

ARRAY EFFICIENT, VM-CENTRIC DATA PROTECTION

HPE Data Protector and 3PAR StoreServ

DECEMBER 2015



One of the biggest storage trends we are seeing in our current research here at Taneja Group is that of storage buyers (and operators) looking for more functionality – and at the same time increased simplicity – from their storage infrastructure. For this and many other reasons, including TCO (both CAPEX and OPEX) and improved service delivery, functional “convergence” is currently a big IT theme. In storage we see IT folks wanting to eliminate excessive layers in their complex stacks of hardware and software that were historically needed to accomplish common tasks. Perhaps the biggest, most critical, and unfortunately onerous and unnecessarily complex task that enterprise storage folks have had to face is that of data backup and recovery. As a key trusted vendor of both data protection and storage solutions, we note that HPE continues to invest in producing better backup software solutions in this space.

HPE has diligently been working towards integrating data protection functionality natively within their enterprise storage solutions starting with the highly capable tier-1 3PAR StoreServ arrays. This isn't to say that the storage array now turns into a single autonomous unit, becoming a chokepoint or critical point of failure, but rather that it becomes capable of directly providing key data services to downstream storage clients while being directed and optimized by intelligent management (which often has a system-wide or larger perspective). This approach removes excess layers of 3rd party backup software products and the inefficient indirect data flows traditionally needed to provide, assure, and then accelerate comprehensive data protection schemes. Ultimately this evolution creates a type of “software-defined data protection” in which the controlling backup and recovery software, in this case HPE's industry-leading Data Protector, directly manages application-centric array-efficient snapshots.

In this report we examine this disruptively simple approach and how HPE extends it to the virtual environment – converging server backup capabilities between Data Protector and 3PAR StoreServ to provide hardware assisted agentless backup and recovery for virtual machines. With HPE's approach, offloading VM-centric snapshots to the array while continuing to rely on the hypervisor to coordinate the physical resources of virtual machines, virtualized organizations gain on many fronts including greater backup efficiency, reduced OPEX, greater data protection coverage, immediate and fine-grained recovery, and ultimately a more resilient enterprise. We'll also look at why HPE is in a unique position to offer this kind of “converging” market leadership, with a complete end-to-end solution stack including innovative research and development, sales, support, and professional services.

CONVERGING DATA STORAGE, BACKUP AND ARCHIVE

A near universal, and seemingly simple problem for IT folks that we often hear is this – “I have very important applications and data that I need to reliably protect.” But in practice a multitude of challenges and implications have historically led to clunky, error-prone, and often incomplete backup and recovery implementations. Here are some problems we especially note in the traditional data center:

- Having to integrate multiple layers of solutions/vendors
- Struggling to manage complex, multi-step, indirect data flows
- Adopting risk due to inability to test or prove out reliable recovery
- Covering less than complete end-to-end operations
- Living with inconsistent backup services (RTO/RPO)
- Failing to meet protection service level agreements (SLAs)
- Suffering backup impacts on production processes
- Providing recovery only as a high-cost all-or-nothing restore
- Finding an effective, reliable data protection strategy for mixed virtual and physical workloads
- Relying on a simple (e.g. server or application) backup utility to fully protect a complex infrastructure

Organizations that have historically tried to tackle these problems have done so by layering up multiple vendor solutions each requiring additional integration, staffing and skills. With every additional puzzle piece the complexity and cost grows, as does the risk that it actually will recover successfully in a crisis. As IT budgets are remaining relatively flat year over year, lack of reliable data protection solutions continues to present a real vulnerability. Fortunately the storage industry has been evolving the capabilities of enterprise arrays, with increased computational power and intelligence converging into the storage directly. One of the bigger enterprise trends we see here at Taneja Group is that IT organizations are looking to their storage vendors to provide more built-in capabilities that require less overt management – enhancing features but at the same time simplifying operations and lowering costs. It's clear that the market is ripe for an aggressive large vendor like HPE to tackle this opportunity by bringing innovations from multiple disciplines together and solve the data protection problem.

