TANEJA GROUP
TECHNOLOGY ANALYSTS

P R O D U C T   P R O F I L E

IBM System Storage SAN Volume Controller 4.3

Proven, Enterprise-Ready and Scalable Storage Virtualization Solution

June 2008

THE NET NET

Storage virtualization has come a long way in the past seven years. After a false start in 2001, fraught with inflated expectations and product deficiencies, the category fell into infamy. Several vendors disappeared, many others repositioned themselves to focus on the Small Medium Business (SMB) space and yet others reinvented themselves with completely different products. Only one company stayed true to the promise of virtualization from the very beginning. That is IBM. With the SAN Volume Controller (SVC) product, launched in July, 2003, the company nurtured the market, in spite of the fact that many in the market didn't even want to say the V-word anymore. IBM persisted, albeit mostly with their own customers, fundamentally because the customer could see the potential of storage virtualization and could count on IBM to support them through the early learning cycles.

The payoff for IBM is huge. They recently boasted the shipment of their 4000th SVC system. SVC is a mature, enterprise-proven product that has demonstrated proven ROI to its customers. Moreover, IBM has shown that SVC and its in-band architecture can indeed scale to handle the largest, most stringent enterprise SAN environments. By doing so, IBM has led the market where others have only slowly followed. IBM's efforts have in fact changed the market, and now it is filled with solutions for storage virtualization. But a casual glance around at the success of the other solutions in this market is telling. EMC is carrying Invista, HDS has successfully brought USP to market, and LSI has acquired the potentially-network-based but practically disappeared StoreAge. HP and Sun resell the controller-based HDS solution, and Dell still has no offerings. While the HDS solution has significant traction, all of the other solutions are still in their infancy when customers are counted. The bottom line is IBM has paved the way to show customers the value of virtualization to the point that the V-word is back in the vocabulary of all storage vendors who have rapidly tried to deliver solutions to the market over the past few years. But the truth is, none of these solutions come close to the success and maturity demonstrated by IBM's SVC.

The value of storage virtualization is unquestioned. It helps rein in storage capital expenses (CAPEX) and operational expenses (OPEX) that are otherwise running amok. It provides a forum to perform storage management in a consistent fashion even while the underlying physical storage is heterogeneous and possesses its own idiosyncrasies. In our view, it is also a key building block for the next generation data center that will focus on delivering a variety of services. IBM knew that and held steady. We believe the payoff until now is a shadow of what it is to come, as IBM ties storage virtualization to other efforts, such as server blades and server virtualization.

TANEJA GROUP
TECHNOLOGY ANALYSTS

PRODUCT PROFILE

In this Product Profile, we briefly define storage virtualization, the various methods of delivering it and the key benefits it brings to the data center. Then we focus in on IBM's SVC, its architecture, its product features and why it has succeeded while others have failed or are struggling. Then we describe where we see it going in the future. There is no better measure of success than customer traction. On that basis IBM stands alone.

The Storage Management Nightmare

It is no secret that the job of the storage administrator has gotten a lot harder over the past decade. Much of the reason for why storage management has become so cumbersome can be traced back to five fundamental challenges that exist in most enterprise data centers.

Challenge #1: Rapid Capacity Growth

IT departments are being asked to store more information longer. The current paradigm of storage management is on a collision course with exponential data growth. IT organizations have a decision to make: they can add more people to keep up with capacity or find a better, more efficient way to manage storage within the current staffing limitations. Something has to give. Taneja Group believes strongly that storage virtualization is a necessity in order to keep up with the ever increasing capacity requirements. Storage virtualization will not be optional in the next generation data center.

Challenge #2: Poor Storage Utilization

Exacerbating this data growth is the fact that the deployed storage capacity is not readily accessible to the hosts that need it. Existing storage practices over-provision disk for performance, because the penalty for running out of capacity is high, and because over-provisioning reduces the need for repeated provisioning in the future. Hence, typical storage utilization rates in most enterprises run in the 25-40% range. Today, low utilization creates more burden than ever before by consuming precious and expensive power, and creating unnecessary heat.

