

Networked Storage Virtualization

Abstract: An ever-growing number of enterprises rely on SAN-based storage to handle their mission-critical data requirements. This, in turn, demands higher levels of flexibility and availability from the storage infrastructure. The answer is networked storage virtualization.

Copyright © 2005 EMC Corporation. All rights reserved.

EMC believes the information in this publication is accurate as of its publication date. The information is subject to change without notice.

THE INFORMATION IN THIS PUBLICATION IS PROVIDED "AS IS." EMC CORPORATION MAKES NO REPRESENTATIONS OR WARRANTIES OF ANY KIND WITH RESPECT TO THE INFORMATION IN THIS PUBLICATION, AND SPECIFICALLY DISCLAIMS IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Use, copying, and distribution of any EMC software described in this publication requires an applicable software license.

Sample Trademark page below.

EMC², EMC, Symmetrix, Celerra, CLARiiON, CLARAlert, DG, E-Infostructure, HighRoad, Navisphere, PowerPath, ResourcePak, SnapView/IP, SRDF, VisualSAN, where information lives, and The EMC Effect are registered trademarks and EMC Automated Networked Storage, EMC ControlCenter, EMC Developers Program, EMC Enterprise Storage, EMC Enterprise Storage Network, EMC OnCourse, EMC Proven, EMC Snap, Access Logix, AutoAdvice, Automated Resource Manager, AutoSwap, AVALONidm, C-Clip, CacheStorm, Celerra Replicator, Centera, CentraStar, CLARevent, Connectrix, CopyCross, CopyPoint, CrosStor, Direct Matrix, Direct Matrix Architecture, EDM, E-Lab, Enginuity, FarPoint, FLARE, GeoSpan, InfoMover, MirrorView, NetWin, OnAlert, OpenScale, Powerlink, PowerVolume, RepliCare, SafeLine, SAN Architect, SAN Copy, SAN Manager, SDMS, SnapSure, SnapView, StorageScope, SupportMate, SymmAPI, SymmEnabler, Symmetrix DMX, TimeFinder, Universal Data Tone, and VisualSRM are trademarks of EMC Corporation. All other trademarks used herein are the property of their respective owners.

All other brand names are trademarks or registered trademarks of their respective owners.

Part Number H1533.1

Table of Contents

- Introduction4**
- Storage Virtualization Technology..... 4
- Keeping Intelligence Closest to What It Controls 4
- Volume Management 5
- Dynamic Volume Migration 5
 - Lease Roll-overs/Technology Refreshes..... 6
 - Performance Moves..... 7
- Point-in-Time Copies 7
- Overall Customer Value 7
 - Flexibility 7
 - Availability 7
 - Optimization 8
- EMC Invista: Implementation9**
- The Platform: Intelligent Switches 9
- Openness and Standard APIs 9
- Physical Architecture 9
- Competitive Approaches10**
- Conclusion.....10**

Introduction

Enterprise computing environments are larger and more complex than ever. Today's IT infrastructures typically include multi-vendor servers, diverse connectivity technologies, and heterogeneous, tiered storage environments. This complexity is not going away; if anything, it may be increasing. Further, as organizations grow, the demand on IT to deliver on the information requirements of the business grows accordingly. Storage infrastructures need to deliver the right information to the right place at the right time, twenty-four hours a day, seven days a week. Storage administrators require the ability to allocate any storage to any application based on changing business needs, and to do so non-disruptively. And they need a uniform, easy way to perform these essential tasks that masks the underlying and increasing complexity of the infrastructure. Today, a new class of storage virtualization technologies are emerging that meet these needs and that ultimately enable organizations to consistently, reliably, and rapidly deliver on the information needs that are driving their business.

Storage Virtualization Technology

By definition, virtualization provides a physical-to-logical storage device abstraction. It presents a simple, consistent representation of a complex infrastructure to entities that consume resources. Virtualization is not a new technology, even in the storage space. Take the example of a storage array. Physically, an array contains a collection of fixed disk drives. However, a connected host does not see a one-to-one mapping of those disks. Instead, it is presented with a set of logical volumes, logical representations which may be the subset or superset of one or more physical disks — whatever the host requires. What is relatively new for storage virtualization technology is locating this capability in the storage network – providing the ability to virtualize storage from multiple connected heterogeneous devices and manage a single logical representation from a single point that encompasses the whole SAN.

Keeping Intelligence Closest to What It Controls

EMC carefully considered which functions could become network-based intelligent services. The proper guiding principal for distributing intelligence in a storage network is this: intelligence belongs closest to what it controls.

