklets spread over multiple drives of the same type. Wide striping provides the full performance capabilities of the array to small volumes without provisioning excess capacity and without creating hotspots on a subset of physical drives. Other chunklets on the drives are available for other volumes.

Physical drives can hold a mix of RAID levels because RAID groups with HP 3PAR StoreServ Storage are constructed from chunklets rather than from whole drives. Different chunklets on a physical drive can be used for volumes with different RAID levels. On a traditional array, a storage administrator might be forced to use RAID 1 for an archival volume in order to use space that is available on a RAID 1 disk even though RAID 5 would deliver adequate performance and better space utilization. The chunklet-based approach employed by the HP 3PAR Operating System allows all RAID levels to coexist on the same physical drives, using the better RAID level for each volume. Additional details about striping are provided in the [“Highly virtualized storage operating system”](#) section.

Bandwidth and communication

The ASICs within each HP 3PAR StoreServ 20850 and 20800 Storage controller node serve as the high-performance engines that move data between three I/O buses, a four memory-bank data cache, and seven high-speed links to the other controller nodes over the full-mesh backplane. These ASICs perform RAID parity calculations on the data cache and inline zero-detection to support the system’s data compaction technologies. CRC Logical Block Guard used by T10-DIF is automatically calculated by the HBAs to validate data stored on drives with no additional CPU overhead. An HP 3PAR StoreServ 20800 Storage system with eight controller nodes has 16 ASICs totaling 224 GB/s of peak interconnect bandwidth.

The single ASIC within each HP 3PAR StoreServ 7200c, 7400c, and 7450c Storage controller node serves as the high-performance engine that moves data between two I/O buses, a dual memory-bank data cache, and three high-speed links to the other controller nodes over the full-mesh interconnect. As with the HP 3PAR StoreServ 20000 Storage series, the ASICs for the HP 3PAR StoreServ 7000 Storage series models perform parity RAID calculations on the data cache and inline zero-detection; CRC Logical Block Guard used by the T10-DIF is automatically calculated by the HBAs to validate data stored on drives with no additional CPU overhead. An HP 3PAR StoreServ 7450 Storage system with four nodes has four ASICs totaling 24 GB/s of peak interconnect bandwidth and eight I/O buses totaling 32 GB/s of peak I/O bandwidth.

Data transfer paths

Figure 9 shows an overview of data transfers in an HP 3PAR StoreServ Storage system with two simple examples: a write operation from a host system to a RAID 1 volume (lines labeled W1 through W4), and a read operation (Gray lines labeled R1 and R2). Only the data transfer operations are shown, not the control transfers.

The write operation consists of:

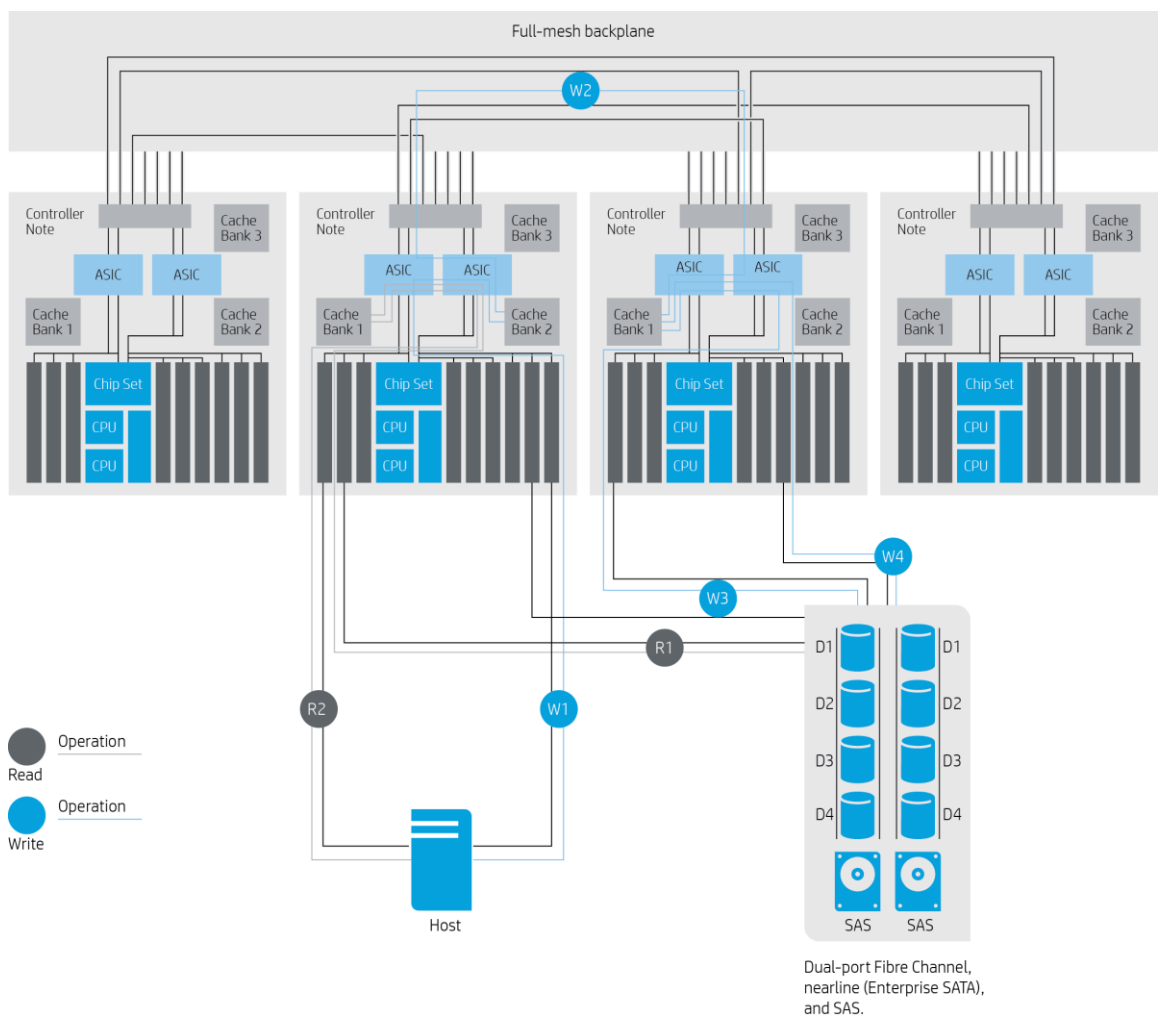
- **W1:** Host writes data to cache memory on a controller node
- **W2:** The write data is automatically mirrored to another node across the high-speed backplane link so that the write data is not lost even if the first node experiences a failure; only after this cache mirror operation is completed is the host's write operation acknowledged
- **W3 and W4:** The write data is written to two separate drives (D1 and D1'), forming the RAID 1 set

In step W2, the write data is mirrored to one of the nodes that owns the drives to which data will be written (in this example, D1 and D1'). If the host's write (W1) is to one of these nodes, then the data will be mirrored to that node's partner. HP 3PAR Persistent Cache allows a node to mirror the write data to a node that does not have direct access to drives D1 and D1' in the event of a partner node failure.