We also see a kind of convergence, or evolution happening with server backup and archive. As data centers grow and consolidate, and new types of workloads like big data and mobile devices proliferate, old approaches to backup are just not keeping pace. Predictable, reliable and consistent recovery operations are becoming more difficult to achieve as a wider variety of workloads are added to the data center. In today's environments, it's no longer sufficient to offer complete system restores from off-site copies as the only method of recovering from disaster or data loss. Users expect, and increasingly utilize, fine-grained archival approaches as a means to get back on their feet if they happen to lose or delete important files. While user-implemented approaches can improve workforce productivity, when these tactics are outside of an IT organization's service offerings they expose a company to unnecessary security risk and haphazard data protection. From an administrative point of view, a good archiving solution reduces the number of secondary copies of data needed (for analysis, archive, backup, dev/test, et.al.) and helps eliminate obsolete data – reducing the online data footprint and required backup capacity. An intelligent archiving solution (like HPE Storage Optimizer, or it's big brother HPE Control Point) can use policy-based classification to identify, manage and archive data to less expensive storage and make backup and recovery process faster and more efficient.

Backup that “Snap of the Array”

Data centers with the highly performant HPE 3PAR StoreServ can take advantage of its excellent snapshot capabilities. Array snapshots allow the storage administrator to make an immediate point-in-time copy of a storage volume and preserve it for use as a form of direct recovery, kept as a part of a recovery chain, as an efficient “offline” copy to drive delta-oriented backup streaming, or layered as a source for linked images and clones.

In general, array snapshots provide inherently efficient storage-level data protection and operational flexibility. But a traditional storage-side snapshot is not aware of the running state of the application or workload at the time of the snapshot, and the actual data snapped may not be “consistent” with the

application's logical state for recovery purposes. Recovering from inconsistent snapshots may be an exercise in futility with lost data, corrupted files, and other data consistency issues.

Operating environments have introduced solutions to address the need for consistency with low-level management layers (e.g. MS VSS, VMware VADP) that temporarily quiesce workloads while storage-level data management processes like snapshots are performed. But these management layers have inherent constraints and require the server or hypervisor (in the case of virtual workloads) to mediate and coordinate the process and data flow of the backup and recovery process. With hypervisor-to-VM densities increasing to 1000's along with increasing expectations for tighter RPO's as more critical workloads are virtualized, the hypervisor can become quite busy just dealing with backup operations, negatively impacting production and limiting the practical extent of data protection services.

In the virtualized environment, an array could directly take on data protection tasks to lighten the load on the hypervisor, improving overall performance and increasing VM agility. With this approach, the hypervisor retains responsibility for coordinating virtual machine backup and restore activities, but no longer would need to manage the movement of backups to an external data protection storage target. This kind of innovation would require strong integration between the backup and recovery software, hypervisor, and the storage array snapshot capabilities in order to accomplish the whole end-to-end protection process in the fastest and most efficient manner.

BRINGING IT TOGETHER - HPE 3PAR AND DATA PROTECTOR

To meet this challenge to help better protect virtual environments, HPE has evolved traditional array snapshot capabilities to a new level of innovation. Starting with their flagship 3PAR StoreServ storage and Data Protector products, they've can now offload resource intensive backup and recovery operations from the hypervisor, freeing it for primary operations.

Formerly, and with other backup and recovery solutions, the hypervisor was instructed to take a software-based snapshot, and then requested to also stream that backup to a data protection storage target, consuming hypervisor resources in contention with the VM's themselves. With this new innovation, Data Protector essentially directs the 3PAR to take a hardware snapshot (once the software snapshot details are set by the hypervisor). This highly efficient array-side snapshot is then leveraged for any required backup streaming, offloading the hypervisor. This array side snapshot is also immediately available for highly efficient recovery purposes.

This advanced integrated backup-snapshot capability has three main benefits to the virtual environment:

1. The time to conduct a backup and the backup's impact on the hypervisor and production environment are minimized.
2. Applications and virtual machines can be "instantly" powered on and recovered from storage snapshots on the array without the need to extract a backup from a traditional external backup storage target.
3. With HPE's 3PAR snapshot capabilities, the whole snapshot management process is streamlined – integrated and automated to greatly improve reliability and recovery speed and consistency (no need to relay on error-prone manual or scripted tasks).

