

Storwize: The Real World Economic Benefits of Primary Optimized Storage

November 2007



With data growth exceeding 50% for most organizations, storage budgets are being devoured by the insatiable need to provision more and more physical storage capacity. Further exacerbating this trend is the fact that organizations' data protection, compliance, test and development, and disaster recovery requirements are expanding too, requiring more data to be stored online longer than ever before. The net result is that not only is primary storage capacity expanding rapidly, but also most organizations have between 4 to 8 copies of mission critical systems data. Something has to give. IT cannot endlessly spend money on purchasing physical storage capacity, let alone be able to afford the OPEX costs to manage all that physical capacity.

As a result, we at Taneja Group believe that Primary Optimized Storage (POS), a new class of compression and data reduction technologies, will become the de facto capability for addressing this dire crisis. Technologies like data de-duplication and single instancing are already well on their way to becoming non-optional capabilities for high growth data protection and archival environments. We believe that it is only a matter of time before IT overcomes its risk aversion and aggressively applies compression and data footprint reduction technologies to the primary storage side too.

As part of this profile, we have had the opportunity to interview two early adopters of Primary Optimized Storage, and understand what kind of Return On Investment (ROI) potential the technology holds. In the following report, we will define this emerging category of Primary Optimized Storage (POS), and introduce Storwize, an early pioneer of these data reduction capabilities. Furthermore, we will discuss how Primary Optimized Storage technology, like Storwize, directly links back to hard dollar ROI for end users, and then summarize our findings from our interviews about the economic benefits that these two end users have enjoyed.

Introducing Primary Optimized Storage

Taneja Group has followed the emergence of Capacity Optimized Storage (COS) technologies since its inception. From the beginning COS technologies like data de-duplication have been focused and engineered for reducing the data footprint of

backup and archival information. Technologies like data de-duplication perform well on sequential streams of information where they can inspect and apply hash based algorithms to reduce copies of reoccurring data patterns

However, over the past 18 months, we have begun to witness a new class of technologies

S O L U T I O N P R O F I L E

and solutions come to market that seek to reduce and optimize the data footprint of primary storage. This class of technologies excels at handling and supporting applications with small block, random I/O profiles. As a result, these primary storage optimization technologies can be used on applications such as databases, email, and file and print serving, along with High Performance Computing applications and their data sets. We have dubbed this class of technologies, Primary Optimized Storage (POS). Prior to POS technologies, there has been no viable approach or algorithms that could do real-time data compression without impacting primary storage performance. POS represents a new locus of innovation in the storage industry.

Meet Storwize

Founded in 2004, Storwize has emerged as the early leader in the POS market. The company has developed substantial intellectual property in terms of creating lossless compression with no performance impact to mission-critical applications. Storwize built their pioneering technology of real-time data compression by coupling the industry-proven Lempel-Ziv compression algorithm with innovative data ordering capabilities.

Unlike backup and archival applications, Storwize's approach works well on small block random I/O workloads that are the province of mission-critical applications. As a result, Storwize has showcased dramatic compression ratios on mission-critical applications workloads such as OLTP databases and file and print serving.

The company's main product line is its STN-6000 series of appliances. The STN appliance is an in band device that sits between hosts servers and storage systems. The appliance is completely stateless, and has no onboard storage. At present, STN series appliances support both CIFS and NFS data access protocols, but Storwize has plans to support both iSCSI and Fibre Channel block protocols in the future.

The product line is composed of two models: STN-6300 for departmental and remote offices and STN-6500 for enterprise and technical applications. The STN-6300 comes standard with 4 main cores and 4GB of memory for compression and data reduction processing and between 4 to 8 ports of connectivity. The STN-6500 is a more powerful model that includes 4 main cores with the potential to add up to 12 additional auxiliary cores and 6 GB of memory and between 8 to 12 ports of connectivity. All the STN appliances can be clustered for fault tolerant access to data.

Storwize STN appliances are unique among other Primary Optimization Storage products on the market today in the following dimensions:

1. **Non-disruptive to Existing Infrastructure** – Storwize requires no forklift upgrade or reconfiguration of existing storage or storage networking to deliver its compression benefits.
2. **No Impact on Performance** – Unlike other compression and data reduction techniques, Storwize's technology does

S O L U T I O N P R O F I L E

not measurably impact or slow down the data path.

3. **Lossless Compression** – Unlike other compression schemes, Storwize’s data reduction techniques preserve complete data fidelity and ensure no data loss as part of the compress and uncompress process.

The Business Case for Primary Optimized Storage

Primary Optimized Storage technology offers one of the most compelling and immediate Return On Investment (ROI) available in infrastructure IT today. Primary Optimized Storage reduces costs and improves efficiency in five key dimensions:

1. **CAPEX Savings on Primary Storage Capacity Purchases.** Primary Optimized Storage obviates the need to purchase more primary physical storage capacity in the near term and greatly reduces the amount of money an organization must spend on storage capacity to keep up with the inevitable data growth in most data centers. POS technology offers the ability to compress or reduce the data footprint by as much as 90%. In fact, a typical real world environment will experience data reductions on the order of 40% to 50%. By cutting the data footprint in half, organizations also cut their storage capacity budgets in half as well. As a result, organizations deploying primary storage optimization solutions realize an immediate ROI and payback on their investment because they can delay or

minimize their future purchasing on physical storage capacity. This represents a direct hard dollar savings for IT.

