



Benjamin Woo
Vice President, Enterprise Storage Systems

Redefining the Foundation for Storage Systems

April 2008

Data storage is changing the way that information is organized, accessed, and preserved. The first generation of external storage systems, Storage 1.0, sought to provide more capacity for growing volumes of data associated with high computational applications. It leveraged newer, low-cost, high-capacity disks, and all intelligence associated with the data resided in the minicomputer or mainframe. Storage 2.0 distinguished itself by taking on a new dimension — connectivity. Organizations wanted storage systems to be a shared resource, so vendors focused on boosting the number of connections between servers and disk storage systems. Standard interconnect options emerged that allowed multiple heterogeneous high-end and midrange systems to access a common disk storage system. IDC believes that the next stage, Storage 3.0, will be based on the premise that storage systems are increasingly built on standardized hardware components, and solutions will be primarily policy driven, services oriented, and object centric. Advanced storage services, such as replication, space-efficient copy, data deduplication, and encryption, will increasingly be delivered as part of role-based storage solutions.

The following questions were posed by Xiotech to Benjamin Woo, Vice President of IDC's Enterprise Storage Systems research, on behalf of Xiotech's customers.

Q. How is enterprise data storage technology evolving, and why is this evolution important to companies?

A. IDC believes that the way users and applications will interact with data storage and the way storage systems are architected are going through yet another revolution. Until now, storage systems have been architected with applications and workloads being the primary driving fact. IDC believes that next-generation storage systems (which we term Storage 3.0) must be designed with a focus on data objects, and not necessarily applications, and driven by policy on a file-based infrastructure.

Three trends in organizations' data and information management requirements will have the greatest impact:

- Greater asset consolidation through the use of server and storage virtualization solutions that allow companies to more efficiently and rationally purchase and deploy physical IT assets
- Higher levels of application and data availability through use of the server and data mobility functions that allow companies to reduce application downtimes and data backup/recovery processes

- More intelligent, long-term handling of more diverse data pools (both file and block based) through use of cluster file systems, data tagging, data deduplication, data archiving, and data destruction technologies that allow companies to "monetize" and secure expanding information assets

In developing the Storage 3.0 concept, IDC made a critical assumption. The assumption is that a robust "protected media subsystem layer" must exist at the nucleus of next-generation storage solutions. In other words, at the core of our concept must exist a set of disk drives that have at least basic protection against faults. Today, the most common form of disk protection is RAID technology. However, IDC also recognizes that RAID may not be the only form of data protection going forward.

Q. What are the standout features of the Storage 3.0 protected media subsystem layer?

- A. The data types we store drive how we manage the data. Two basic considerations are: What do we store, and for how long do we keep it? If we associate the fact that there is a cost in terms of determining and managing what data is appropriate to destroy, then it might be easier (and potentially less costly) to simply store all the data forever.

Similarly, it could even be argued that the predictive potential of historical data may be valuable, so keeping everything forever could actually be beneficial. As a result, having a scalable, extensible protected media subsystem layer becomes paramount. IDC forecasts that content growth, particularly for the purposes of long-term data preservation, will exceed 120% through 2011. As a result, end users will demand greater durability and reliability from their storage subsystems. They will also demand performance characteristics that not only support the current capacity demands but also scale in parallel with future capacities.

The Storage 3.0 protected media subsystem layer is unique in that the Storage 3.0 concept does not define any technical standards or interconnects. Instead, it allows storage system vendors to decide the appropriate interconnect based on available technologies or cost factors. It also allows storage system vendors to decide the appropriate disk technology again based on available technologies or cost factors.

The lack of dependency on a predetermined interconnect or disk technology means that multiple technologies can persist at any given time. This provides the ultimate investment protection, flexibility, and scalability.

As technologies change, there will be a decreasing desire to perform "forklift" upgrades in order to adopt and move to newer, more effective or efficient technologies. Within the current forecast period through 2011, while the capital cost of technology acquisition will be relatively flat, expenditure on the maintenance and management of technology will increase dramatically. As a result, customers will demand a predictable and manageable total cost of acquisition over a longer useful life.