Our guiding principle of “keeping intelligence closest to what it controls” prevents us from developing technology for technology's sake. Network-based storage virtualization technology should not be recreating core array functionality in the network. Our goal is to add the appropriate intelligence into the storage network to provide new capabilities that do not already exist elsewhere. It means solving problems without creating new ones, and being able to embrace the existing environment.

EMC believes that storage arrays should always remain in control of data preservation functions. This includes implementing RAID levels, caching, checksums, hardware resilience, proactive error detection, and so forth. These functions are often compute-intensive and require rigorous test and qualification cycles.

Duplicating or recreating these preservation functions at the network layer does not deliver new business value, and could negatively affect data integrity and performance. A storage virtualization design should not, for example, cache data or hold state information. When that is done, the caching or holding entity is responsible for the integrity and reliability of data. That is already done at the array level and would be a duplication of technology and an additional resource burden. EMC believes that the virtualization layer should be as non-intrusive as possible; storing data in the virtualization layer violates this key principle.

Servers should primarily be focused on running business applications. From an infrastructure capability perspective, the server should control application-centric functions such as clustering and application failover. Today, however, there are many storage management operations that are also done on the server such as volume management (a data presentation function) that can potentially move off the host. Going forward, there will still be a need to have path management capabilities from the server to the network and storage for failover and load balancing, but this should be a relatively “thin” layer of software. This would enable servers to be used primarily for their core application-centric functions.

What remains and what could move into the network is data presentation and data access functions. These functions become essential capabilities of a network-based storage virtualization product. Customers who deploy network-based storage virtualization will enjoy additional flexibility to choose where to locate these functions as best suits their needs.

Let's discuss these actual functions in detail.

Volume Management

Network-based volume management is the basis for what many think of as “virtualization.” This is the ability to create and configure virtual volumes from a heterogeneous storage pool and present them to hosts.

Networked storage virtualization solutions can aggregate capacity from many different types of storage arrays. Users dedicate capacity from a given array they wish to virtualize, then import that capacity into a storage pool. Individual volumes which then get created and presented to hosts are subsets of the virtualized capacity from these pools. The volume size and structure of these virtual volumes are completely independent of structural restrictions imposed by the arrays.

Once created, hosts are presented the virtual volume from the network. This virtual volume appears to a host just like a standard SCSI volume from an array. The core function of network-based volume management is the mapping of the virtual volume to the corresponding physical storage.

Volume management thus becomes centralized in the infrastructure and there is no need to load or configure software or configure on individual hosts. Today, an administrator can spend 20 to 30 percent of their time on volume management-related tasks. By centralizing volume management in the network, IT organizations can significantly reduce the amount of time spent on these types of tasks, freeing up resources to focus on other areas that are more critical to the business.

Dynamic Volume Migration

Once the foundation of a virtual volume is in place, additional high-value services can be layered on top to take advantage of the flexibility that is offered. One such service is dynamic volume migration. As shown in figure 1, dynamic volume migration allows storage administrators to move volumes from one location to another while the application remains online.

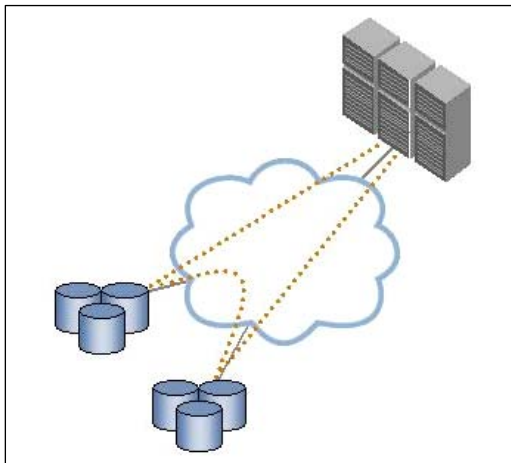


Figure 1: Dynamic Volume Migration

To move a virtual volume, the virtualization software performs a re-direction of I/O from one physical location to another. Despite the fact that the I/O is physically redirected to a new location by the virtualization software, the address of the virtual volume presented to the host never changes. This is accomplished through virtual addressing (Figure 2). This allows the process to be transparent and non-disruptive to the host. Additionally, since the copying and remapping is done by the virtualization system, no host cycles are required, freeing servers to be dedicated to their proper application-centric function.