The read operation consists of:

- **R1:** Data is read from drive D3 into cache memory
- **R2:** Data is transferred from cache memory to the host

Figure 9. Data transfer paths



I/O bus bandwidth is a valuable resource in the controller nodes, and is often a significant bottleneck in traditional arrays. As the data transfer example in Figure 9 illustrates, I/O bus bandwidth is used only for data transfers between the host-to-controller node and controller node-to-drive transfers. Transfers between the controller nodes do not consume I/O bus bandwidth.

Processor memory bandwidth is another significant bottleneck in traditional architectures, and is also a valuable resource in the controller nodes. Unique to the HP 3PAR StoreServ Storage system, controller node data transfers do not consume any of this bandwidth. This frees the processors to perform their control functions far more effectively. All RAID parity calculations are performed by the ASICs directly on cache memory and do not consume processor or processor memory bandwidth.

Sharing and offloading of cached data

Because much of the underlying data associated with snapshot volumes is physically from other VVs (snap VVs and/or the base VV), data that is cached for one VV can often be used to satisfy read accesses from another VV. Not only does this save cache memory space, but it also improves performance by increasing the cache hit rate.

In the event that two or more drives that underlay a RAID set become temporarily unavailable (or three or more drives for RAID MP volumes)—for example, if all cables to those drives are accidentally disconnected—the HP 3PAR Operating System automatically moves any “pinned” writes in cache to dedicated Preserved Data LDs. This helps ensure that all host-acknowledged data in cache is preserved so that it can be properly restored once the destination drives come back online without compromising cache performance or capacity with respect to any other data by keeping cache tied up.

On flash-based systems, autonomic cache offload mitigates cache bottlenecks by automatically changing the frequency at which data is offloaded from cache to flash media. This helps ensure high performance levels consistently as workloads are scaled to hundreds of thousands of IOPS.

Pre-fetching

The HP 3PAR Operating System keeps track of read streams for VVs so that it can improve performance by “pre-fetching” data from drives ahead of sequential read patterns. In fact, each VV can detect up to five interleaved sequential read streams and generate pre-fetches for each of them. Simpler pre-fetch algorithms that keep track of only a single read stream cannot recognize the access pattern consisting of multiple interleaved sequential streams.

Pre-fetching improves sequential read performance in two ways:

- The response time seen by the host is reduced.
- The drives can be accessed using larger block sizes than the host uses, resulting in more efficient operations.

Write caching

Writes to VVs are cached in a controller node, mirrored in the cache of another controller node, and then acknowledged to the host. The host, therefore, sees an effective response time that is much shorter than would be the case if a write were actually performed to the drives before being acknowledged. This is possible because the mirroring and power failure handling help ensure the integrity of cached write data.

In addition to dramatically reducing the host write response time, write caching can often benefit back-end drive performance by:

- Merging multiple writes to the same blocks so that many drive writes are eliminated
- Merging multiple small writes into single larger drive writes so that the operation is more efficient
- Merging multiple small writes to a RAID 5 or RAID MP LD into full-stripe writes so that it is not necessary to read the old data for the stripe from the drives
- Delaying the write operation so that it can be scheduled at a more suitable time

Additional benefit to write caching is provided via Adaptive Flash Cache; while the technology is aimed to accelerate reads, the fewer random read request to the HDD tier translate to more pages available for writes.

Fast RAID 5

The architectural design of the HP 3PAR StoreServ Storage systems and HP 3PAR Operating System enables RAID 5 redundancy with performance levels that are on par with RAID 1 mirroring. This implementation combines the HP 3PAR ASIC, a large, battery-backed memory cache, and wide striping for reducing spindle contention to offer performance that approaches that of RAID 1, thus reducing the performance impact typical of RAID 5 on legacy storage architectures.

For certain workloads, Fast RAID 5 can provide higher performance than RAID 1. The write-back cache in HP 3PAR StoreServ Storage systems allows sequential writes (as generated by transaction journals, logs, and similar performance-sensitive workloads) to be collected until a full parity group can be written, reducing disk I/O traffic and possible back-end bottlenecks. Given its layout algorithm, Fast RAID 5 is appropriate for volumes that are dominated by read activity. HP 3PAR StoreServ Storage systems allow selection of the number of data blocks per parity block (N+1) to suit different needs. For RAID 5, 3+1 is the default, but any value from 2+1 to 8+1 can be selected. Higher values of N result in higher storage efficiency, but can reduce the chances for full-stripe writes. HP customers using HP 3PAR StoreServ Storage arrays typically choose HP 3PAR Fast RAID 5 for most or all volumes, as Fast RAID 5 minimizes the performance disadvantages associated with traditional RAID 1 while providing greater storage efficiency.

HP 3PAR RAID MP (Fast RAID 6)

Exponential growth in disk capacity without commensurate improvements in reliability or performance results in greater risk of data loss. For example, consider the 600 GB Fibre Channel (FC) disks and 2 TB nearline (Enterprise SATA) disks available on HP 3PAR StoreServ Storage systems. The capacity difference alone implies that reconstruction of a failed disk upon replacement can be expected to take more than six times as long with the 2 TB disk. The nearline disks are slower, too, which further increases the mean time to repair (MTTR) relative to smaller, faster Fibre Channel disks. A longer MTTR creates a larger window during which a second disk failure could cause data loss when using RAID 1 or RAID 5. Fast RAID 6 was created to address this problem. Like Fast RAID 5, Fast RAID 6 uses distributed parity, but it stores two different parity values in a manner that allows the data to be reconstructed, even in the event of two drive failures.

