



Snapshot
6 June 2002

IBM Articulates a Storage Strategy for Block Virtualization

By Joyce Tompsett Becknell

IBM is tackling the difficult issue of growing complexity in storage management. Leveraging its knowledge of RAID controller design combined with experience in server and clustering architectures, IBM recently announced plans for several forthcoming products, which address various aspects of storage networks. This report looks at IBM's plans for a storage virtualization engine and its use in the company's customer base.

The Storage Conundrum

Conventional wisdom states that storage in the enterprise doubles year on year. Unfortunately the storage resides on technologies of varied capability and vintage, dedicated to numerous applications. Due partly to relatively low cost, but mainly for reasons of application incompatibility, IT managers have tended to overspend on storage when purchasing new systems, with an end result of excess capacity going unused or inefficiently deployed.

In recent years, many IT managers turned to storage area networks (SAN) to manage their mushrooming storage. SAN solutions detach storage devices from individual servers and relocate them in pooled arrays where resources can be more easily managed and capacity better utilized. As a result, intelligence has migrated to the storage controller, which coordinates advanced functions such as file system to storage block level mapping, failover, data migration, or copying. The benefit of this technology is that storage is centralized in fewer locations, making it easier to manage. Additionally, as an application grows, more storage can easily be allocated from the SAN as most products support Linux, UNIX, and Microsoft operating systems.

While SANs enjoy continued success, some enterprise customers have begun to encounter manageability challenges as their SANs grow larger or more diverse. As the number of arrays grows or the variety of controllers used within a SAN increases, the ability to manage across its breadth and depth becomes more difficult. With these customers in mind, IBM has announced a solution that aims to address these environments.

The IBM Block Level Virtualization Solution

In keeping with the Storage Networking Industry Association (SNIA) model of block aggregation, the IBM virtualization engine will sit in the network, not at the host or inside the RAID controller. Instead, the virtualization engine is located in the data path, with the primary goal of moving intelligence from individual controllers into the network thus, in a sense, becoming the controller of storage controllers.

The IBM product will provide clusters of paired nodes to manage the pool of disks. Each node is xSeries-based (the IBM Intel-based server) with up to 4GB of read/write cache mirrored across the pair of nodes. Virtual volumes are then shared across the nodes. Customers purchase a minimum of one pair of nodes, and the first iteration of the product will scale up to four pairs of nodes in an individual cluster, with future plans to scale higher.

In essence, IBM has designed the virtualization engine to replace the function of the intelligent storage controller that sits on each individual storage array. It provides many of the features customers currently expect from a controller. A dual path option is provided in the SCSI driver that handles auto-failover if a node crashes. Additionally, nodes work in pairs. When a node receives a write, it writes it not only to its own cache,

The Sageza Group, Inc.
836 W El Camino Real
Mountain View, CA 94040-2512

650-390-0700 fax 650-649-2302
London +44 (0) 20-7900-2819
Munich +49 (0) 89-4201-7144
Amsterdam +31 (0) 35-588-1546

Copyright © 2002 The Sageza Group, Inc.
May not be duplicated or retransmitted without written permission

sageza.com

but also to its partner node's cache before it sends an acknowledgement of the write. Should one node in the pair die, the remaining node will empty the write cache to the storage device and then proceed in write-through mode. There is also an auto restart of the node and then re-admission to the cluster once the node is running again.

IBM claims their philosophy is to create a product that is robust enough for an enterprise environment, while at the same time being simple to manage. One difference between this and other similar products is that IBM has clustered the nodes rather than assigning a portion of the storage to each node. This means that each node sees the complete storage environment. This is crucial to administration and management planning.

IBM believes this architecture is appropriate because:

- The engine sits in the network. It avoids the communication problems of multiple device drivers talking to multiple operating systems or OS versions. While an out-of-network implementation is not necessarily a problem when the system is running as planned, it can make for difficult coordination when something goes wrong.
- All of the data can be seen by all of the nodes, which allows for better policy-based management — the current holy grail of the storage management world.
- Multiple types of RAID controller are supported in this environment, which is the likely real-world scenario. This solution does not require that all RAID controllers be IBM products.

Who Benefits from Block Level Virtualization

The goals of this product are threefold:

- First, this architecture should increase storage administrator productivity, because individual servers are isolated from the physical assets of storage. Managing, adding, or migrating disks can be done transparently to the application.
- Advanced functions are now provided from a common platform. Disaster recovery (or other functions) can also be accomplished from a single point in the SAN in a common way for all data.
- Capacity utilization is arguably improved. Extra capacity can be reallocated as needed, even across differing storage systems, and this is transparent to the application, the operating system, and the hardware.

Broadly speaking, enterprise customers will find their SAN environments falling into two categories:

- Those that require monolithic storage growth supported by large, intelligent RAID controllers. These customers are likely to be using IBM's Shark or other similar products.
- Those that require more modular or incremental storage growth. These are the customers who will find the most benefits from the virtualization engine. This includes:
 - Customers with large numbers of arrays that run on multiple platforms (such as mixed UNIX and NT environments);
 - Customers with multiple RAID controllers from multiple vendors with varying levels of controller intelligence;
 - Telcos and ISPs that are looking to add nodes in a modular fashion.

Customers with storage environments resembling those mentioned above who are looking to gain greater robustness and manageability of data storage over the next year should check with IBM. This product is expected in 2003 but planning requires that customers begin to evaluate large storage environments now to check solution appropriateness for their needs.