Details on Data Protector

Data Protector is HPE's flagship backup and recovery software solution in use at over 45,000 customers worldwide, protecting roughly half of the Global 500. Data Protector has evolved over the years into a scalable, reliable, and high performance data protection solution that can solve enterprise backup challenges for both virtual and physical workloads on heterogeneous

infrastructures. Data Protector offers many advanced features including multi-hypervisor support, fine-grained object recovery, built-in disaster recovery, snapshot integration with a variety of storage arrays, instant recovery of mission-critical applications, and integral support now for both 3PAR StoreServ and StoreOnce.

Data Protector is a fully GUI driven and automated data protection solution. It doesn't require fancy scripting or complicated gyrations to setup and ensure important applications are covered. When integrated with the 3PAR StoreServ, there is perhaps the smallest impact on the application and server platforms of any competitive solution we've seen.

Data Protector's main job of course is to manage the backup and recovery process, end-to-end. Data Protector in this case can manage multiple storage snapshots and replicas, and recover from those snapshots directly as well as from local or remote replicas as necessary. Recovery from local snapshots is near instant and especially efficient when compared with recovering from secondary storage, tape, or offsite copies (note – IT still needs to make downstream replicas for full data protection, they just won't need to access them in any recovery scenario where a relevant storage snapshot is available).

Furthermore, we found that Data Protector provides the unique ability to off-load 3PAR snapshots to an intermediary backup to disk target, HPE SmartCache. SmartCache can catalog, store and maintain snapshots online for rapid recovery operations, including single item recovery, instant VM power-on recovery, and even live VM migration. Importantly, SmartCache is an extension of the backup server allowing the administrative staff to keep instant recovery instances for longer periods of time without demanding capacity on the primary storage array. SmartCache can use any type of storage on a Linux or Windows machine.

As we've been discussing, HPE's Data Protector can now make direct use of the 3PAR StoreServ to provide for the most demanding RPO/RTO scenarios. Data Protector offers multiple options in the virtual environment – traditional agent-based backups (with an agent in the virtual machine), agentless backup (using the hypervisor to coordinate and transmit the backup to a storage target), and hardware-assisted agentless backup (where it offloads the backup process by managing the creation of snapshots on the storage array and streaming the backup set). With the hardware-assisted version, Data Protector is effectively extending its "Zero Downtime Backup and Instant Recovery" capability to offload the resource overhead that traditional backup and recovery solutions place on virtual infrastructure. The result is that virtual machine backups can now occur at high frequency with little overhead and recovered immediately to a consistent state when necessary.

In the virtual environment, VM's can be instantly restored and actually powered-on from both StoreServ snapshots and from SmartCache backup stores. Additionally, restored VM's can be migrated live back to a production VMware environment while running. This is useful for verifying backup/restore functionality as well as improving the RTO of any given VM.

Data Protector can also be configured to work with 3PAR's Remote Copy facility. If a primary StoreServ is configured to mirror itself via remote replication to a secondary array, Data Protector can offload all backup operations to that secondary. This scheme reduces any backup overhead on the primary, and provides an additional recovery alternative.

We should also mention that HPE Data Protector is based on HPE's Adaptive Backup and Recovery (ABR) vision to use integrated, continuous analytics to better optimize and automate backup and recovery processes. An ABR approach bakes enterprise class management right into the data protection process to handle what otherwise can quickly become difficult and complex to operate at scale. ABR solutions would include providing actionable recommendations to optimize backups and ensure recoverability, real-time views into the operational health of the data protection infrastructure, and the

ability run simulations to determine the impact a new workload might have on the existing environment.

An adaptive backup and recovery (ABR) approach eventually drives fully automated operations based on “self-learning” analytics. With every release of Data Protector and ecosystem solutions (i.e. 3PAR StoreServ), HPE is maturing their ABR approach in practice. IT leaders are increasingly investing in operational analytics because real value can be realized through making data-driven optimizations (instead of always throwing additional hardware at virtual infrastructure challenges).