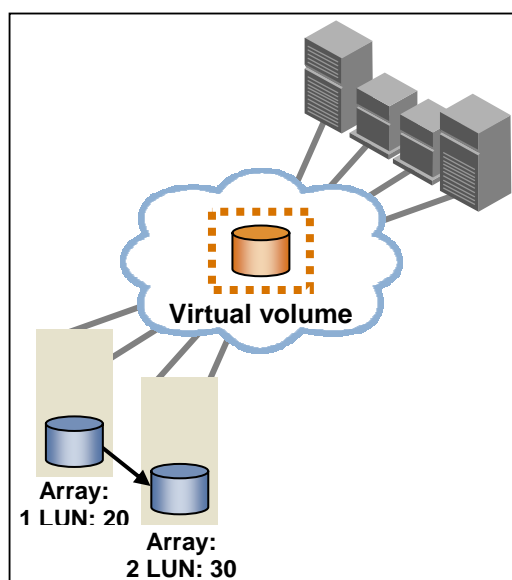


Figure 2: Virtual Addressing

There are a number of very valuable real-world uses for dynamic volume migration:

- Supporting lease roll-overs / technology refreshes
- Performing volume moves or changes in response to changing performance or cost requirements in support of ILM strategies
- Network-based local replication increases flexibility across tiered storage infrastructures

Lease Roll-overs/Technology Refreshes

One significant challenge organizations face today is getting new arrays deployed in their environment. Today's migration solutions are highly complex, manual, and disruptive, making the process extremely labor intensive, complex, and inefficient. Dynamic volume migration capabilities, as part of a network-based storage virtualization system can significantly improve the process.

An organization can bring a new array into the environment, virtualize enough capacity to hold the source data, and then set a mobility job in motion to copy the data from the source to the new array. All the copying occurs in the background, without host impact. Once the copy is complete and consistent, the cut-over occurs and I/O is redirected to the new array.

Once the I/O has been redirected, the old array can be removed from the environment. All of this occurs online, while the application accesses the virtual volumes that have been moved to the new array (target). Furthermore, there is no performance penalty on the hosts as there would be with a host-based volume manager.

Performance Moves

Dynamic volume migration can also help facilitate performance moves. Picture a configuration with medium- and high-performance storage pools. A storage administrator observes that a volume in the medium pool is not meeting a specified SLA. Therefore, to remain in compliance with the service level that has been guaranteed to the application owner, a job is set in motion to migrate this volume into the high-performance pool.

Just as in the lease-rollover example, a copy is performed in the background with no host impact and no application interruption. Once the copy is complete and consistent, I/O is redirected to the new volume in the high-performance pool. In the future, policy-based automation capabilities will transparently and automatically optimize the performance of the virtualized storage infrastructure in accordance with business requirements.

This seamless, heterogeneous information mobility capability, based on dynamic business requirements, is key to support advanced information lifecycle management strategies.

Point-in-Time Copies

The ability to make a copy of a virtual volume is another piece of functionality that can be provided by a network-based storage virtualization solution. Point-in-time copies can be used for back-ups, data warehouse loads, or by test systems to reduce the access load on production storage. The ability to make these point-in-time copies on any storage device that is virtualized adds a great deal of flexibility to the operational scenarios that employ them.

In general, we believe that replication can and should exist in multiple places in an IT environment. Where and how replication is done (in the network using a storage virtualization product, on disk arrays using software such as EMC® TimeFinder®, or on the host, using software such as EMC FullTime™ RepliStor®) depends on a customer's particular use case. Customers will evaluate which solution is most appropriate based on their specific business requirements. EMC is the only vendor offering such a broad range of choices in this area, and point-in-time copy functionality as part of a network-based storage virtualization system adds another option.

Overall Customer Value

Deploying a network-based storage virtualization solution that implements these “proper” access and presentation functions delivers significant benefits in the areas of infrastructure flexibility, availability, and optimization.

Flexibility

Individual storage products today provide varying degrees of flexibility. Data can be moved within an array manually, and in some cases automatically based on policies. Data associated with an application can be moved from one storage array to another array, typically between storage devices from the same vendor. Most of these operations today are either disruptive to the application, or they require manual processes.

Storage virtualization provides the flexibility to leverage the entire multi-vendor storage environment to respond to business demands. While many customers have deployed tiered storage environments, extracting the value from such an environment is achieved when information can be moved seamlessly anywhere in a multi-vendor storage environment, based on application requirements.

Availability

High availability in the usual sense is an attribute of a system that allows it to circumvent unplanned outages. As organizations look to have more and more of a 24x7 operations capability, eliminating unplanned downtime is not enough; planned downtime must also be addressed. Planned downtime may result from operational requirements such as data center moves, upgrades, or platform migrations across the infrastructure, regardless of storage vendor. Storage virtualization can help address outages due to planned downtime, enabling a true non-disruptive operations capability. One of the key requirements to enable non-disruptive operations is the mobility of data across a heterogeneous storage infrastructure.

