



High-Performance Storage Virtualization Architecture

A StoreAge White Paper

The Virtualization Architects

www.solutioncentre.co.uk

The Best Storage Virtualization Architecture

A StoreAge White Paper

Copyright © StoreAge Networking Technologies Ltd. 2001.

All rights reserved.

No part of this publication may be reproduced in any material form (including photocopying, or storing it in any medium by electronic means whether or not transiently or incidentally to some other use of this publication) without the prior written permission of the copyright owner, or under the terms of a license issued by the copyright owner.

The information contained in this document is subject to change without notice. StoreAge is neither responsible for nor liable to anyone in connection with this document.

StoreAge, SVM and MultiView are trademarks of StoreAge Networking Technologies Ltd. All other marks mentioned in this material are the property of their respective owners.

Contact Information:

StoreAge Networking Technologies Ltd.
Gutwirth Science Park, Technion City
Haifa 32000, ISRAEL
Tel: +972-4-832-9030 Fax: +972-4-832-9040
<http://www.Store-Age.com>

Solution Centre Limited
Vickers House
Priestley Road
Basingstoke, RG24 9NP.
Tel: 01256 818600
Fax: 01256 819600
E-Mail: sales@solutioncentre.co.uk
www.solutioncentre.co.uk

<http://www.store-age.com>

Abstract

The fast growth in storage capacity and processing power in enterprise installations coupled with the need for high availability, requires Storage Area Network (SAN) architecture to provide seamless addition of storage and performance elements without downtime. Virtualization provides consolidation and uniform management of storage resources for SANs with multiple servers and multiple storage sub-systems. The industry goal is for storage virtualization that is inexpensive, does not affect performance, and that is flexible enough to be applied to various types of systems and networks without requiring significant investments in time or resources. Currently, storage virtualization is implemented at three architectural levels: (a) at the storage sub-system, (b) in the host and (c) in the storage network. Each of the methods provides specific advantages but is limited in its capabilities. This paper presents the available storage methods and a virtualization method that provides the storage management benefits offered by existing techniques without the corresponding drawbacks.

The Need for Storage Virtualization

The term "Storage Virtualization" is used to indicate the operation of dividing the available storage space into "Virtual Volumes" without regard to the physical layout or topology of the actual storage elements (e.g. Disk Drives, RAID sub-systems etc.). Typically, virtual volumes are presented to Operating Systems as an abstraction of physical disks and are used by these OS as if they were such disk drives.

Storage Area Networks impose the need of managing large amounts of storage in a uniform way and from a central location. The fast growth in storage capacity and processing power in many enterprise installations, coupled with the need for high availability and 7x24 operation, requires from the SAN architecture to enable seamless addition of storage and performance elements without downtime. These goals are best achieved via virtualization of the storage. The main difference between a virtual volume and what is generally termed Logical Unit (LUN) is that a virtual volume can be created, expanded, deleted, moved, and selectively presented, independent of the storage sub-systems on which it resides. A virtual volume may consist of storage space in different storage sub-systems, with different characteristics (e.g. number and size of LUNs) and can be expanded with any available storage. The importance of this capability could be exemplified in handling peak activities of e-commerce web sites. Using a Virtualization Manager capable of adding/removing servers on the fly as well as instantly moving volumes from server to server is the best solution to manage the large variance between normal and peak times.

The independence of virtual volumes from the physical storage is essential to enable the migration and replacement of obsolete storage subsystems without performing shutdown or causing data access interruptions. In addition, virtualization of storage enables the usage of software applications instead of human administrators to perform volume allocations. Data management is significantly enhanced through Storage Applications such as Snapshot, Remote Mirroring, Virtual Tapes, etc. that require the creation and expansion of virtual volumes for their use. Server-Free Backup can be implemented only

with Virtualization infrastructure. In summary, a SAN cannot be truly useful without storage virtualization.

Storage Virtualization should provide the following functions without affecting inherent performance:

- ✓ Universality - the creation of a unified view of the virtual storage regardless of the physical elements or topology
- ✓ Single Point of Management - as dictated by the enterprise policy server, Operating System and File System independence - to enable the construction of the most cost effective SAN
- ✓ Storage sub-system independence - to enable unbiased selection of best of breed system
- ✓ Heterogeneous SANs support - to enable smooth growth in changing market conditions
- ✓ Flexible allocation of storage space to servers - to effectively meet users' needs
- ✓ Support the overall specific SAN criteria - e.g. High Availability, Scalability, Security etc.