Challenge #3: Tiered Storage

Storage administrators are being asked to wring costs out of their infrastructure by ensuring that the data is stored on the most cost-efficient media possible. Typically, the value of data decays over time. Therefore, it does not make sense to store seldom accessed information on the highest cost storage systems and media. To cut costs, storage administrators must create tiers of different types of storage based on performance and cost per capacity ($/TB) metrics. They must continually ensure that the data is stored on the most efficient storage available, redistribute data among disk types (FC, SATA, SAS, enterprise, mid-range, near-line, MAID), and ensure that protection practices such as replication are consistently maintained across tiers. In truth, migrating data is disruptive, and then breaks many complex relationships between replicas and data protection systems. This makes storage tiering extremely complex if not impossible, and often thwarts data tiering efforts and the enormous potential cost savings.

87 Elm Street, Suite 900 Hopkinton, MA 01748 Tel: 508-435-5040 Fax: 508-435-1530 www.tanejagroup.com

Challenge #4: Non-Disruptive Data Migration

In today's world, IT systems are expected to always be operational. However, storage administrators are often required to take storage offline to migrate data between arrays or change the storage infrastructure. In fact, storage administrators are expected be able to do technology refreshes, vendor/equipment swap outs, and re-configuration activities as part of routine data center and storage maintenance. These actions prevent applications from accessing data and thus increase application downtime. The cost of this downtime can dramatically impact a corporation's bottom line and its reputation. Therefore, storage administrators need a way to perform storage changes and data migrations between arrays and different types of storage media while still maintaining continuous availability for the applications and their data.

Challenge #5: Data Protection & Disaster Recovery

Crafting a disaster recovery and data protection plan for enterprise data is a non-trivial endeavor. The management of snapshots, backup, replication, and mirroring technologies imposes a tremendous level of administrative complexity on the storage organization. Storage administrators must now protect each application and its data and cope with the nuances of a variety of heterogeneous storage array vendors and products. Furthermore, DR and data protection compound the already acute storage management problem.- an administrator must now cope with managing two copies of the same data across two locations while ensuring its consistency.

Storage Virtualization Defined

As Taneja Group defines it, storage virtualization abstracts the idiosyncrasies of the individual storage devices and provides a single management point for all storage devices. At its core, the storage virtualization layer pools physical storage from multiple, heterogeneous network storage devices and presents a set of storage volumes for hosts to use. In addition to creating storage pools composed of physical disks from different arrays, storage virtualization provides a wide range of services, delivered in a consistent way, that include:

- Basic volume management(including LUN masking, concatenation, volume grouping and striping)
- Support for tiered storage
- Non-disruptive data migration
- Data protection and disaster recovery functionality, such as snapshots and Asynchronous, Semi-synchronous, or Synchronous mirroring

Storage virtualization alleviates the five challenges listed above of managing a large enterprise-class storage environment.

Virtualization Architectures

In a virtualized SAN fabric, there are three ways to deliver storage virtualization services: in-band, out-of-band or split path architecture for intelligent devices (SPAID). Before we delve into the architectural specifics, it is important to understand that a typical I/O path can be deconstructed into three separate paths or streams – the metadata, control, and data path. The metadata path controls mapping between virtual volumes and physical devices. The control path maintains the interface between the metadata path and the data path software. Lastly, the data path contains the actual information that needs to be transmitted between host and storage.

In the case of in-band architectures, the metadata, control, and data path processing are all performed by the same computing element – typically an Intel-based appliance. In other words, all three are "in the path." For out-of-band implementations, the metadata management and the control path processing are both performed by a separate compute engine, distinct from the compute engine that processes the data path software. Software (agents) must be installed on each host and these are responsible for the high performance direct transfer of data from/to the host to/from storage. Lastly, a SPAID architecture is a combination of in-band and out-of-band approaches together. A SPAID system leverages the port-level processing capabilities of an intelligent switch alongside a separate device for control path and/or metadata management, and thus eliminates the need for host level agents, while maintaining a separation of control and data paths.