HP 3PAR RAID MP (multiple, distributed parity) initially supports dual parity and is capable of supporting higher parity levels in the future.

Environments such as highly consolidated virtual host environments tend to have unusually high data protection requirements due to the large number of users that could be affected by data loss, and so demand the highest level of data protection.

High I/O loads make RAID 6 problematic on traditional arrays; the implementation of RAID MP on HP 3PAR StoreServ Storage arrays is the only choice that provides the extra level of data protection without compromising I/O performance.

Data compaction

Reducing capacity requirements by mitigating overprovisioning; using enterprise-class data compaction technologies; and enabling fast, simple, automated space reclamation are essential to the industry-leading efficiency of HP 3PAR StoreServ Storage systems. Compaction technologies such as thin provisioning, thin deduplication, and thin reclamation offer efficiency benefits for primary storage that can significantly reduce both capital and operational costs with spinning media and SSDs.

Thin technologies can vary widely in how they are implemented, and this can greatly impact the ability to reduce capacity requirements and extend SSD life span without forcing performance trade-offs. Not only is HP 3PAR StoreServ Storage viewed as the industry's thin technology leader, but third-party testing and competitive analysis also confirm that HP 3PAR StoreServ Storage offers the most comprehensive and efficient thin technologies among the major enterprise storage platforms.⁴

In addition to efficiency enhancements related to performance such as system-wide striping and Fast RAID implementations discussed in previous sections, HP 3PAR StoreServ Storage offers the most comprehensive set of thin technologies available to drive up resource utilization while protecting array performance. The following sections describe various data compaction and underlying technologies. For an expanded discussion, refer to the brochure "[Thin Deduplication: HP 3PAR StoreServ Storage with Thin Technologies for data compaction](#)" and the [HP 3PAR Thin Technologies white paper](#).

⁴ "HP Thin Technologies: A Competitive Comparison," Edison Group, September 2012, hp.com/h20195/v2/GetDocument.aspx?docname=4AA4-4079ENW.

HP 3PAR ASIC for inline data reduction

At the heart of every HP 3PAR StoreServ node there is an HP 3PAR ASIC that features an efficient, silicon-based zero-detection and hashing mechanism. This unique hardware capability gives HP 3PAR StoreServ Storage the power to perform inline deduplication and remove allocated but unused space inline and non-disruptively.

The zero-detect capability can recognize an incoming write request of 16 KiB of zeros and either not allocate space for the zero block or free the space if it was already allocated for that block. All this happens in cache, and therefore no zeroes are written to the back end. When a read request comes in for a block that is unallocated, the HP 3PAR StoreServ will immediately return zeros back to the host.

Many other storage arrays do not detect blocks of zeroes on write. Instead, the zeros are written to disk and a scrubbing process later detects these zeroed blocks and discards them. With this approach, the zeroed blocks consume space until they're scrubbed, and therefore they may not be available for use by other volumes when needed. Also, there is increased load placed on the storage as the scrubbing process examines the block contents on the physical storage.

Built-in zero detection can be controlled per Thin Volume, and it is enabled by default. The HP 3PAR ASIC is also the only solution in the industry with a built-in, silicon-based hash calculation engine. With HP 3PAR Thin Deduplication software, the process of calculating the hash signatures for incoming data and verifying reads are offloaded to the HP 3PAR ASICs, freeing up processor cycles to deliver advanced data services and service I/O requests. This hardware-assisted approach, together with a unique Express Indexing feature, enables extremely efficient, extremely granular block-level inline deduplication that carries multiple benefits, including increased capacity efficiency, flash performance protection, and flash media lifespan extension. Unlike other approaches, HP 3PAR Thin Deduplication software performs a full check on all data before marking it as duplicated, which is essential to ensuring data integrity for mission-critical environments.

HP 3PAR Thin Technologies highlights

- HP 3PAR Thin Technologies are completely automated.
- HP 3PAR Operating System (OS) software uses a reservationless, dedicate-on-write approach to thin technologies and virtual copies that draws and configures capacity in fine-grained increments from a single free space reservoir without pre-dedication of any kind.
- HP 3PAR Thin Technologies use an allocation unit size of just 16 KiB, so you don't have to worry about small writes consuming megabytes or even gigabytes of capacity.
- HP 3PAR StoreServ is a storage platform built from the ground up to support thin technologies by reducing the diminished performance and functional limitations that plague bolt-on thin solutions.

HP 3PAR Thin Deduplication and Thin Clones

HP 3PAR Thin Deduplication software with Express Indexing relies on the HP 3PAR ASICs to generate and assign a hash key to each unique incoming write request. Express Indexing, a mechanism that accelerates data signature comparison, is used for ultrafast detection of duplicate write requests in order to prevent duplicate data from being written.

When a new I/O request comes in, the Express Indexing feature performs instant lookups using metadata tables in order to compare the hash keys of the incoming request to signatures of data already stored in the array. When a match is found, Express Indexing flags the duplicate request and prevents it from being written to the back end. Instead, a pointer is added to the metadata table to reference the existing data blocks. To help ensure data integrity, HP 3PAR Thin Deduplication software relies on the controller node ASICs to perform a bit-to-bit comparison before any new write update is marked as a duplicate.

With this approach, the CPU-intensive jobs of calculating signatures for incoming data and verifying reads are offloaded to the ASIC, freeing up processor cycles to deliver advanced data services and service I/O requests. Hardware-accelerated thin deduplication delivers a level of capacity efficiency that is superior to other approaches without monopolizing CPU resources and degrading performance, thereby delivering the only primary storage deduplication solution in the industry that is truly enterprise-class. ASIC-assisted, block-level deduplication takes place inline, which provides multiple benefits, including increasing capacity efficiency, protecting system performance, and extending flash media life span.

Without the purpose-built HP 3PAR ASICs and HP 3PAR Express Indexing, other storage architectures lack the processing power to simultaneously drive ultrafast inline deduplication and the high performance levels demanded by flash-based media. HP 3PAR Thin Deduplication software enables these functions to take place without contention, without sacrificing performance, and while concurrently delivering advanced data services such as replication, federated data mobility, and quality of service-level enforcements.

