Product Brief

Storage Virtualization isn't About Storage

Date: January 2010 **Author:** Steve O'Donnell, Senior Analyst Data Center Practice

Abstract: *Virtualization is at the center of all 21st Century IT systems, yet many CIOs fail to fully understand all of the benefits it can deliver to the data center operation. When we think of virtualization, we think compute, network, and storage—and we mostly think about driving up utilization on each. Storage controllers have always offered the ability to carve out pieces of real storage from a large pool and deliver them efficiently to a number of hosts, but it is storage virtualization itself that offers improvements that drive operational efficiency. IBM has been quietly addressing storage virtualization with SAN Volume Controller (SVC) for the last six years, building up a significant technical lead in this space.*

Overview - Storage in the Data Center

Strange but true: most infrastructure architectures are deliberately designed from the outset to need little or no change over their lifetimes. There are two main reasons for this:

1. Change often means outages and customer impact and must be avoided
2. Budgets are set at the beginning of a project and getting more cash later is tough

Typically, then, applications are configured with all of the storage capacity they need to support the wildest dreams of their business sponsors (and then some extra is added for contingency by IT). Equally, storage is always configured with the performance level (storage tier) set to cope with the wildest transactional dreams of the business sponsor (and guess what? IT generally adds a bit more for good measure.).

No wonder storage is now one of the largest cost components involved in delivering and running a business application.

Endemic Over-provisioning

The storage salesperson loves this approach. Always over-provisioning gets her more volume at a higher price than the much more sensible approach of buying what is really needed, just in time, would. In fact, many vendors' storage sales strategies are aimed at just that issue, making it easy for the customer to over-provision.

Actually, the work behind capacity and performance management is boring, tedious, and needs active participation from both the business unit and application development groups. Many organizations cut those IT service management jobs a few business cycles back as no one would notice immediately. The infrastructure architect knows this and builds in so much additional performance and capacity that it covers for the lack of attention to the capacity and performance management process.

Here is a real life (and personal) example of what actually happens on the ground:

An application has just been rolled out that, at launch, actually needs 20 GB of storage and enough performance to support 3 transactions per minute, peak. It is provisioned with 700 GB of storage on Tier 1 (320 GB, 15k Fiber Channel disks)—enough performance and capacity to deal with the best possible business case in 10 years.

When the application got ported to a new hardware platform five years later, it was actually using 40 GB of storage at 60 transactions per minute, peak. To put some perspective around the commercial impact, over the five year hardware life, this level of over-provisioning consumed more than 6 megawatt hours of excess electricity and thousands of dollars in capital depreciation and maintenance costs.

This example is repeated again and again across IT systems, making the cost of failing to do the job right, billions of dollars of waste, and many gigawatt hours of wasted energy and unnecessary data center plant.

Even if some capacity and performance management had been performed that was able to identify in-life just how over-provisioned applications are, the difficulty of migrating data and applications to a lower tier of storage, plus the resultant service outage, would make the exercise impossible to justify.

Operational Paralysis

In many high risk, high impact environments, such as wholesale banking and manufacturing, (where application outage costs can be greater that the cost of the hardware), provisioning totally separate infrastructure for each business unit is an entirely accepted practice. The thinking is that by separating the infrastructure, change risk can be mitigated and obtaining change approval or permission from the business to schedule an outage is easier.

The problem here is not the change itself, but the impact of the change. Changes that reduce service availability are unpopular, are difficult to manage, and sometimes significantly delay implementation. Software and firmware patches often can't be installed; other times, unreliable hardware can't be replaced—all because doing so would cause a service impacting outage.

Well run IT infrastructure organizations set themselves up to avoid the need for changes that impact service availability. Unfortunately, changes to storage have been notoriously difficult to execute without impacting servers and applications. Typically, applications are very unforgiving of data unavailability, even if the outage is only for a short time.

The Continuous Migration Architecture

In addition to improved utilization, virtualization offers the prospect of isolating the application from the underlying hardware to such an extent that application migration is viable in real time as part of normal business operations. At ESG, we refer to this as the "continuous migration architecture (CMA)." It delivers simultaneously on operational flexibility, business agility, and reduced costs. In effect, a CMA is the embodiment of the private cloud, including compute, storage, and network virtualization.

