

# STORAGE SWITZERLAND REPORT

## MAKING SENSE OF AUTOMATED TIERING



Eric Slack, Senior Analyst

Storage tiering is a technology solution to one of the fundamental IT challenges; data growth and cost containment. But while the concept is simple, its implementation has proved to be complex. Identifying and moving data sets between the different tiers of storage turned out to be much harder than just creating the storage tiers themselves. With SSDs, deduplicated archives and cloud storage added to the tiering mix, the process got even more complex and automation may be the answer. Making sense of automated tiering would seem to be a prerequisite to effectively leveraging storage tiers to solve cost and data growth challenges.

Automated tiering can be defined as the movement of data from one class of storage to another, without direct human intervention, usually triggered by a condition or event such as age or date of last activity. For diagrams and more explanation on this process please see our latest [screencast](#). Originally implemented in applications such as archiving, automated tiering has historically been focused on moving data from more expensive tiers of storage to less expensive ones as it becomes less active. While these early examples of automated tiering moved data 'down' in tier, newer implementations are also moving

data 'up' in tier, especially to SSD storage, in an effort to improve application performance and efficiency over disk-only systems. Another new use case for automated tiering is the movement of archive-ready data to a 'cloud tier'.

### Why Automated Tiering

Storage tiering is pretty well understood as an optimization strategy, but regarding viability and implementation, the 'devil's in the details'. How data is identified and moved is the question. Manual data migration to support a storage tiering strategy could be done in some instances, but it's not typically efficient or particularly cost effective. For many organizations, there's just too much data and not enough available time to devote to a slow, manual process, especially downward movement at the file level. Tracking data and implementing move policies manually is simply not feasible for IT organizations, considering the number of people who may be involved. With the advent of solid state disk (SSD) storage, it's more important than ever to keep the highest tiers fully utilized. For these use cases, automated tiering is an appealing alternative to a manual migration process.

## Implementation Methods

Automated tiering systems transparently place data on different tiers of storage, based on criteria set by users - like access history or age, etc. A number of disk array systems have implemented automated tiering at a block level, using the storage controller as a means of matching data usage patterns to storage types. These systems divide files up into blocks and place each of these blocks on a tier based upon the appropriate move rules. When a file is requested it can be assembled from blocks that are on these different tiers. Vendors' solutions differ in the granularity with which they create data blocks and the levels of storage supported. Some also have advanced features like QoS which prioritizes blocks and guarantees a level of availability or performance by keeping them on higher storage tiers.

While block-based tiering is relatively common, the focus has been on file-based automated tiering systems of late, more specifically NAS devices, possibly due to the predominance of file data as percentage of overall storage growth. This technology also enables tiers to be in different physical storage devices and, in the case of the cloud, in different geographic locations. File-based automated tiering is implemented as either a cache appliance or as a file virtualization appliance.

## Caching Devices

'Pure' caching appliances sit in the data path, and as the name suggests, cache, or store data temporarily to improve performance. They're designed to accelerate NAS devices supporting traditional file systems or NAS-hosted databases. Regarding use cases, this is not a 'migrate down' technology or a consolidation tool. Instead, it's used to supply RAM storage, basically a tier 0, for tier 1 arrays during periods of highest activity.

'Persistent' caching appliances also sit in-line but are used to store files for longer periods, not just during peak usage.

They typically contain fast disk and RAM or SSD and can be thought of as a performance tier for other NAS devices. Use cases are for files that need extra performance, full time, like NAS-hosted databases or storage for high-transaction VMs. These cache devices can also provide a performance tier for an existing, slower NAS, enabling it to be redeployed as an archive and extending its useful life. Like 'pure' cache, persistent cache appliances are also not a 'migrate down' technology and typically require a certain amount of modification to host file mapping to be implemented.

## File Virtualization

File virtualization systems, like the [ARX series from F5](#), abstract the physical location of a file from the end user requesting it. The file virtualization appliance receives all file requests from users and routes them to the storage devices that currently hold each file. It also can move files to different physical storage devices, based upon file attributes. This isn't a 'global file system', which is typically exclusive to a single vendor or operating system. As a tiering engine, this technology can move data up or down in tiers, and when implemented as a gateway, can even be used to send files to cloud storage. Another destination for 'archive tier' data can be a deduplication appliance, as long as it presents itself as a file system. This enables archive data to be greatly condensed, reducing storage space while eliminating the requirement for backups to those data sets.

Being independent from the storage, as opposed to the controller-based block tiering discussed previously, file virtualization can be used to logically consolidate different NAS systems or existing file servers. This target or consolidated NAS can also be a single system that contains multiple tiers of storage, even SSD. By including an SSD tier, a single NAS utilizing mid-tier disk arrays can often provide the performance necessary to replace traditional tier 1 storage. The cost savings potential is significant.

While always a compelling concept, the realities of identifying and moving data to the appropriate storage 'bucket' in a tiered architecture have limited its success for many organizations. This situation was probably made worse with the addition of SSD as a performance enhancing 'tier 0', and cloud storage as an archive option. Automated tiering may be the technology that's been lacking in the effective implementation of this storage strategy.

Block-based automated tiering, as implemented in the storage controllers of some disk arrays, provides a way to optimize the higher-cost storage in these products and improve overall performance. But file-based solutions seem to offer more potential benefits. Caching devices can accelerate performance for NAS-based applications or provide a performance tier to extend the useful life of an existing NAS. File virtualization appliances can automate the movement of files across storage platforms and across the enterprise. It can consolidate NAS devices, support all internal storage tiers and provide a gateway to the cloud as an off-site archive tier.

### **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>.