An extension of HP 3PAR Thin Deduplication for server virtualization environments, HP 3PAR Thin Clones software enables the creation of non-duplicative VM clones with Microsoft® Hyper-V and VMware® ESXi. These VM clones are created instantly by leveraging copy offload for VMware vStorage API for Array Integration (VAAI) and Microsoft Offloaded Data Transfer (ODX) technology without increasing capacity consumption on the HP 3PAR StoreServ Storage system. HP 3PAR Thin Clones software leverages HP 3PAR Thin Deduplication to update the metadata table without copying data, relying on inline deduplication technology to reduce capacity footprint as new write requests come in.

HP 3PAR Thin Provisioning

Since its introduction in 2002, HP 3PAR Thin Provisioning software has been widely considered the gold standard in thin provisioning. This thin provisioning solution leverages the system's deduplicate-on-write capabilities to make storage more efficient and more compact, allowing customers to purchase only the disk capacity they actually need and only as they actually need it.

HP 3PAR Thin Persistence

HP 3PAR Thin Persistence software is an optional feature that keeps TPVVs and read/write snapshots of TPVVs small by detecting pages of zeros during data transfers and not allocating space for the zeros. This feature works in real time and analyzes the data before it is written to the source TPVV or read/write snapshot of the TPVV. Freed blocks of 16 KB of contiguous space are returned to the source volume, and freed blocks of 128 MB of contiguous space are returned to the CPG for use by other volumes.

HP 3PAR Thin Conversion

With HP 3PAR Thin Conversion software, a technology refresh does not require terabyte-for-terabyte replacement, but instead offers the opportunity to eliminate a significant amount of legacy capacity through fat-to-thin conversion made possible by the HP 3PAR ASIC. In fact, the HP Get Thin Guarantee program stands behind the ability of new HP 3PAR StoreServ Storage customers to reduce storage capacity requirements by a minimum 50 percent by deploying any model HP 3PAR StoreServ Storage system and using HP 3PAR Thin Conversion and Thin Technologies to convert traditional volumes on legacy storage to TPVVs or TDVVs on the new system.⁵

HP 3PAR Thin Conversion software makes this possible by leveraging the zero-detection and inline deduplication capabilities within the HP 3PAR ASIC and a unique virtualization mapping engine for space reclamation that powers the simple and rapid conversion of inefficient "fat" volumes on legacy arrays to more efficient, higher-utilization "thin" volumes on the HP 3PAR StoreServ Storage array. Virtual volumes with large amounts of allocated but unused space are converted to TPVVs or TDVVs that are much smaller than the original volumes. During the conversion process, allocated but unused space is discarded and the result is a TPVV that uses less space than the original volume. If TDVVs volumes are chosen all duplicate data will be detected during the migration and only one copy of the data will be written to the HP 3PAR StoreServ.

HP 3PAR Thin Copy Reclamation

An industry first, HP 3PAR Thin Copy Reclamation software keeps storage as lean and efficient as possible by reclaiming the unused space resulting from deleted virtual copy snapshots and remote copy volumes. This solution builds on a virtualization mapping engine for space reclamation called HP 3PAR Thin Engine, which is included as part of the HP 3PAR Operating System.

HP 3PAR Thin Copy Reclamation software is an optional feature that reclaims space when snapshots are deleted from a system. As snapshots are deleted, the snapshot space is reclaimed from a TPVV or fully provisioned VV and returned to the CPG for reuse by other volumes. Deleted snapshot space can be reclaimed from virtual copies, physical copies, or remote copies.

⁵ Subject to qualification and compliance with the Get Thin Guarantee Program Terms and Conditions, which will be provided by your HP Sales or Channel Partner representative. More information is available at hp.com/storage/getthin.

HP 3PAR Virtual Copy

HP 3PAR Virtual Copies or snapshots are used to provide a point-in-time virtual copy of a VV and ensure that the original data can always be obtained should a problem occur when updating the data on a VV. HP 3PAR Virtual Copies rely on an efficient COW mechanisms (copy-on-write) called Delayed copy-on-write (DCOW) to improve performance from the host perspective. HP 3PAR StoreServ delayed COW, relegates the reading of the original data, updating of the base volume with the new data, and updating of the SD area with the original data to a background process, after the write update has been acknowledge to the host.

HP 3PAR Virtual Copies are always thin, non-duplicative (only one copy of the change blocks if kept and then referenced) and reservationless, they can also have variable quality of service as the user can specify a different CPG for snapshot data; for example user data (SD) on SSD and snapshots data (SD2) on a more cost efficient SAS tier. Snapshot of TDVVs (Thinly Deduped Virtual Volumes) that reside on an SSD tier will deduplicate snap data.

Thanks to HP 3PAR StoreServ efficient metadata handling, users can configure thousands of read-only and read-write snapshots and enables flexible management as promotion from any snapshot without destroying other snapshots.

Autonomic storage management

The HP 3PAR Operating System helps simplify, automate, and expedite storage management by handling provisioning, tiering, and change management autonomically and intelligently, at a subsystem level, and without administrator intervention.

The system's user interfaces have been developed to offer autonomic administration, which means that the interfaces allow an administrator to create and manage physical and logical resources without requiring any overt action. Provisioning does not require any pre-planning, yet the system constructs volumes intelligently based on available resources, unlike manual provisioning approaches that require planning and the manual addition of capacity to intermediary pools.

Ease of use

The HP 3PAR Operating System reduces training and administration efforts through the simple, point-and-click **HP 3PAR StoreServ Management Console** and the scriptable HP 3PAR Command Line Interface (CLI). Both management options provide uncommonly rich instrumentation of all physical and logical objects for one or more storage systems, thus eliminating the need for the extra tools and consulting often required for diagnosis and troubleshooting.

Open administration support is provided via SNMP, Storage Management Initiative Specification (SMI-S), and Web Service API.

Provisioning is managed intelligently and autonomically. Massively parallel and fine-grained striping of data across internal resources assures high and predictable service levels for all workload types. Service conditions remain high and predictable as system use grows or in the event of a component failure, while traditional storage planning, change management, and array-specific professional services are eliminated.