2. **CAPEX Saving on Secondary and Tertiary Storage.** By compressing and reducing the data footprint at its source, data that is copied, replicated, or backed up remains compressed and the initial savings and compression effects are magnified throughout the storage environment. From our research, it is not uncommon that between 4 to 8 copies of production data is made through disaster recovery, data protection, and test and development processes. For example, an organization with 10TBs of original primary data where six separate copies of the data are created for test/dev, data protection and DR purposes can free up 35 TBs of storage capacity assuming that they can achieve a 50% reduction on the original 10TBs. By reducing data at its source, the downstream effects of Primary Optimized Storage technology can significantly reduce the data footprint across the organization and result in tremendous hard dollar savings.
3. **Power, Cooling, and Space Savings.** A spillover effect of reducing capacity is that Primary Optimized Storage also reduces the amount of power, cooling, and floor space needed in the data center. In short, each spinning disk that is eliminated results in less power consumption to spin it, less power needed to cool it, and less rack space needed to house it. Given that power, cooling, and space represent one of the largest line item expenses in data center operations,

S O L U T I O N P R O F I L E

reducing these costs can be a critical ROI driver for Primary Optimized Storage technology. Furthermore, Taneja Group believes that as storage capacities continue to scale into the petabyte range, capacity optimization technologies will become non-optional in terms of intelligently managing data growth and curbing power and space requirements. In short, Primary Storage Optimization technology becomes a key fulcrum for achieving green storage initiatives.

- OPEX Savings from Less Storage.** A second spillover effect of reducing deployed capacity is the reduction of administrative cost. There is a direct relationship between the number of TBs deployed and the number Full Time Equivalents (FTEs) needed to manage that capacity. Most IT departments staff their storage teams based on a fixed staff ratio (e.g. Managed Capacity in TBs to FTEs). By reducing total amount of deployed capacity and mitigating future capacity additions, IT no longer needs to increase headcount aggressively. Moreover, critical personnel are now available for more strategic IT projects and tasks. As a result, Primary Optimized Storage technology allows companies to contain their OPEX spending and achieve better productivity from their current staffing levels.
- Bandwidth Efficient** – As a result of reducing the primary data footprint, organizations can replicate the same amount of data using less bandwidth and do it faster than they could before. Therefore, POS technologies improve DR

RPOs and allow end users to dramatically reduce their telecom bill.

- Reduced Backup Window** – As backup windows continue to shrink and data continues to grow, users are under intense pressure to complete their backups in the allotted window. A secondary effect of POS is that it reduces the amount of actual data that has been backed up. As a result, POS technology allows users to speed backups, meet their backup SLAs, and minimize the amount of system downtime required for data protection.

Two Real World Storwize Deployments

We will now turn our attention to briefly describe two real world deployments of Storwize's Primary Optimized Storage solution and the economic benefits that they have observed. Taneja Group had the opportunity to speak at length to end users at each organization and interview them about their usage model and the economic benefits that they have observed over the course of their deployment.

International Bank

We had the opportunity to interview an IT manager at a large international bank who was an early adopter of Storwize and had had the product deployed in their storage environment for over a year. The bank has offices and subsidiaries throughout North America, Latin America, and Europe. It also operates investment banking services by underwriting and investing in companies and

S O L U T I O N P R O F I L E

provides trust services to individuals and businesses.

The Storwize deployment front ends a Network Appliance FAS 940 filer with 14TB of raw capacity. The NetApp filer support a mix of applications and workloads from standard file serving and user home directories to the storage of Access databases and other business critical Line of Business (LOB) application data. In sum, the Storwize and NetApp filer support over 1,000 concurrent users.

As a major financial services institution, the bank is naturally very cautious in adopting new technologies. Despite the perceived risks of deploying a technology in the data path, the bank's IT staff chose to deploy the Storwize product due its potential for exceptionally high ROI and near immediate payback on their initial investment. The bank is always looking to cut costs and drive greater operating efficiencies and Storwize offered a unique solution unavailable elsewhere in the industry.

After a year of deployment, the IT staff at the bank reports that Storwize has allowed them to delay additional storage capacity purchases, freeing up significant CAPEX dollars for other IT initiatives. In fact, even though the bank's unstructured data is growing at over 100% a year, they have not had to purchase new storage capacity since deploying Storwize. Furthermore, their initial investment was paid back within 6 months of deployment. Given the tremendous ROI and strong payback of the Storwize solution, the bank has chosen to expand its deployment of Storwize and has placed another order for

four more STN appliances for another portion of their file storage environment.

