

Brochure

# Thin Deduplication

HP 3PAR StoreServ Storage with Thin Technologies for data compaction



# Use data compaction to start thin, get thin, and stay thin

## HP 3PAR StoreServ capacity efficiency benefits

### Start thin:

- Just-in-time, reservationless thin provisioning eliminates pre-allocation and pooling
- HP 3PAR Virtual Copy software enables reservationless snapshots that only consume capacity for changed data
- HP 3PAR Thin Clones software instantly creates non-duplicative VM clones for Microsoft® Hyper-V and VMware® ESXi
- System-wide striping simultaneously allows high disk utilization and performance
- RAID 50/60 uses granular, 1 GB chunklets instead of disks for increased capacity efficiency

### Get thin:

- The new HP 3PAR Gen5 Thin Express ASIC features a new deduplication engine to power inline dedupe for block and file compaction without compromising performance or scale
- HP 3PAR Thin Deduplication software with Express Indexing delivers inline, block-level deduplication without performance or capacity inefficiency tradeoffs
- Built-in, zero-detection mechanism drives efficient inline zero block deduplication at the hardware layer
- HP 3PAR Thin Conversion software leverages zero-detection to drive inline, hardware-accelerated “fat-to-thin” volume conversions

### Stay thin:

- HP 3PAR Thin Persistence software reclaims allocated but unused space associated with deleted data within volumes
- Inline deduplication increases capacity efficiency, protects flash performance, and extends flash media life span
- HP 3PAR Thin Copy Reclamation reclaims unneeded space from snapshots and remote copies

More energy efficient, more compact, and more predictable than spinning drives, the data center of the future runs on flash. All-flash arrays are capable of giving you incredible performance in a superbly efficient and increasingly cost-efficient package. However, getting the most out of your flash investments requires not just a flash-optimized hardware architecture, but a software stack optimized for flash. With a flash-optimized hardware and software architecture that includes unique efficiency mechanisms like thin provisioning, ASIC-based block-level deduplication, and space reclamation, HP 3PAR StoreServ Storage has you covered.

With the new Gen5 Thin Express ASIC, HP 3PAR StoreServ Storage continues to set the gold standard for hardware-accelerated thin technologies that drive up capacity efficiency and extend flash media life span. These technologies, combined with the industry’s largest flash drive, bring the cost of all-flash storage down to \$1.50 USD per usable gigabyte<sup>1</sup> and open up new possibilities for using flash to accelerate all of your applications.

Whether you choose a midrange flash array or an enterprise flash array, an all-flash configuration or a converged flash system that also supports spinning media, HP 3PAR StoreServ Storage offers a modern architecture that extends your flash storage investments without compromising performance, resiliency, or Tier-1 data services such as local and remote replication. A flash-first hardware architecture and a software stack optimized for flash let you start thin, get thin, and stay thin to lower up-front costs and extend media life span so you can get the most out of your flash investment.

## Why deduplication for primary storage?

Enterprises adopting flash to solve specific performance problems are seeing secondary benefits such as more predictable performance, lower operational costs, and greater productivity that are convincing them that flash is the choice moving forward. And as the mainstreaming of flash continues, deduplication for primary storage arrays is critical. Despite the fact that flash continues to become more and more affordable, the cost differential between solid state drives (SSDs) and hard disk drives (HDDs) still requires compaction technologies like thin provisioning and deduplication to elevate flash-based media cost-efficiency.

Widespread deployment of server virtualization is also driving the demand for primary storage deduplication. The potential benefits of deduplication correlate directly with data redundancy. For example, virtual machine (VM) images and client virtualization environments with hosted virtual desktops are both characterized by a high degree of data redundancy, meaning that these are two use cases where primary deduplication fits well.

The main issue that primary deduplication typically faces is that, particularly in virtualized environments, primary storage arrays are subjected to unpredictable performance demands that can require simultaneously low latency and high throughput. The impact of deduplication on performance is determined by various parameters such as whether deduplication takes place inline or as a background process and the level of granularity used for deduplication operations. This means that deduplicating data at a fine level of granularity while simultaneously delivering space savings improvements generally requires a lot of CPU processing power and memory—more than most primary storage arrays have to spare. This can force tradeoffs that restrict the overall efficiency of primary deduplication.

