

STORAGE SWITZERLAND REPORT

ARCHITECTING STORAGE NETWORKS FOR DATA DELIVERY VS. DATA SERVICES



George Crump, Senior Analyst

Over the past few years storage systems have become increasingly feature-laden. Capabilities like snapshots, replication and thin provisioning are now commonplace. These features represent services that the storage system can provide to make the protection and provisioning of data easier. What has been lost in the proliferation of data services functionality, however, is getting data from server to storage and back again as fast as possible – data delivery.

To start with, it's important to remember that every storage system has a storage processor. Similar to the way a processor on a server is the engine that drives its applications, the processor in a storage system provides the basic storage capabilities for that system including: RAID functions, disk I/O and LUN management. Storage processors also enable the growing number of advanced data service capabilities discussed earlier. Storage systems can support these new features because storage processors have historically had spare cycles, since the drives that sit behind the controllers haven't typically generated enough I/O to keep the CPUs busy. As a result, advanced data service capabilities have continuously been added to storage systems, and each new feature has come at the price of a decrease in available storage processor resources.

Prior to SSD, the way the typical storage vendors improved performance was with more disk drives. Their strategies included increasing array drive counts, short-stroking

drives (writing only the faster, outer cylinders) or just specifying faster disk drives. But using these methods to address data delivery is expensive and inefficient. Even for applications that can justify the investment, these drive-based performance solutions become far too complicated. Additionally, there are likely other applications that could benefit from a performance boost but can't justify that level of investment. Many IT organizations are looking to solid state storage devices to provide a more efficient and less complex solution.

Traditional storage vendors know there is a growing performance challenge and have also begun to turn to SSD (solid state disk) to help address the problem, but have found it challenging to help customers determine which applications to move to SSD and how to migrate the active components of those applications. There is also sometimes lack of expertise on the storage manager's part to identify which applications should be moved to faster or slower tiers of storage. This lack of expertise is compounded by the lack of tools to help customers move data between SSD and other types of storage. Finally, there is a lack of time to properly manage this new tier and make sure that only the most active data is on it. SSD becomes one more thing in the storage manager's day to manage and tweak, costing time that most storage managers simply don't have. This has led to the recent surge in automated tiering technology that automatically places active data on a faster tier of storage, typically SSD, by some storage vendors.

The challenge, however, is that many storage systems have become so bogged down by features that they do not have the ability to support fully the incredibly high I/O of memory-based storage technology and the storage system's software actually becomes part of the problem. Also, the storage system's architecture, which was designed for legacy mechanical drives, was never designed to handle the I/O capabilities of SSD.

In both cases, this usually means a move to a newer storage system or in some cases can require a total replacement of storage vendors. If the data center is ready for a storage refresh then that may be acceptable, but in tight economic times customers are more often looking for a way to optimize performance without making these kinds of investments. What's needed is the ability to separate the data delivery function from the data services function. This would allow customers to improve their I/O response times without having to replace and learn a new set of data services processes.

Companies like [Avere Systems](#) are abstracting data delivery from data services with purpose-built appliances and software focused on performance. These systems, as we describe in our recent article "[Storage Performance Sprawl](#)", provide an alternative to automated tiering from traditional suppliers and allows for a broader treatment of that technology. Also, by focusing on the data delivery aspect exclusively, this technology gains several advantages over the traditional storage manufacturers.

First and most importantly, a separate automated tiering system and the supporting software are purpose-built to leverage DRAM and Flash based storage. There is significantly less system overhead in these types of appliances compared with traditional storage and the separation of data delivery from data services greatly improves performance by removing bottlenecks. With a purpose-built appliance, not hindered by a dizzying array of data services, the premium storage is not hindered by bottlenecks elsewhere in the system, as can be the case with traditional storage array processors that must support those extensive data services features.

Second a separate automated tiering system can be a significantly more cost effective way to introduce high performance storage into the environment across a broad

range of servers and applications. It extends the usefulness of the existing storage platform by accelerating its performance when needed. Compared to the cost of replacing the storage platform with a new platform, a purpose-built automated tiering system will be able to generate a much faster return on investment.

Third a more focused company like Avere can quickly be competitive in the market by focusing on the largest potential problem, performance, without having to reinvent the 'data services wheel'. A good example of this is found in Avere's ability to leverage both DRAM and Flash SSD in the same system, which many of the legacy providers can not do. Using these technologies in conjunction with one another, it is possible to get around the limitations of both. DRAM, while offering the highest level of performance, is more costly and volatile, meaning that a failure can result in data loss. Flash on the other hand doesn't offer the same performance, especially write performance, that DRAM does. With the ability to focus, Avere can deliver the full promise of automatically placing the data on the right tier at the right time. Inactive data can reside on the original NAS storage device, while data needing performance can move up to high speed SAS, SSD or even DRAM as the need demands. Again all of this done automatically, so the storage manager can optimize both performance and cost without spending hours a day managing the process.

Finally, a purpose-built automated tiering system allows the storage manager to continue to count on the data services that they have today. No relearning or changing the process is required. This is especially important when it comes to the data protection services like snapshots and replication. Inserting an appliance to improve data delivery performance typically does not affect the procedures already in place for data protection.

If a storage refresh is not in the immediate future for your data center but performance problems are a reality NOW, leveraging an automated tiering system that is focused on data delivery is worthy of consideration. It may even delay the future storage upgrade even further than originally planned. A purpose-built automated tiering system allows you to optimize performance with minimal risk to your existing procedures.

About Storage Switzerland

Storage Switzerland is an analyst firm focused on the virtualization and storage marketplaces. For more information please visit our web site: <http://www.storage-switzerland.com>.