Where Virtualization Lives

Storage virtualization services, like volume management, snapshots, and replication, can reside at the host, network, or storage device level. Traditionally, storage intelligence has lived at either the host-level with a software volume manager, like Veritas Volume Manager, or in the RAID controller of a storage device.

However, with the advent of network-based storage virtualization, this intelligence is being pushed into the network. At the network layer, there are three broad platform categories for delivering storage virtualization services in the storage area network (SAN) fabric: general-purpose appliances (e.g. IBM SVC, FalconStor IPStor, Datacore SANsymphony), purpose-built appliances (PBA) (e.g. Emulex 765-S), Array controllers (e.g. Hitachi Data Systems USP-V and USP-VM), and intelligent switches (e.g. Cisco MDS 9000).

Meet IBM SAN Volume Controller (SVC)

IBM System Storage SAN Volume Controller (SVC) 4.3 is the crown jewel in IBM's storage strategy and portfolio. IBM launched SVC in July, 2003. To date, IBM has shown impressive market adoption, out executing the early pioneers and many of the nimble startups that initially evangelized the storage virtualization product category.

IBM boasts a fourth generation product with over 4000 SVC systems shipped and five years of market acceptance. SVC is widely viewed as the most successful storage virtualization in the market today. We think these statistics speak to the maturity of the product, the value of a high touch consulting approach, and the superior execution of the IBM Storage business unit.

The Product

IBM and its Business Partners sell the SVC software as an appliance-based system that combines IBM SVC software and hardware. SVC 4.3 supports 4Gbps Fibre Channel SAN speed, but can auto negotiate down to 2Gbps speed, if necessary. The SVC software can support up to 1024 attached hosts running virtually any operating system including Microsoft Windows, HP-UX, IBM AIX, Sun Solaris, Linux and VMware and can export up to 8,192 virtualized LUNs. As well, IBM sports one of the largest storage array support lists on the market, encompassing over 130 different storage systems.

The SVC is based on a clustered, redundant, highly scalable architecture. SVC is deployed only in clustered pairs of servers—each with 8GB of cache—running SVC software. A pair of SVC servers (or "nodes") is known as an I/O group. In order to ensure redundancy, a single SVC node is not a supported configuration. Adding another I/O group (that is, two SVC nodes) can increase cluster performance and bandwidth. A maximum of 4 I/O groups (a total of 8 nodes) and 8PB of capacity can be added together in a single SVC cluster. Therefore, customers can start small and scale as their storage needs and I/O throughput profile changes over time.

SVC comes with the full breadth of features most users expect from a storage virtualization product, including snapshots in the form of IBM's FlashCopy, thin provisioning by way of Space Efficient Virtual Disks (SEVs), and synchronous or asynchronous replication via Metro Mirror or Global Mirror. These features are in addition to the fundamental storage capabilities of storage pooling, provisioning, volume mirroring, and volume copying/migration.

SVC Architecture

IBM SVC is what is known as an in-band virtualization appliance. Traditionally, the knock on in-band virtualization appliances has been that they introduce a "bump in the wire" and could not scale to match the level required of enterprise-class SAN environments. However, IBM SVC has allayed these scalability fears and proven that an in-band approach with a large amount of mirrored cache can indeed scale to handle the most stringent enterprise IOPS and availability requirements. For example, SVC boasts the highest IOPS rate (272,500 IOPS) of any storage virtualization system as measured by the Storage Performance Council SPC-1 benchmark. The SPC-1 benchmark simulates read-write random I/O workloads, like Online Transaction Processing (OLTP) databases and email and is an accurate approximation of a typical enterprise I/O pattern. SVC has also laid claim to the top spot (7.08 GB/s) in the SPC-2 benchmark, a benchmark that simulates large, sequential I/O processing. More detailed information regarding SVC SPC-1 and SPC-2 results is available at http://www.storageperformance.org/results.

Moreover, IBM's SVC team has a fanatical focus on availability, and their efforts pay off for customers. To date, field data across all deployed SVCs demonstrates availability comparable to or better than enterprise-class disk systems.