What Value Does Virtualization Deliver?

Storage virtualization is all about operational efficiency, service levels, and flexibility. Consolidation is nice, but it is only a small piece of the overall value. Storage virtualization, on the other hand, is NOT about disconnecting the logical from the physical—*it is actually about breaking the link between the physical storage and the application*. It is about being able to perform normal maintenance activities with no service impact—and being able to make very significant changes to the underlying physical environment without anyone noticing. Compute, network, and storage virtualization technologies all deliver two key benefits:

- Operational flexibility – the ability to separate the application from the underlying physical hardware
- High utilization of assets – sharing the resources between multiple consumers

It is the first benefit that makes storage virtualization important for the CIO. By separating the physical storage from the application, significant changes can be made without a service outage. The quantity of allocated storage can be changed transparently, storage volumes can be moved from one tier of storage to another, and whole storage controllers can be swapped in and out—all without any impact whatsoever.

The second benefit is also a benefit of the standard storage controller: enabling the sharing of a large pool of disks between many consumers in a resource efficient way. Uniquely, IBM SVC adds the ability to concatenate stranded storage between controllers, reducing wastage and improving storage utilization.

IBM SAN Volume Controller (SVC)

Like many Tier 1 storage controllers, SVC offers enhanced features that further improve the utilization of storage resources (such as thin provisioning) while improving their resilience (via synchronous and asynchronous replication).

Additionally, SVC offers this functionality across all supported storage controllers. In many cases, functionality such as replication support between different vendors' equipment or, in some cases, between different models of the same vendor's storage controllers is not generally available elsewhere. SVC acts as an integration layer, offering a single administrative toolset to the multi-vendor storage administrator.

IBM has delivered 19,000 SVC nodes that are running in more than 6,000 SVC systems and, as a result, has built up a huge amount of experience with most configurations and environments. IBM's approach reflects the reality of today's data center: the need to support multiple different hardware and software platforms is a consolidated and simplified way.

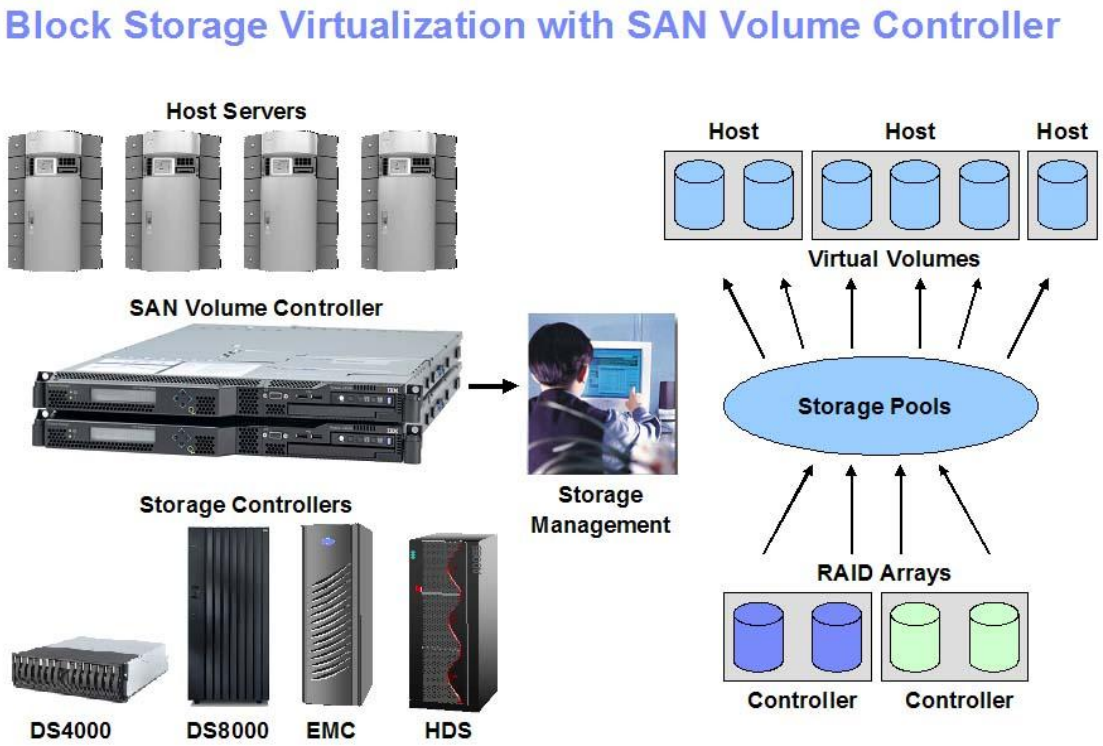
The ability to work across multiple server virtualization hypervisors (VMware, Hyper-v, XEN, PowerVM, zLPAR) means that data center managers can fully support multiple virtualization environments (spanning Oracle Linux [XEN], Microsoft [Hyper-v], and VMware) all at the same time and from the same virtualized storage platform.