Given the wide variety of available storage media types, from solid state disks based on DRAM to very dense, capacity-optimized magnetic disk drives, optical platters, and linear tape, the performance characteristics of any given piece of data can be addressed in the most granular way possible.

Q. What are the business benefits of next-generation storage?

- A. Two key notions behind the Storage 3.0 concept are the ability to derive cost efficiencies through process improvement and the ability to improve value derived from persisting data.

Keeping data for the sake of keeping data is costly. Keeping data ultimately requires management. Given the time scales required by certain regulations such as HIPAA, data will ultimately require migration from one generation of storage system to another. As a result, vendors that focus on the extending the generational time frames of product life cycles will be better positioned. In fact, it is arguable that no organization can drive enough cost out of data persistence or management to make the economic equation viable. So if one cannot drive enough cost out, then it is imperative that one derive value from the data. The existence of software applications such as business analytics, data mining, and decision support systems provides the ability to sustain value from historical data over time. In fact, if predictive analysis based on historical data is effective, the value from historical data can actually exceed that of active data!

The Storage 3.0 concept attempts to address these issues by suggesting the use of policies that tightly integrate into existing directory structures and security policies to provide the capability of equating the value of data with the value of the underlying disk.

An example might be as follows. In a given environment, the current disk storage in the protected media layer is made up of 500GB SATA disks. The organization's data policy is defined whereby a certain set of data must reside on the lowest-cost disk storage available, so that set of data will therefore persist on the 500GB SATA drives.

However, six months later, a new subsystem made up of 1TB disk drives is introduced to the protected media layer and as such has a lower \$/GB. That set of data would automatically be migrated to the new 1TB drives to comply with the organization's data policy. Predictable and manageable support costs and longer useful lives (through extending product life cycles) will help to further lower the total cost of ownership. This in turn will serve to further drive out the cost involved with long-term data persistence.

In doing so, cost efficiencies can be maintained and data can be persisted without human intervention.

Q. What needs to happen for Storage 3.0 to become a market reality?

- A. Perhaps the single greatest challenge for the Storage 3.0 concept to become a market reality is integration of the various layers introduced by the Storage 3.0 concept. In many cases, further technology convergence still needs to occur. In particular, greater convergence between data security, network security, and directory structures needs to occur to enable a global policy engine to operate effectively.

With Storage 3.0 comes also the concept of object-based storage. While file systems are sufficient surrogates for a true object-based storage environment, many enhancements need to occur that can exist only through the development of standards.

While much of the focus will be on the value storage services layer, the Storage 3.0 concept could not exist without a strong, scalable, extensible protected media layer. Some forward-looking vendors have recognized the need to develop and deliver a new foundation for storage that focuses on reliability, performance, and scalability. Storage systems are available today that can virtually eliminate drive-related service events, for example, while delivering performance that scales with capacity.

ABOUT THIS ANALYST

Benjamin Woo is Vice President of IDC's Enterprise Storage Systems research. He is responsible for advising IDC's clients on the evolution of and trends relating to data storage systems; the impact storage systems have on adjacent technologies, including servers, software, and virtualization; and best practices in the deployment of storage systems in relation to numerous industry dynamics.

ABOUT THIS PUBLICATION

This publication was produced by IDC Go-to-Market Services. The opinion, analysis, and research results presented herein are drawn from more detailed research and analysis independently conducted and published by IDC, unless specific vendor sponsorship is noted. IDC Go-to-Market Services makes IDC content available in a wide range of formats for distribution by various companies. A license to distribute IDC content does not imply endorsement of or opinion about the licensee.

COPYRIGHT AND RESTRICTIONS

Any IDC information or reference to IDC that is to be used in advertising, press releases, or promotional materials requires prior written approval from IDC. For permission requests, contact the GMS information line at 508-988-7610 or gms@idc.com. Translation and/or localization of this document requires an additional license from IDC. For more information on IDC, visit www.idc.com. For more information on IDC GMS, visit www.idc.com/gms.

Global Headquarters: 5 Speen Street Framingham, MA 01701 USA P.508.872.8200 F.508.935.4015 www.idc.com