These performance and availability results are a true accomplishment. IBM, through its focus on performance, scalability, and availability approaches has shown that an in-band approach to virtualization can, in fact, scale and meet the availability demands of the large enterprise.

IBM SVC's highly scalable in-band architecture, and robust feature set and capabilities provide five crucial benefits to end users that differentiate SVC from the other virtualization players.

Benefit #1: Heterogeneous Storage Pools

SVC provides the ability to centrally manage all storage as a single pool or multiple pools. From a pool, SVC presents one or more virtualized disks, potentially composed of storage from different vendors. A storage pool simplifies the provisioning of storage to applications and hosts and allows under-utilized capacity to be freed. Although all storage virtualization players virtualize LUNs from different vendors and support storage pooling, IBM has the wherewithal and the relationships to ensure that a customer is supported and can combine different vendors' storage without hassle or incident. With release 4.3, SVC has further broadened what was already a broad storage subsystem support matrix. SVC currently supports storage from all major vendors including Dell, HP, IBM, EMC, Sun, NetApp, and Hitachi. Furthermore, IBM Subsystem Device Driver (SDD), an included software component of SVC, supports multipathing across all types of storage. In the past, early virtualization adopters faced added deployment complexity because they had to use a different multipath I/O solution for each vendor's storage. The included SDD also translates into significant savings for an SVC customer, given that multipathing software is quite expensive and is required for each supported server/storage device.

While SVC has always enabled heterogeneous storage with SVC-based replication and volume provisioning, with version 4.3 of SVC, IBM also extends thin provisioning – IBM's Space-Efficient Virtual Disks (SEV) – and space-efficient, delta-only incremental FlashCopy across heterogeneous storage ("Space-Efficient FlashCopy"). The benefit is that now, every pool of storage in the enterprise can be enabled with the newest capacity optimization features available, and managed with a single consistent toolset.

Simply put, few solutions on the market can bring this level of feature set to bear on such a broad list of different storage systems and hosts.

Benefit #2: Central Control Point for All Storage Devices

SVC simplifies the storage administrator's job by providing a central, single interface for managing how storage is allocated and consumed by hosts. With SVC, the administrator now has a central control point for seeing which hosts are consuming what storage, the different performance characteristics of the various storage volumes, and how much capacity is available for provisioning. In addition, SVC is based around SMI-S so it is open and integrates with existing storage management tools that may already be deployed in the data center.

This makes storage management in a heterogeneous environment simpler and allows administrators to manage more TBs.

Benefit #3: Tiered Storage

SVC supports the ability to create a single pool or tier of storage or if need be, create multiple tiers or pools of storage. With SVC enabled storage tiering, enterprises can create multiple storage pools with different characteristics that meet different SLAs and price/performance characteristics. This is a key building block for Information Lifecycle Management (ILM). For example, one pool might contain all high performance Fibre Channel disks and another pool might contain SATA disks for a better cost/ capacity ratio. Another pool may have more stringent DR requirements and thus merit asynchronous replication to another data center.

Moreover, in our view, SVC addresses two of the primary deterrents to migrating to tiered storage architectures in the enterprise. First, storage managers are often challenged to implement storage tiering because heterogeneous storage systems will wreak havoc with array-specific replication and backup scripts, and increase ongoing management overhead. Second, storage managers are resistant to tiering their data because any oversights when data is moved can have detrimental performance impacts on key applications. SVC creates tiers with a single uniform layer of management, making a single approach to replication and data protection work across all storage, and mitigates the impact of lower performing tiers with SVC's cache.

SVC, working with IBM TotalStorage Productivity Center, provides the ability for an administrator to automatically migrate data from one tier to another, based on a stated policy, and without disruption to applications. Storage administrators can ensure that the least used/less valuable data resides on the cheapest storage while the most mission critical data resides on the highest performance, most highly available storage. By putting this intelligence in the fabric, the storage administrator has a single point to implement and tweak ILM policies across the entire storage infrastructure.