Existing Storage Virtualization Methods

Virtualization at the Host Level

One method of virtualization is via storage management software that runs at the server level. The main advantage of this method is that it enables multiple storage subsystems to work in parallel with multiple servers. A key difficulty with this method is that it assumes a prior partitioning of the entire SAN resources (disks or LUNs) to the various servers. Virtualization is only performed on pre-assigned storage, losing a key advantage of SANs as well as the independence of volumes from servers. Typically, entire LUNs are assigned to specific servers, thus limiting the amount of servers that can use the same storage subsystem. Further, a virtual volume created over the storage space of, say, two LUNs will not be easily moved to another server, especially if there are other volumes created over the same LUNs. Virtualization at the host level typically also require augmenting the management function with some parallel mechanism of zoning and LUN masking and relying on LAN connectivity for synchronization between servers, which may affect the reliability of the entire SAN.

Virtualization at the Storage Sub-system Level

This method was first implemented in mainframe environment in the 90's and is one of the most common storage virtualization solutions in use today. The virtualization functions are achieved by creating virtual volumes over the storage space of the specific storage subsystem. Pooling all SAN storage resources and managing virtual volumes across several storage subsystems requires that this method be augmented by other means which in general will only be practical in homogeneous SANs using a single type of RAID subsystem. Creating virtual volumes at the storage system level provides server independence, but typically will render a uniform management structure only for homogenous environment and even then only with limited flexibility.

Network Based Virtualization

This is the most interesting approach and the most likely to win out due to its neutrality towards both storage and servers. A key requirement from Storage Virtualization is "making disparate storage look and behave like a common storage resource". Network-based virtualization assures vendor neutrality.

Network-based virtualization is being implemented in two major architectures:

- Symmetric approach - separate appliances in the data path of the storage network infrastructure
- Asymmetric approach - separate appliances installed out of the data path of the storage network infrastructure.

Symmetric (In the Data Path) Virtualization

The general architecture of symmetric approach (sometimes referred to as In-the-Data-Path) is depicted in Fig. 1. This approach calls for the appliance (a computing platform and associated memory) to be installed between the storage users (i.e. the servers on the SAN) and the storage resources, at least in terms of data flow. The key drawback of the symmetric approach to virtualization is that it creates a SAN bottleneck and thus limits the SAN performance and scalability and significantly complicates the design of large-scale highly available configurations.

The symmetric virtualization concept requires that all data from all application servers pass through one single computer. To avoid serious performance difficulties, this computing platform needs to be capable of sustaining the throughput of the entire SAN, typically resulting in expensive hardware configuration. Even with high performance hardware, scalability remains an issue as the symmetric appliance has a fixed maximum bandwidth.

As an example, in a SAN with 50 application servers, the throughput required for each can reach 100 MB/sec. This value * 50 servers = 5 Gbytes/sec. While it is fairly easy to install multiple RAID subsystems with aggregated performance exceeding this requirement, a typical symmetric appliance based on a standard Pentium III processor with a 64/66 PCI bus can deliver no more than 520 MB/sec (that is 260 MB/sec sustained rate as data needs to enter and exit through the PCI bus on the appliance) – a far cry from

the specified requirement. Needless to say, as the SAN becomes larger, this problem becomes more acute.

The hardware cost to construct a symmetric appliance that will support 20 servers with reasonable performance could reach \$25,000 and more; this cost is at least doubled if a High Availability configuration is required.

Some symmetric designs are attempting to circumvent the performance problem through adding a cache facility to the appliance. This approach further loads the appliance's memory subsystem. For many reasons cache is much more effective when it is distributed (i.e. part of the RAID subsystem) rather than centralized (i.e. part of the storage management appliance).

The use of caching also complicates High Availability (HA) configurations, which try to avoid single points of failure, as well as Scalable configurations, where additional appliances are added to respond to increased load. If caching is used with multiple appliances, some kind of cache coherency strategy is required. Experience with RAID controllers shows that this can be very complicated (and expensive).

In order to achieve high performance on the server side, while using two (or more) HBAs, a software driver should be installed on each server. That feature is common to the symmetric and asymmetric solutions. Both need software to be installed on the server in order to create redundant and high performance (multiple paths) solutions for increasing throughput.