<sup>1</sup> Requires the use of 3.84 TB cMLC SSDs and data compaction technologies.

## The key enabling mechanisms for HP 3PAR Thin Deduplication with Express Indexing

HP 3PAR Gen5 Thin Express ASIC features a new dedupe engine that powers inline deduplication for block/file compaction without compromising performance or scale.

- Built-in data signature generation and bit-to-bit compare on match protect data integrity
- Hardware offload engines identify duplicated data without impacting performance or latency
- Fast lookup tables store location pointers to accelerate data access

Designed for mission-critical environments, HP 3PAR StoreServ Storage offers the only solution in the industry that uses built-in, silicon-based mechanisms and a patented technology to protect flash performance while delivering extremely efficient, highly granular block-level deduplication.

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. To understand how this approach is unique and why it's superior, it is important to first understand how HP 3PAR Thin Technologies drive data compaction through features that provide a broad range of capacity efficiency benefits and leverage hardware acceleration to preserve high performance and ensure the resiliency demanded of Tier-1 storage.

## HP 3PAR Thin Technologies for data compaction

Compaction technologies such as thin provisioning and thin reclamation offer efficiency benefits for primary storage that can significantly reduce both capital and operational costs with spinning media and SSDs. However, 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 tradeoffs. Not only is HP 3PAR StoreServ Storage viewed as the industry's thin technology leader, but third-party testing and competitive analysis confirm that HP 3PAR StoreServ offers the most comprehensive and efficient thin technologies among the major enterprise storage platforms.<sup>2</sup>

Unlike competitive offerings, the "thin" mechanisms unique to HP 3PAR StoreServ Storage allow the platform to offer a range of thin technologies that do not require pre-planning or up-front space reservations.<sup>3</sup> These technologies can reduce capacity requirements by 75 percent or more through data compaction.<sup>4</sup> In addition, these data compaction technologies protect SSD performance and extend flash media life span while ensuring resiliency.

HP 3PAR Thin Technologies—including HP 3PAR Thin Provisioning, Thin Conversion, Thin Deduplication, Thin Clones, Thin Persistence, and Thin Copy Reclamation—achieve data compaction through leveraging built-in hardware capabilities and Express Indexing.

- Inline, real-time HP 3PAR Thin Conversion only available with HP 3PAR StoreServ Storage provides simple data mobility across clusters or generations of storage systems without downtime, completely changing IT refresh cycles.
- HP 3PAR Thin Persistence and Thin Copy Reclamation apply these unique mechanisms to ensure that capacity remains thin by reclaiming allocated but unused space at a granular level.
- HP 3PAR Thin Deduplication and related HP 3PAR Thin Clones software take thin efficiency to the next level when used with flash arrays or any system with an SSD tier.

## Enterprise-class deduplication

HP 3PAR Thin Deduplication software is able to deduplicate data inline with a high degree of granularity to provide capacity efficiency that is superior to other approaches without monopolizing CPU resources or compromising data integrity.

HP 3PAR StoreServ Storage employs purpose-built HP 3PAR Gen5 Thin Express ASICs at the heart of each controller node. These ASICs feature efficient, silicon-based mechanisms to drive inline deduplication, including a dedicated deduplication engine. The ASICs generate and assign signatures to each unique incoming write request.

Express Indexing, a mechanism that accelerates data signature comparison, is used for ultra-fast detection of duplicate write requests in order to preventing duplicate data from being written. When a new I/O request comes in, Express Indexing performs instant lookups using metadata tables in order to compare the signatures 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 ensure data integrity, HP 3PAR Thin Deduplication software relies on the controller node ASICs to perform bit-to-bit comparison before any new write update is marked as a duplicate.