Benefit #4: Non-Disruptive Data Migration

SVC allows a storage administrator to migrate data from one storage tier to another regardless of whether the storage is from one supplier or not. This migration occurs without any downtime to the application or the user. The application interface does not change either. There are serious benefits that accrue with this facility even when the storage is all from one vendor. For example, even in an all IBM environment, the administrator can non-disruptively migrate data from a DS4000 to a DS8000 system.

By empowering the administrator to seamlessly migrate data, SVC allows customers to break their dependence on a single storage supplier. This is by far the most sought after feature from any virtualization product and is amply proven in the case of SVC.

Benefit #5: Thin Provisioning, Data Protection, and Disaster Recovery in the Fabric

Storage virtualization in the fabric provides the ability to implement data protection and disaster recovery strategies at a central point across the entire storage infrastructure. This simplifies the complexity and hassles of managing different vendors' technologies and ensures that LUNs are protected with the appropriate SLA level.

SVC provides both volume-level snapshots and remote replication services for all virtualized disks. IBM refers to its snapshot functionality as FlashCopy. FlashCopy provides point-in-time copies – either full volume copies or copy on write. Every FlashCopy replica can be used as a fully writeable volume, with up to 256 FlashCopy targets per source volume. FlashCopy consequently is a crucial technology for both storage utilization and data protection. Since FlashCopy replicas are copy-on-write and can be connected to SEVs, administrators can use FlashCop replicas written on thin-provisioned volumes to massively reduce wasted storage space (space required for replicas can be reduced by 75% or more). This is especially true for virtual servers, where a single golden boot image can be copied multiple times to create thin-provisioned, fully modifiable boot images for every virtual guest in the enterprise. Moreover, administrators can use SVC's Space-Efficient FlashCopy to take periodic point-in-time (PIT) copies that incrementally capture every change to a volume since the last time a FlashCopy was executed. Since the captured changes are only incremental, these FlashCopy replicas then provide highly space efficient, multiple PIT copies for data recovery or testing. The space efficiency of these FlashCopy replicas allow the enterprise to keep many more PITs on-line for either volume mounting and file recovery, or nearly instantaneous rollback of a complete volume in the case of data corruption. Additionally, a backup application can mount these FlashCopy replicas and backup the data to tape for offsite safe-keeping. While thin-provisioning alone is often expected to increase storage utilization by as much as 50%, when coupled with space-efficient FlashCopies in testing and development or server virtualization environments, we expect end users will see even greater savings.

In addition, SVC supports remote replication in two ways: Metro Mirror and Global Mirror. Metro Mirror is a synchronous remote replication technology that typically operates within a Metropolitan Area Network (MAN) whereas Global Mirror is an asynchronous remote replication technology that is not distance limited and can replicate data across a metropolitan area network or wide area network. Both functions operate in the same manner and copy the changes made to a virtualized volume between two SVC Clusters in different locations (they may also be used in an "intra-cluster" mode with just a single cluster, often used for testing). Metro Mirror and Global Mirror are critical building blocks in any disaster recovery strategy where a company must recover quickly with minimal data loss.

It is important to understand that the bottom-line value of these capabilities in the network is tremendous. Without virtualization, storage managers must acquire and maintain separate data protection and replication feature licenses for each array in their enterprise. Moreover, when arrays are replaced, such licensing must be purchased anew. With these features provided through a uniform network layer like SVC, several benefits are apparent: licensing is simplified; the costs of feature licenses for new storage arrays are avoided; and on-going operational expense from management overhead associated with multiple dissimilar tools is greatly reduced.

Taneja Group Opinion

Of the major storage vendors, IBM has been the most aggressive in terms of evangelizing the need for storage virtualization and helping to drive awareness and adoption of the technology.

From our vantage point, we see three crucial keys to IBM's success – a comprehensive product feature set designed for the enterprise, a device support matrix that is second to none, and the backing of IBM Global Services and a large army of IBM Business Partners.