IT management at the bank has observed that data reduction rates vary depending on the application and file types. They estimate that on their workload and data, Storwize has reduced their overall data footprint between 30 to 35%. In short, the bank's IT staff estimate that they have achieved the following economic savings as a result of deploying Storwize:

- Saved tens of thousands of dollars by not having to buy additional NetApp storage capacity.
- Saved between \$50,000 to \$60,000 on power, cooling, and floor space.
- Saved additional CAPEX by not having to purchase a VTL solution. Storwize allowed them to shorten their backup window to an acceptable operational level, obviating the purchase of a VTL.

Subsidiary of Fortune 500 Energy Company

Seismic interpretation and analysis are heavily computational and I/O intensive tasks that require massive amount of processing and raw storage capacity. As a result, energy companies are drawn to any solution that offers the ability to control storage costs and reduce power, cooling, and space requirements of that storage. Primary Storage Optimization is such a solution.

Taneja Group had the opportunity to interview an IT manager at the software and services arm of a major Fortune 500 energy

S O L U T I O N P R O F I L E

company. The subsidiary specializes in developing solutions that aid in the interpretation, visualization, and analysis of geological and geophysical data. Their solutions are in use throughout the oil and gas industry and help energy companies validate prospects, define reservoirs, and manage risk throughout the oil and gas exploration process.

Over the past ten years, this energy subsidiary has evaluated various compression technologies in the hope of finding a solution that it can pair with its software applications. However, it never was able to find a completely lossless compression or data reduction technology with no performance impact. Storwize is the first and only solutions that they have tested extensively and deployed at a number of their accounts that meets these criteria.

Through its extensive testing, IT management at the subsidiary concluded that Storwize allows their clients to save 40% on their storage capacity costs. In fact, the subsidiary has conducted a variety of benchmarks on different types of data and applications related to seismic interpretation and analysis. For example, the subsidiary tested Storwize on oil and gas specific file types, along with application text files and Oracle data files. Across these different file types and application usage models, the IT team observed that Storwize delivered an average compression of between 40% and 50%. The lowest observed reduction was 20%, while the highest compression rate was on the order of 70% to 80%. From the subsidiary's point of view, Storwize represented a slam dunk ROI and immediate

payback for their clients, given the petabytes of seismic data deployed by many of their Oil & Gas clientele.

According to the IT team, another highly prized benefit of the Storwize solution was the fact that it is a completely non-disruptive addition to an existing NAS infrastructure. This is a critical point for the subsidiary's customer base. Many of them are unwilling to consider data reduction technologies that require a rip and replace of their existing Network Appliance storage systems and management tools.

The icing on the cake for the subsidiary is that they were actually able to benchmark and demonstrate that using Storwize made their software applications perform faster. They observed that Storwize accelerate the read and writes of large sequential portions of data. For example, if their application needed to write 10GBs of data, the write would be compressed to 6GB and would complete approximately 10% faster than a 10GB write without Storwize. Since their applications are constantly reading and writing large seismic data sets, the performance speedup from Storwize was significant and allowed their clients to complete business critical analysis faster, compressing the time to find oil and gas deposits and improving their profitability.

Lastly, IT management was very impressed with Storwize's ability to enable cost-effective, rapid DR for their clients. Because of the large file sizes in seismic datasets, replicating data from site to site can be an expensive and slow process. By leveraging Storwize's data reduction technology, the

S O L U T I O N P R O F I L E

subsidiary could compress their files by 40% before replicating them to a secondary DR facility. The subsidiary observed that they were able to replicate the data faster and use less bandwidth, resulting in better RPOs and lower bandwidth costs.

Taneja Group Opinion

It is rare that we come across a technology in infrastructure IT that has such a compelling and clear ROI. Moreover, this is not a ROI based on soft costs and unproven assumptions. Primary Optimized Storage delivers hard dollar cost savings through the immediate reduction of the primary, secondary, and tertiary data footprints. Moreover, these obvious savings result in trickle down savings on power, cooling, space and improved data protection and disaster recovery SLAs. Truly, POS technology offers a compelling 360 value proposition unrivaled in the storage industry today.

It is one thing for a vendor to assert and quantify the economic benefits of POS, but it

is quite another thing when end users testify to the efficacy and overwhelming savings that this technology can deliver. Real world proof points are the ultimate litmus test of a technology value. Clearly, Storwize end users have seen dramatic economic benefit from their solution.

With data de-duplication and single instance storage becoming accepted and adopted, we believe that data footprint reduction technologies for primary storage are close behind. In short, we believe that Primary Optimized Storage will become a non-optional capability for users' storage infrastructure going forward. Frankly, the economics are too compelling for it not to become ubiquitous. Storwize has emerged as a pioneer and early leader of this rapidly evolving segment. It is our opinion that any end user who is grappling with significant data growth and must contain costs around their storage infrastructure would be wise to consider Storwize. We don't think that they will be disappointed.

***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 which may appear in this document.*