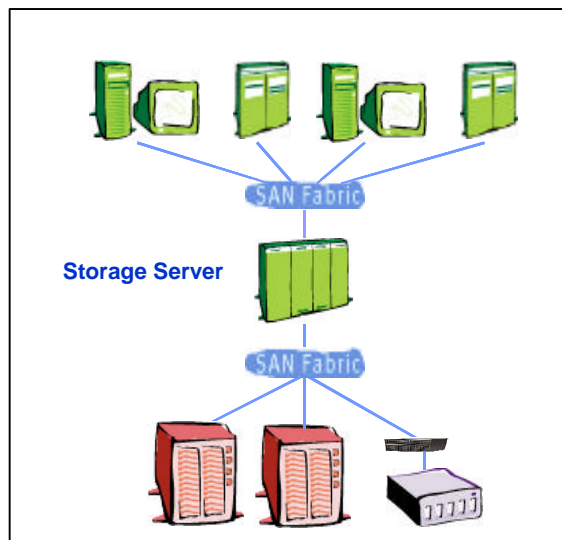


Figure 1

Asymmetric (Out-of-The-Data-Path) Virtualization

This method uses a combination of an Appliance and Agents to create and manage virtual volumes while enabling direct data transfer between server and storage subsystems (see Fig. 2). By having multiple storage subsystems working in parallel with multiple servers, total performance is increased up to the maximum of the FC fabric bandwidth or even more.

In the Asymmetric approach (sometimes referred to as Out-of-The-Data-Path), the handling of metadata is done separately from the data path. The appliance serves as a Metadata Center that “sees” the physical storage and allocates virtual volumes, while an Agent at each of the servers on the SAN is responsible for the actual virtual volume mapping. The Agent retrieves the volume configuration from the appliance and presents virtual volumes to the operating system as if they were disk drives. When the operating system sends an I/O to the virtual volume, the Agent intercepts the I/O, translates the volume’s logical address to the physical address, and sends the I/Os directly to the storage devices.

This architecture maintains the flexibility of symmetric virtualization without incurring the degradation of performance or the high cost of hardware. The appliance can be a small and inexpensive unit, because it does not have to handle actual data transfers.

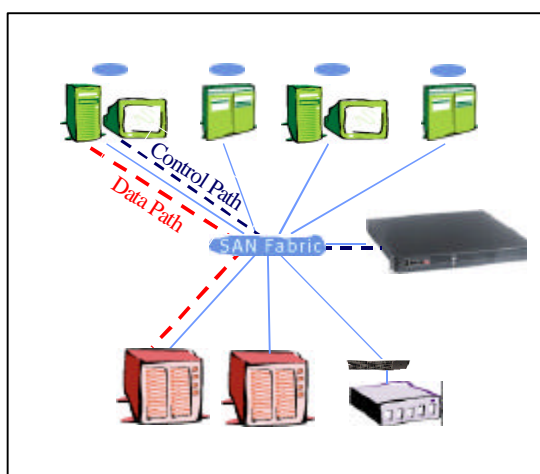


Figure 2

High Availability (HA) SAN configurations can be implemented through simple redundancy of the appliance. Very high degrees of scalability can be achieved through loosely connected storage domains and similar designs, benefiting from the key advantage of the asymmetric concept, i.e. that there are never pending I/Os stored in the appliance nor there is a need for clustering, cache coherency algorithms and other complex mechanisms which limit the practical expansion of real life enterprise SANs.

HA server configurations are achieved (like in the symmetric approach) by installing multiple HBAs in the server and having a S/W Agent who implements Load Balancing and Fail Over functions.

Comparison of Various Virtualization Methods

Method	Host Based	Storage Based	Symmetric Virtualization	Asymmetric Virtualization
Universal	-	-	+	+
Single point of management	-	+	+	+
Server OS & FS independency	-	+	+	+
Storage subsystem independency	+	-	+	+
Flexibility	-	-	+	+
Storage Pooling	-	+	+	+
Performance	+	+	-	+
SAN scalability	Poor	Poor	Poor	High
SAN HA configuration	+	+	Complex	+
Relative management cost	Medium	High	High	Low
S/W to be installed on servers	+	+ *	+ *	+ **
Security	-	+	High security level since Server cannot "see" the storage	StoreAge SVM increases security level by smart algorithms of "Automatic switch zoning".

* Since in most of SANs, servers installed with Multiple HBAs in order to have HA configuration, additional software is needed in order to control this HA implementation.

** StoreAge SVM Agent supports multiple HBA's configuration over all HBAs vendors.

Conclusion

The best way to implement storage virtualization is by the asymmetric architecture using a combination of SAN appliance and agents. This method provides significant scalability, performance, reliability and price advantages over alternative architectures. StoreAge's SVM (Storage Virtualization Manager) is an implementation of this preferred approach.