Managing performance is paramount to ensuring critical applications have the resources needed at the times of peak demands. These resources rely on a tool or tools to report metrics in a timely and accurate fashion. The monitoring of performance allows the storage administrator the flexibility to view real time data as well as performance trends by using historical data stored within a database. The tool which is used with HP 3PAR StoreServ arrays is called System Reporter.

HP 3PAR System Reporter software is a flexible, intuitive, performance- and capacity-management tool that aggregates fine-grained performance and capacity usage data for HP 3PAR StoreServ Storage, regardless of location. HP 3PAR System Reporter simplifies performance monitoring and helps create chargeback reports and resource planning.

HP 3PAR System Reporter software is installed and runs by default on HP 3PAR StoreServ systems, and automatically starts collecting performance data from the local system, there is no need to configure anything and HP 3PAR StoreServ Management Console can be used to run historical capacity and performance reports.

HP 3PAR Autonomic Groups takes autonomic storage management a step further by allowing both hosts and VVs to be combined into groups or sets that can then be managed as a single object. Adding an object to an autonomic group applies all previously performed provisioning actions to the new member. For example, when a new host is added to a group, all volumes that were previously exported to the group are autonomically exported to the new host with absolutely no administrative intervention required. Similarly, when a new volume is added to a group, this volume is also autonomically exported to all hosts the group has previously been exported to—intelligently and with no administrator action required.

In fact, management of the HP 3PAR StoreServ Storage system requires only knowledge of a few simple, basic functions:

- Create (for VVs and LDs)
- Remove (for VVs and LDs)
- Show (for resources)
- Stat (to display statistics)
- Hist (to display histograms)

Although there are a few other functions, these commands represent 90 percent of the console actions necessary, returning simplicity to the storage environment. Both the CLI and the HP 3PAR StoreServ Management Console communicate with the corresponding server process on the HP 3PAR StoreServ Storage system over TCP/IP over the on-board Gigabit Ethernet port on one of the controller nodes.

Management of the HP 3PAR StoreServ Storage system benefits from very granular instrumentation within the HP 3PAR Operating System. This instrumentation effectively tracks every I/O through the system and provides statistical information, including service time, I/O size, KB/s, and IOPS for VVs, LDs, and physical drives (PDs). Performance statistics such as CPU utilization, total accesses, and cache hit rate for reads and writes are also available on the controller nodes that make up the system cluster.

These statistics can be reported through the HP 3PAR StoreServ Management Console or through the CLI. Moreover, administrators at operation centers powered by the leading enterprise management platforms can monitor MIB-II information from the HP 3PAR StoreServ Storage system. All alerts are converted into SNMP Version 2c and Version 3 traps and sent to any configured SNMP management station.

An even more powerful and flexible way to manage HP 3PAR StoreServ Storage systems than through the CLI and the HP 3PAR StoreServ Management Console is through the use of the HP 3PAR Web Services API. This API enables programmatic management of HP 3PAR Storage systems. Using the API, the management of volumes, CPGs, and VLUNs can be automated through a series of HTTPS requests. The API consists of a server that is part of the HP 3PAR Operating System and runs on the HP 3PAR StoreServ Storage system itself and a definition of the operations, inputs, and outputs of the API. The software development kit (SDK) of the API includes a sample client that can be referenced for the development of customer-defined clients.

HP's commitment to the OpenStack community brings the power of OpenStack to the enterprise with new and enhanced offerings that enable enterprises to increase agility, speed innovation, and lower costs. Over the past two years, HP has been a top contributor to the advancement of the OpenStack project. HP's contributions have focused on continuous integration and quality assurance, which support the development of a reliable and scalable cloud platform that is equipped to handle production workloads. To support the need that many larger organizations and service providers have for enterprise-class storage, HP has developed the HP 3PAR StoreServ Block Storage Drivers, which support the OpenStack technology across both iSCSI and Fibre Channel (FC) protocols. This provides the flexibility and cost-effectiveness of a cloud-based open source platform to customers with mission-critical environments and high resiliency requirements.

For more information about HP 3PAR StoreServ support for OpenStack, please see the [OpenStack HP 3PAR StoreServ Block Storage Driver Configuration Best Practices](#).

Data optimization

HP 3PAR StoreServ storage offers several products that can be used for service-level optimization. These solutions match data to the most cost-efficient resource capable of delivering the needed service level at any given time.

HP 3PAR Dynamic Optimization is a powerful software feature that enables storage administrators to perform several online volume optimizations:

- Online data movement of existing volumes to different RAID levels and different tiers of storage. For example, application that requires high performance only during certain windows can be tuned to RAID 1 on SSDs and during lower activity periods moved to more cost-effective RAID 6 storage on nearline disks.
- Convert existing volumes to a different volume type. For example, Thinly Provisioned volumes on a HDD CPG to a Thinly Deduped volume on an SSD tier. Or the conversion of a full volume to a thinly provisioned volume. This conversion happens transparently and non-disruptively.

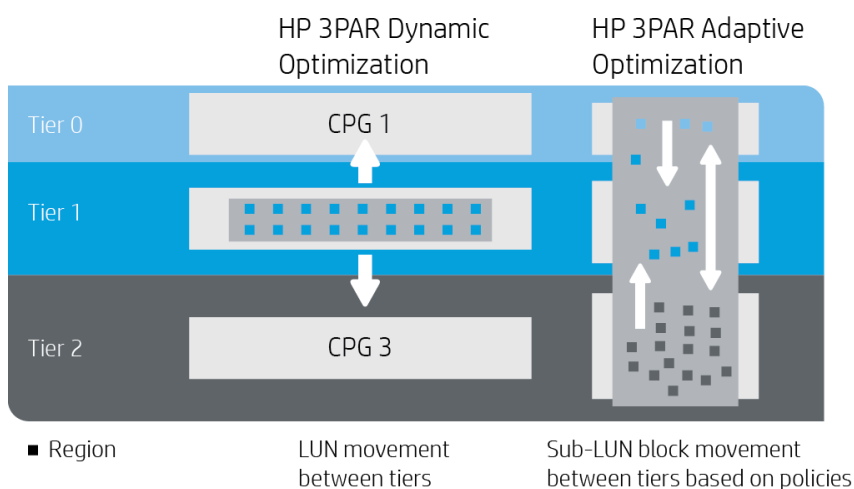
HP 3PAR Adaptive Optimization software is another autonomic storage tiering tool that takes a fine-grained, policy-driven approach to service-level optimization. HP 3PAR Adaptive Optimization works by analyzing performance (access rates) for sub-volume regions, then selects the most active regions (those with the highest I/O rates) and uses the proven sub-volume data movement engine built into the HP 3PAR Operating System to relocate those regions to the fastest storage tier available. It also moves less active regions to slower tiers to help ensure space availability for newly active regions.