The ability to work across multiple storage controllers—including those from EMC, Hitachi, HP, Pillar, IBM, Sun, Fujitsu, and others—enables incredible operational flexibility. Universal connectivity enables the introduction and retirement of storage systems in a flexible and non-disruptive way. It also allows for flexible storage tiering so when the workload increases on an application and the relevant disk component is unable to sustain the performance required, migration to a higher performance platform or tier can be managed transparently.

A less important, but nevertheless relevant, issue is the ability to introduce competition into the storage buying process because the issue of compatibility with what you already have is gone. SVC customers have the advantage of a single storage driver for a single device type across a heterogeneous storage estate.

Remote and local site mirroring between storage controllers of different types (e.g., CLARiiON to DMX, IBM to HP) means that cost effective disaster recovery platforms can be constructed from the best fit platforms.

Figure 1. IBM SVC Connectivity Diagram



Source: IBM Corp, 2009.

Solid State Support in the Virtualization Layer

Solid state disks (SSD) offer the potential of extreme IO performance with low latency and low seek times, providing service for the most demanding applications. Conventionally, SSDs are either located inside the host computer (which undoubtedly gives the lowest possible latency and IOPS potential) or in a storage controller. The host computer location option makes high availability clustering and cloud operating systems difficult to support. The storage controller option is ideal for high availability designs, but maximum IOPS are generally limited by the performance of the central controller and shared cache designs.

SAN Volume Controller offers an innovative, alternative solution with SSD support at the virtualization layer. A fully configured SVC system can support up to 32 SSD modules spread across eight SVC engines configured to provide 800,000 read IOPS within a 2.4 TB volume. By sharing the load between high performance computers (Quad Core Intel 2.4GHz with 24 GB of RAM) connected by four 8Gbps FC ports, SVC leverages an excellent scale-out design to deliver great performance.

Although these extreme levels of performance can only be achieved in a fully configured SVC cluster, much smaller high performance designs can be assembled, starting at two SSD modules spread across a pair of SVC engines in a cluster.

Why is This Important?

Organizations that fail to adopt virtualization and the continuous migration architecture will find the costs of running their IT organizations growing faster than IT budgets will allow. The opportunity to make changes and react to business demands will diminish as the cost of maintaining the legacy becomes all consuming and the cost of change becomes unaffordable.

Businesses facing an unresponsive IT organization have typically responded in predictable ways, such as:

- Do nothing and let more responsive businesses win
- Outsource IT
- Make a once every five year major IT investment to clean up the mess
- Get a new CIO/CTO

Enterprises should not "pay and pray" that things will get better. Those organizations that do so will find costs increasing rapidly whilst business agility disappears. In order to attain any CMA, virtualization must play a part. An investment made in virtualization will pay back in four key areas:

1. Reduction of operational cost – reduced headcount
2. Reduction of change risk
3. Reduction of over-provisioning and over-spending risks – reduced capital and operational costs
4. Improved change cycle time, improved business agility

The Bottom Line

By adding storage virtualization to a toolset that already contains server and network virtualization, IT organizations have the opportunity to ready themselves for a continuous migration architecture where all applications are isolated from their underlying physical systems. The continuous migration architecture is the embodiment of the promise offered by the private cloud: flexible, autonomic computing that simultaneously improves business flexibility while reducing the cost burden of IT.