First, IBM SVC is a fourth generation product that has prospered where other storage virtualization products have struggled. SVC is a truly enterprise-tested product as evidenced by the more than 4000 systems deployed with customers. IBM has shown through those customer implementations and its strong performance benchmarks that SVC can scale to meet the largest, most challenging SAN environments. Moreover, SVC brings a feature set to any SAN environment that encompasses every feature that enterprise users demand from a storage virtualization layer, including: uniform multi-pathing; thin-provisioning; copy-on-write, fully modifiable snapshots; volume mirroring; and asynchronous and synchronous replication. There is no other storage virtualization product in the market today with such a track record of performance, customer adoption, and systematically growing capabilities.

Second, although this may be blindingly obvious, a major determinant of SVC success has been that IBM has truly built a comprehensive device support matrix that includes all major storage suppliers and operating system providers. It is easy to say you are open, but it is another thing altogether to have the rest of the ecosystem acknowledge you as an open vendor. This is a testament to IBM's wherewithal within the industry. Without a comprehensive matrix, customers would not have the confidence to make strategic commitments to SVC as the linchpin of their storage virtualization strategy.

Third, IBM can deliver the entire end-to-end solution and ease customers through the transition to a new approach to storage management. Many of IBM's startup competition have been unable to field the services capability needed to migrate customers to a new virtualized infrastructure. A trusted professional services organization like IBM Global Services (IGS) or one of IBM's Business Partners can get customers over the line and enable them to realize a tangible ROI – a key determinant to future success. The services and support backing of

IBM and its Business Partner community is a significant differentiator for SVC.

Virtualization technology delivers many benefits, as we have outlined above. But it is an invasive technology in that it modifies the format in which data is stored on the physical devices. In other words, using a specific virtualization product is always a strategic decision for the customer. This is particularly true when the virtualization device is an in-band one, such as SVC. If it doesn't work flawlessly and scale well the impact on the IT infrastructure can be dramatic. IBM seems to have solved these issues and the product has reached a level of maturity and broad acceptance. IBM is uniquely capable of benefiting from this maturation process. Customers can be confident that IBM will stand behind their product and be accountable for problem resolution between SVC and the rest of the ecosystem.

We are optimistic that storage virtualization technology, in general, has matured to the point where it is delivering on the kind of hard ROI savings that had been promised five years ago. In the case of IBM, we have spoken with several adopters who have reaped significant economies from SVC in their environments.

Moreover, storage virtualization today stands poised to be a key enterprise technology moreso than ever before. That is specifically because of the runaway adoption of server virtualization, and the tremendous benefits that storage virtualization can deliver in virtual server infrastructures. While we've identified how thin provisioning can benefit virtual server environments, storage virtualization can also enable higher levels of availability, uniform and host efficient replication, and make the virtual infrastructure more flexible with easy, broad-reaching virtual guest movement and storage tiering. It is easy to envision adaptable virtual infrastructures that can dynamically move virtual guests across different performance tiers of storage in response to constantly changing processing demands.

Taneja Group has always been a strong proponent of the trend toward intelligence moving into the network. It has never been a question of if, but only a question of when mainstream adoption of network-based storage virtualization would reach critical mass. We think that it has. Taneja Group continues to believe that core volume management, replication, mirroring, data protection, and disaster recovery services will ultimately migrate from the host and storage array into the fabric. That is not to say there is a "winner take all" dynamic operating in terms of where different services ultimately reside. Clearly, there are many storage services that are best done at the host (e.g. file system level snapshots) or on a storage device (e.g. RAID). Nonetheless, we envision a world where services, such as those mentioned above, will reside predominantly in the network.

In short, storage virtualization is coming of age and SVC is well positioned to capitalize on that trend. IBM is already the dominant supplier of storage virtualization today and has the most mature, robust, widely supported offering in the marketplace. They are a formidable force and customers are rewarding them for it. But IBM's success has not gone unnoticed by their competitors. Other vendors have accelerated their product introductions and development now that IBM has blazed the trail to success. But, as we all know well by now, it is hard to catch the vendor that has a five plus years lead on others. Any way one looks at this it is ultimately IBM's customer that wins!