Traditional storage arrays require storage administrators to choose between slow, inexpensive storage and fast, expensive storage for each volume—a process that depends on the knowledge of the application’s storage access patterns. Moreover, volumes tend to have hotspots rather than evenly distributed accesses, and these hotspots can move over time.

Using HP 3PAR Adaptive Optimization, an HP 3PAR StoreServ Storage system configured with nearline disks and Fast Class disks plus a small number of SSDs can approach flash performance at little more than the cost per megabyte of SATA-based storage, adapting autonomically as access patterns change. HP 3PAR Dynamic Optimization and HP 3PAR Adaptive Optimization are shown in Figure 10.

For more information about HP 3PAR Adaptive Optimization Software, please see the [HP 3PAR Adaptive Optimization white paper](#).

Figure 10. HP 3PAR Dynamic Optimization and HP 3PAR Adaptive Optimization



Autonomic rebalancing

The HP 3PAR Operating System automatically creates a balanced system layout by mapping VVs to many LDs, which are composed of chunklets drawn from many physical disks. As new hardware, drives, drive chassis, and nodes are added to a system, the existing data can be laid out to use the new components and benefits from system-wide striping.

The HP 3PAR Autonomic Rebalance feature provides the ability to analyze how volumes on the HP 3PAR StoreServ Storage system are using physical disk space and makes intelligent, autonomic adjustments to help ensure better volume distribution when new hardware is added to the system. This rebalancing is achieved via the “tunesys” command.

Proactive support

HP Support for HP 3PAR StoreServ Storage provides a global support infrastructure that leverages advanced system and support architectures for fast, predictive response and remediation. The **HP 3PAR Secure Service Architecture** provides secure service communication between the HP 3PAR StoreServ Storage systems at the customer site and **HP 3PAR Central**, enabling secure diagnostic data transmission and remote service connections. Key diagnostic information such as system health statistics, configuration data, performance data, and system events can be transferred frequently and maintained centrally on a historical basis. As a result, proactive fault detection and analysis are improved and manual intervention is kept to a bare minimum.

This implementation provides automated analysis and reporting that delivers accuracy and consistency, full system information in hand that reduces onsite dependencies, and fully scripted and tested automated point-and-click service actions that reduce human error.

HP 3PAR StoreServ Storage systems include a dedicated **service processor**, a server that monitors one or more HP 3PAR StoreServ systems and enables remote monitoring and remote servicing of the array. The service processor is a physical server that is external to the HP 3PAR StoreServ Storage system and communicates to it via TCP/IP. A virtual service processor is available for the HP 3PAR StoreServ 7000 Storage series. Service processor with a redundant power supply is also available for the HP 3PAR StoreServ 20000 series for higher availability.

The service processor functions as the communication interface between a customer's IP network and HP 3PAR Central by managing all service-related communications. It leverages the industry-standard Hypertext Transfer Protocol Secure (HTTPS) to secure and encrypt data for all inbound and outbound communications. The information collected via the service processor is sent to HP 3PAR Central. This information includes system status, configuration, performance metrics, environmental information, alerts, and notification debug logs. No customer data is sent.

The data sent is used by HP support teams to proactively monitor the array and contact the customer if potential issues are discovered. Customers are warned proactively about potential problems before they occur. In the case of switch issues, the customer is advised of an issue and replacement parts are dispatched. Trained HP service personnel can service the system at the customer's convenience. If the service processor cannot dial HP for any reason, both the HP 3PAR StoreServ Storage system and HP support centers will receive alerts.

The service processor is also used to download new patches, maintenance updates, and new firmware revisions; it will store them and push them to the HP 3PAR StoreServ Storage system for software upgrades. If remote access is needed for any reason, the customer can configure inbound secure access for OS upgrades, patches, and engineering access. If the customer's data center does not permit "phone home" devices, then all alerts and notifications will be sent to the customer's internal support team. The customer can then notify HP support of an issue or suspected issue, either over the phone or via the Web.

In addition, all HP 3PAR StoreServ Storage systems support a complimentary support service known as Over-Subscribed System Alerts (OSSA) in addition to and concurrent with automated remote monitoring, alerting, and notification. This automated monitoring tool performs proactive utilization checks on key system elements utilizing data that resides at HP. This data is collected periodically from the system and sent to HP. The intent is to provide valuable information such as storage node CPU utilization, disk IOPS, the number of host initiators per port, and other metrics to keep the HP 3PAR StoreServ Storage system running optimally.

For more information about HP 3PAR Secure Service Architecture, please see the [HP 3PAR Secure Service Architecture white paper](#).

HP 3PAR Storage Federation

Storage federation is a native functionality built into the HP 3PAR Operating System that enables users to move data and workloads between arrays without impacting applications, users, or services. Simply and non-disruptively shift data between any model of HP 3PAR StoreServ Storage system without additional overhead impact to hosts management layers or appliances. With a very broad support matrix of Host Operating Systems, seamless bidirectional data mobility on HP 3PAR StoreServ Storage systems can also improve availability in not just physical hosted environments but also Hypervisor based environments like VMware, Microsoft Hyper-V etc.

Storage federation is the delivery of distributed volume management across self-governing, homogenous peer storage arrays, all managed by a single management interface call HP 3PAR StoreServ Management Console. Federated data mobility allows live data to be easily and non-disruptively moved between HP 3PAR StoreServ Storage systems without any impact to host performance of any complexity of an added appliance in the data path. This is very similar to the virtual machine mobility enabled by products like VMware vMotion, Microsoft Hyper-V Migration etc. but in the case of HP 3PAR Storage Federation, data volume mobility is enabled between storage systems. Storage federation on HP 3PAR StoreServ Storage systems is peer-to-peer relationship, something that is native to the system itself.