<sup>2,3</sup>HP Thin Technologies: A Competitive Comparison, Edison Group 2012. [h20195.www2.hp.com/V2/GetDocument.aspx?docname=4AA4-4079ENW&cc=us&lc=en](http://h20195.www2.hp.com/V2/GetDocument.aspx?docname=4AA4-4079ENW&cc=us&lc=en).

<sup>4</sup>As compared to not using data compaction technologies. Based on a 4:1 data compaction ratio.

# 75%

Capacity savings achieved by HP 3PAR StoreServ Storage with Thin Deduplication software.<sup>5</sup>

Customize your IT lifecycle management, from acquisition of new IT, management of existing assets, and removal of unneeded equipment.  
[hp.com/go/hpfinancialservices](http://hp.com/go/hpfinancialservices)

Developing solutions for major social and environmental challenges  
[hp.com/hpinfo/globalcitizenship](http://hp.com/hpinfo/globalcitizenship)

With this solution, the CPU-intensive jobs of calculating signatures for incoming data and verifying reads are offloaded to the ASICs, freeing up processor cycles to deliver advanced data services and service I/O requests. This hardware-assisted approach enables inline deduplication that carries multiple benefits, including increased capacity efficiency, flash performance protection, and flash media life span extension.

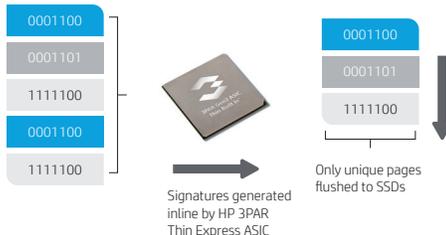
Without the purpose-built Thin Express ASICs and Express Indexing, other storage architectures lack the processing power to simultaneously drive ultra-fast inline deduplication and the high performance levels demanded by flash-based media. HP 3PAR Thin Deduplication software enables these functions to 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.

HP 3PAR StoreServ Storage with Thin Deduplication software is the only solution in the industry that uses silicon-based signature generation, allowing inline deduplication to take place while protecting flash performance levels. Best of all, it's built into the HP 3PAR Operating System, so comes at no additional cost.

## Free up flash capacity with HP Express Protect

One of the best ways to get more out of your all-flash HP 3PAR StoreServ Storage system is to offload snapshot data to cost-effective, deduplicating HP StoreOnce Systems using a feature called Express Protect, available with HP StoreOnce Recovery Manager Central (RMC) software. This feature gives your applications additional protection, accelerates recovery, and makes your flash array more efficient through cost-effective, off-array snapshot data retention.

**Figure 1.** ASIC-based signature generation for inline deduplication



Here's how it works. With Express Protect, you can create application-consistent snapshots on your array and back them up directly to an HP StoreOnce physical or virtual appliance without involving a media server or backup agent software. Only incremental changes are sent over to the HP StoreOnce appliance. These incremental changes are used to create and maintain an up-to-date, synthetic full backup—so when it's time to recover, you have a full backup ready and waiting without the need to apply incremental backups.

With no load on the array, no media server, and no dependency on backup software to move data, recovery time objectives (RTOs) are decreased to seconds or minutes. Elimination of the media server and associated backup software also means greater simplicity and significantly lower cost as there is no capacity-based licensing involved. But that's not all. You also get additional protection against file loss or data corruption beyond the oldest snapshot stored on your array. And since your snapshot data resides on the StoreOnce appliance and not your array, you aren't consuming valuable flash capacity. You also get protection against a wider range of failure scenarios such as storage platform outages<sup>6</sup>. Today this solution supports VMware and integration with other applications is under development.

Learn more at  
[hp.com/go/StoreServ](http://hp.com/go/StoreServ)

<sup>5</sup>As compared to not using data compaction technologies. Based on a 4:1 data compaction ratio.

<sup>6</sup>See ESG, Examining the Benefits and Strategic Impact of HP's Converged Storage Vision, December 2014.

Sign up for updates  
[hp.com/go/getupdated](http://hp.com/go/getupdated)

     
Share with colleagues

  
Rate this document