Data Protector readily works with HPE StoreOnce, HPE’s world-class backup target. StoreOnce gives Data Protector a fully federated backup target architecture to work with. StoreOnce’s fully deduped backup storage can be deployed in many different, and all entirely interoperable ways, including as a virtual appliance. StoreOnce implements advanced schemes like Adaptive Bidding/Routing to efficiently dedupe on a global basis across multiple backup target pools. Together, Data Protector and StoreOnce form a highly optimized and complete global data protection solution for StoreServ arrays.

HPE 3PAR StoreServ

HPE 3PAR StoreServ is HPE’s highly capable and competitive enterprise class primary storage array, with many advanced feature sets. Fundamentally, the innovative 3PAR architecture provides a scalable, agile, and highly performant base for advanced enterprise storage capabilities.

The StoreServ can easily use its differential snapshot capability to take periodic, say hourly, snapshots. These are thin by nature, and so are highly efficient. At any point a full restore image can be had by synthetically combining the required differential snapshots as needed. This is much more efficient and reliable than a scheme in which recovery attempts have to carefully restore a complete full snapshot plus a series of incremental snapshots. It’s also possible here to just restore the delta differential snapshots as desired.

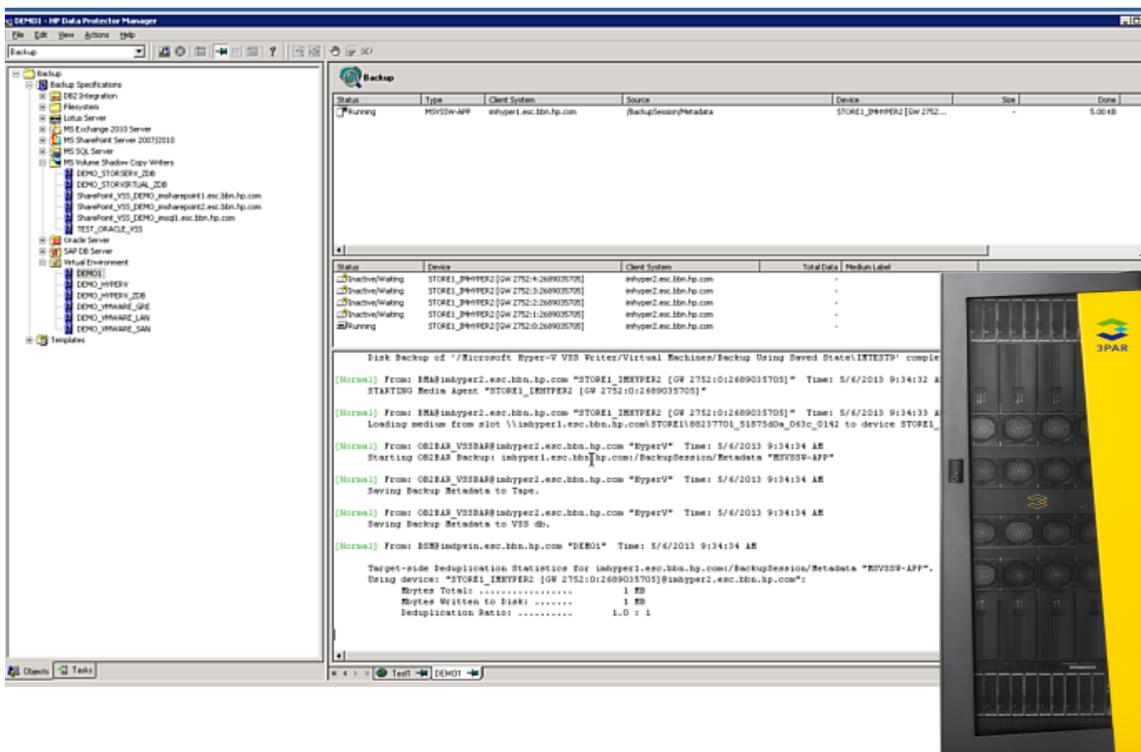


Figure 1 - HPE 3PAR StoreServ Snapshot

It's worth pointing out that by doing differential snapshots the whole backend process is made vastly more efficient as well. Full snapshots require a lot of bandwidth and processing compared to just processing (replicating, deduping) differential snapshots.