Note

That Storage Federation is different from hierarchical virtualization, which is the delivery of consolidated or distributed volume management through appliances that hierarchically control a set of heterogeneous storage arrays. Hierarchical virtualization, also sometimes referred to as external storage virtualization, adds a new layer that has to be purchased and managed. A new layer not only introduces additional fault domains that need to be handled but also compromises the overall functionality of the entire system to be at the lowest common denominator.

In contrast, storage federation on HP 3PAR StoreServ Storage systems delivers the following benefits:

- Keeps costs low (no redundant layer of intelligent controllers)
- Reduces failure domains (no additional layers)
- Maintains functionality of each of the peers
- Simplifies administration
- Reduced complexity of Interoperability between storage systems

Storage federation has emerged as a way to address and improve storage agility and efficiency at the data center and even metropolitan area level. The primary building blocks of HP 3PAR Storage Federation are:

- HP 3PAR Peer Motion
- HP 3PAR Online Import
- HP 3PAR Peer Persistence

HP 3PAR Peer Motion

HP 3PAR Peer Motion software is the first non-disruptive, do-it-yourself data mobility tool for enterprise block storage that does not require any external appliance to be included in the data path, nor does it introduce any additional overhead on the host resources. Unlike traditional block migration approaches, HP 3PAR Peer Motion enables non-disruptive data mobility of storage volume(s) between any HP 3PAR StoreServ Storage system that are a part of the Storage Federation, without complex planning or dependency on extra tools. HP 3PAR Peer Motion offers bidirectional data mobility which helps setup your federation upfront and create a setup with back and forth data movement between two HP 3PAR StoreServ Storage system within the same data center.

Federated multi-array bidirectional data mobility between HP 3PAR StoreServ Storage systems is easy to implement and manage via the HP 3PAR StoreServ Management Console (SSMC). Simple predefined workflows have been implemented within the SSMC that give you the ability to move data—for example, moving all volumes associated with a host from one HP 3PAR StoreServ Storage system to another or moving individual volumes for workload balancing purposes—all with just one click.

The HP 3PAR Peer Motion software leverages the same built-in technology that powers the simple and rapid inline thin conversion of inefficient fat volumes on source arrays to more efficient, higher-utilization thin volumes on the destination HP 3PAR StoreServ Storage system.

HP 3PAR Peer Motion software allows all HP 3PAR StoreServ Storage systems to participate in peering relationships with each other in order to provide the following flexibility benefits:

- **Federated workload balancing**—Moves workloads from over-utilized assets to underutilized ones
- **Federated asset management**—Non-disruptively adds new storage to the infrastructure or migrates data from older systems to newer ones
- **Federated thin provisioning**—Manages storage utilization and efficiency at the data center level, not the individual system level

For more information about HP 3PAR Peer Motion, please see the [Storage Federation for non-disruptive data mobility white paper](#).

HP 3PAR Online Import

Based on HP 3PAR Peer Motion technology, HP 3PAR Online Import software leverages federated data mobility on the HP 3PAR StoreServ Storage array to simplify and expedite data migration from HP EVA Storage, EMC VMAX, EMC VNX, and EMC CLARiiON CX4 arrays, Hitachi Data Systems (HDS), Hitachi TagmaStore Network Storage Controller (NSC), Hitachi Universal Storage Platforms (USP), and Hitachi Virtual Storage Platforms (VSP) Storage systems. With HP 3PAR Online Import software, migration from these platforms can be performed in only five stages:

1. Set up the online import environment
2. Zone the host to the new system
3. Configure host multipathing
4. Unzone from the source, and start the migration
5. Validate the application

HP 3PAR Online Import will virtualize your HP EVA Storage, EMC storage or HDS Storage systems, so you can migrate from your legacy Storage platforms to HP 3PAR StoreServ Storage system, without adding any external appliance in the data path or impacting your host performance. With an easy, Do-It-Yourself migration, HP 3PAR Online Import software offers the ability to move the host data with minimal to no-disruption (depending on the Host OS) from EMC VMAX, EMC VNX and EMC CLARiiON CX4, HDS NSC, HDS USP or HDS VSP Storage arrays to any model of HP 3PAR StoreServ Storage system. The orchestration of data mobility provided for EMC and HDS Storage systems is via an easy to use command driver tool called HP 3PAR Online Import Utility.

HP 3PAR Online Import for EVA uses HP EVA Command View as the orchestration platform to enable direct migration of data from a source HP EVA Storage system to a destination HP 3PAR StoreServ Storage array without requiring host resources for data migration. The entire process can be completed with only minimal to no disruption (depending on host OS), sending EVA virtual disk and host configuration information to the HP 3PAR StoreServ Storage array without the need to change host configurations or interrupt data access in most cases.

For more information on HP 3PAR Online Import for HP EVA, see the [HP EVA P6000 to HP 3PAR StoreServ Online Import Best Practices white paper](#). For more information on HP 3PAR Online Import for EMC and HDS Storage systems, see the [HP 3PAR Online Import Software for EMC Storage Solution Brief](#), and the [HP 3PAR Online Import Data Sheet](#).