Recently, HPE 3PAR has made big announcements centered on flash storage technologies, including an all-flash version (e.g. StoreServ 7450) of their storage array. While many traditional storage arrays are beginning to offer all-flash versions, they are mostly constrained by non-flash internal designs. HPE designed 3PAR from the start to leverage multiple tiers of storage, and with massive inline processing power was already fundamentally flash-ready, so it's truly a platform for the future.

Other notable features include an Adaptive Flash Cache and Express Writes to accelerate performance, Thin Deduplication on all SSD tiers in all StoreServ models, and a Flash Advisor to figure out how to maximize flash investment.

Benefits of HPE Data Protection Solution

It's worth noting that customers that have already 3PAR StoreServ can further reduce their TCO significantly by leveraging Data Protector's strong integration with the 3PAR.

We see HPE as a big innovation, full stack vendor that can provide lots of benefit to its clients. Other factors in selecting HPE should include considering its:

- Global presence
- Vast partner network
- Deep relationships with other big vendors (e.g. Microsoft, SAP)
- Trusted sales force that can help with complex situations
- Established long-term future
- History of technical innovation

HPE has long held a stellar reputation for technical leadership. In every HPE business unit we find best-of-breed products backed by engineering excellence. In hardware units like storage, products like 3PAR StoreServ can face down competition on both performance and cost, while in software HPE continues to re-invigorate product lines like Data Protector with world-class capabilities today while drawing out exciting and believable futures going forward.

Clearly, a critical improvement that HPE has been working on in the last couple of years is bringing separate products to market together as complete solutions that provide customers with even better returns – solutions that directly address business challenges and that deliver great business value. HPE under Meg Whitman has committed to putting their customers first, and has worked hard to tether their various technological domains to service what customers really need.

TANEJA GROUP OPINION

This idea of using Data Protector to drive 3PAR intelligently in the virtualized environment not only beats out most other data protection approaches, it significantly shows where the reinvigorated HPE is capable of going. The broader application of using one world-class solution as the larger intelligence to drive optimized functionality out of another world-class solution may actually be unique to HPE as the only complete full-stack technology solution provider.

We've long thought that HPE's Data Protector and 3PAR StoreServ (as well as HPE StoreOnce with its federated dedupe) individually and now particularly as a jointly valuable solution can disrupt the data protection competition. Both products already have large install bases, and those customers will find great value by leveraging the cross-innovation HPE is bringing here. In addition to the large base of existing product deployments ripe for cross-implementation, we feel that HPE will be able to garner a large increase in market share also with new solution-level customers and expansion opportunities into virtualized environments at many levels.

HPE Data Protector and 3PAR StoreServ individually offer a lot of value, but even more so together. HPE seems to be doing all the right things to bring these products to market closer together with a more seamless customer experience. We expect HPE to bring this same solution focus across other components of their complete IT portfolio to offer more business and IT service focused solutions, and become even more of a trusted business level “partner” to their customers.

Still, while HPE is helping simplify and consolidate the IT data center for their customers and working on becoming a whole stack solutions provider, they aren’t aiming to lock-in clients with HPE only solutions. Rather, it seems more of a convenience and TCO plan of attack – HPE can “converge” valuable solutions for clients saving up front time, money, and risk to get things going and yet still support standards, interoperability, heterogeneity in the data center without imposing vendor lock-in.

HPE, already a long recognized leader in enterprise solutions and technology, is ready to roll-out new game-changing technologies and also new ways to leverage those technologies as solutions with real business value across the broader IT market. Given some of the recently observed activity with other large IT vendors, HPE has a real opportunity to become THE trusted, complete, one-stop solution provider.

NOTICE: The information and product recommendations made by the TANEJA GROUP are based upon public information and sources and may also include personal opinions both of the TANEJA GROUP and others, all of which we believe to be accurate and reliable. However, as market conditions change and not within our control, the information and recommendations are made without warranty of any kind. All product names used and mentioned herein are the trademarks of their respective owners. The TANEJA GROUP, Inc. assumes no responsibility or liability for any damages whatsoever (including incidental, consequential or otherwise), caused by your use of, or reliance upon, the information and recommendations presented herein, nor for any inadvertent errors that may appear in this document.

HPE Document #4AA6-3387ENW