HP 3PAR Peer Persistence

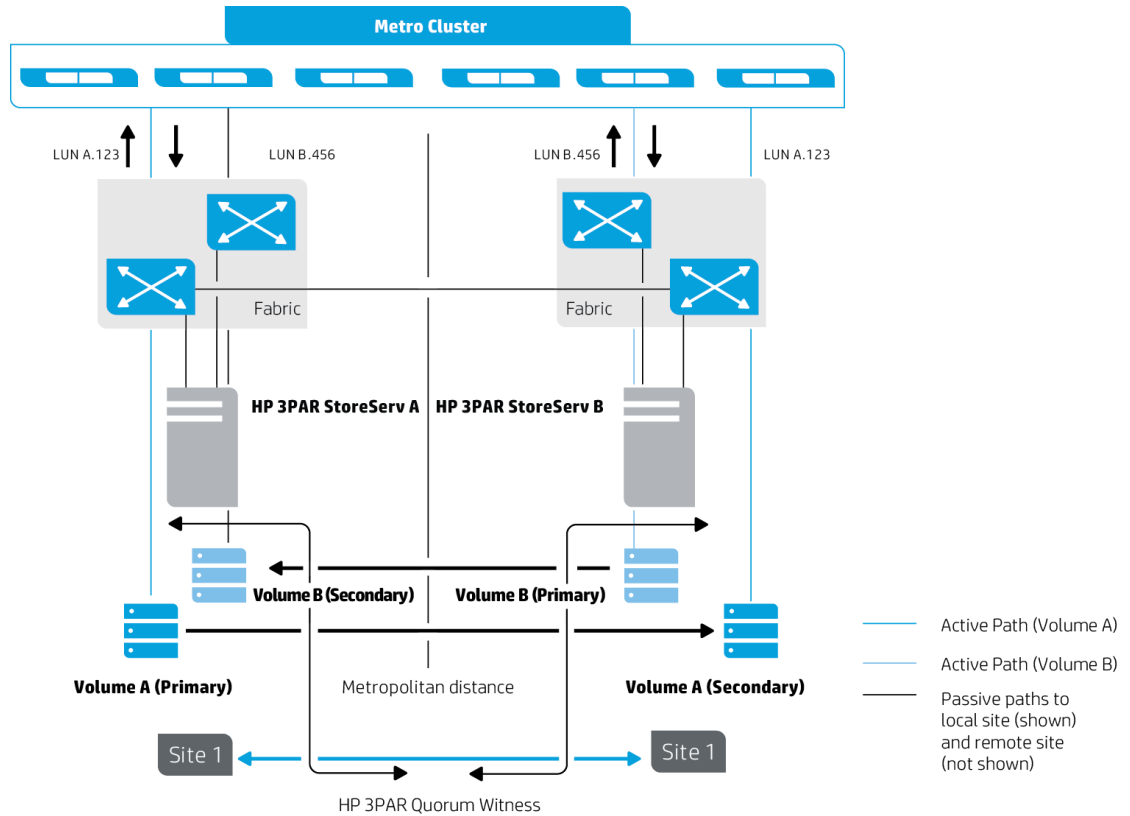
HP 3PAR Peer Persistence software enables HP 3PAR StoreServ Storage systems located within a metropolitan distance to act as peers to each other for delivering a high-availability, transparent failover solution for the connected VMware vSphere, Microsoft Hyper-V and Microsoft Windows® clusters. HP 3PAR Peer Persistence allows an array-level, high-availability solution between two sites or data centers where failover and failback remains completely transparent to the hosts and applications running on those hosts. Unlike traditional disaster recovery models where the hosts (and applications) must be restarted upon failover, HP 3PAR Peer Persistence allows hosts to remain online serving their business applications, even when the serving of the I/O workload migrates transparently from the primary array to the secondary array, resulting in zero downtime.

In an HP 3PAR Peer Persistence configuration, a host cluster would be deployed across two sites and an HP 3PAR StoreServ Storage system would be deployed at each site. All hosts in the cluster would be connected to both of the HP 3PAR StoreServ Storage systems. These HP 3PAR StoreServ systems present the same set of VVs and VLUNs with same volume WWN to the hosts in that cluster. The VVs are synchronously replicated at the block level so that each HP 3PAR StoreServ Storage system has a synchronous copy of the volume. A given volume would be primary on a given HP 3PAR StoreServ Storage system at any one time. Using Asymmetric Logical Unit Access (ALUA), HP 3PAR Peer Persistence presents the paths from the primary array (HP 3PAR StoreServ Storage system on which the VV is primary) as “active/optimized” and the paths from the secondary array as “standby” paths. Issuing a switchover command on the array results in the relationship of the arrays to swap, and this is reflected back to the host by swapping the state of the paths from active to standby and vice versa. Under this configuration, both HP 3PAR StoreServ Storage systems can be actively serving I/O under normal operation (albeit on separate volumes).

With HP 3PAR Peer Persistence, Oracle RAC stretch cluster support is now available on Linux® (RHEL 6) over FC, FCoE, and iSCSI protocols.

For more information, see the [HP 3PAR Peer Persistence white paper](#).

Figure 11. Transparent failover with HP 3PAR Peer Persistence software



Summary

Virtualization, cloud computing, and ITaaS are driving new requirements around storage agility and efficiency that are pushing legacy architectures to their brink. With a modern, scale-out architecture that is designed to meet these new demands, HP 3PAR StoreServ Storage delivers the flexibility to respond flexibly and efficiently to change, allowing you to:

- Consolidate with confidence onto a multi-tenant platform with six nines of availability
- Deliver uncompromising QoS for even the most demanding workloads
- Accelerate performance with a flash-optimized architecture featuring inline deduplication
- Cut capacity requirements by 50 percent and double virtual machine density
- Respond 8X faster with autonomic management
- Seamlessly refresh storage and maintain load balancing across arrays
- Painlessly migrate to Tier-1 storage built for cloud computing and ITaaS

HP 3PAR StoreServ Storage does this all while driving up efficiency and resource utilization with hardware acceleration that enhances performance and lowers total cost of ownership for storage. With a range of models that all leverage the same scale-out architecture and a single operating system to bring Tier-1 data services to the midrange, deliver all-flash array performance, and provide mission-critical resiliency and QoS, HP 3PAR StoreServ Storage is the only primary storage platform you will ever need for block, file, and object access.

For more information

For detailed and up-to-date specifications on each of these products, please refer to the product QuickSpecs:

- [HP 3PAR StoreServ 7000 Storage QuickSpecs](#)
- [HP 3PAR StoreServ 7450c Storage QuickSpecs](#)
- [HP 3PAR StoreServ 20000 Storage QuickSpecs](#)

Learn more at

hp.com/go/3PARStoreServ

Sign up for updates
hp.com/go/getupdated



Share with colleagues



Rate this document

© Copyright 2011, 2014–2015 Hewlett-Packard Development Company, L.P. The information contained herein is subject to change without notice. The only warranties for HP products and services are set forth in the express warranty statements accompanying such products and services. Nothing herein should be construed as constituting an additional warranty. HP shall not be liable for technical or editorial errors or omissions contained herein.

Intel and Intel Xeon are trademarks of Intel Corporation in the U.S. and other countries. Linux is the registered trademark of Linus Torvalds in the U.S. and other countries. Microsoft and Windows are trademarks of the Microsoft group of companies. Oracle is a registered trademark of Oracle and/or its affiliates. VMware is a registered trademark or trademark of VMware, Inc. in the United States and/or other jurisdictions.