A recent EMC study found that end users spend over \$130,000 per year in staffing, professional services, and downtime costs to manage tasks associated with data migrations: mapping, planning and scheduling, provisioning, data movement, application reconfiguration, and retiring the source. Scheduling (i.e., setting planned downtime windows) is the number one problem area mentioned to manage and minimize the downtime required. Virtualizing the environment allows customers to move more towards a mindset and goal of continuous data availability.

Optimization

Storage virtualization can be a key enabler for service level driven automation to move applications to the proper tier of storage based on policies. We discussed such an example in the section on dynamic volume mobility. This seamless mobility capability based on changing performance requirements is a key element in supporting advanced information lifecycle management strategies.

Some individual storage arrays, such as EMC Symmetrix[®], have the capability to automate performance tuning through intelligent load balancing and optimized data placement within the array. Organizations want the ability to optimize the entire storage infrastructure across multi-vendor arrays in a similar way. Virtualization would allow them with the capability to better leverage resources by optimizing data placement in the infrastructure overall enabling them to balance service-level requirement with cost.

Storage virtualization includes all of these capabilities, while having no negative impact upon availability. It delivers a centrally managed environment with the ability to optimize resources through the automated movement of applications, maintain proper levels of performance based on application requirements, and deliver the most effective allocation of storage resources across an organization based on business policy.

EMC Invista: Implementation

EMC's network-based storage virtualization product is called EMC Invista™. Key to its ability to deliver the benefits of virtualization is the product design and unique architecture.

The Platform: Intelligent Switches

EMC has chosen to leverage intelligent SAN switches as the platform for deployment of network-based storage virtualization. Intelligent switches have specialized port-level processors (ASICs) that can perform the basic operations of storage virtualization (I/O redirection) at wire-speed. The intelligent switch is thus a high-performance platform that can be easily incorporated into the SAN infrastructures that exist today. EMC is working with all of our current switch partners, Brocade, Cisco, and McData so that customers can deploy EMC Invista using intelligent switches from the vendor of their choice.

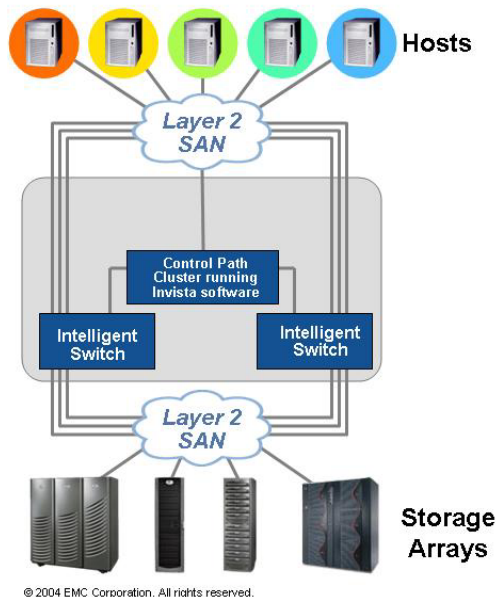
Openness and Standard APIs

It is critical to implement on an intelligent switch platform in an open, non-proprietary fashion. EMC made the key architectural decision to write to vendor APIs rather than develop custom code for each supported switch platform. One of the key reasons this makes sense is the advantage of being able to grow with the switch vendors as their hardware matures, which protects customers against future switch hardware changes. If an application is written in a proprietary way to a specific hardware implementation, when that hardware becomes obsolete, the application will have to be rewritten to work with the next-generation of technology. This will have the effect of requiring a “rip and replace” of the existing deployment. If, on the other hand, the application is written to an established API, and that API remains consistent across hardware generations, the upgrade to future generations of technology should be seamless. Writing to an API also allows software developers to focus on applications, not core platform functionality, resulting in a much more rapid cycle of product improvements and enhancements.

Today, each vendor has its own API, but EMC is collaborating closely with our switch partners and other ISVs to develop a standard API that will work across all intelligent switch platforms. This emerging standard, FAIS (the Fabric Application Interface Standard), is being created in the INCITS T11.5 committee, a subgroup of the committee that develops the Fibre Channel standards. Our initial work with individual vendor APIs will position us to rapidly support the FAIS standard when it is ratified and implemented by the switch vendors.

Physical Architecture

EMC Invista



© 2004 EMC Corporation. All rights reserved.

Deploying EMC Invista involves both hardware and software. The EMC Invista software component runs on an out-of-band appliance called the Control Path Cluster (CPC). The CPC configures and interacts with intelligent switch(es) of the customer's choosing. The intelligent switches may be new switches, or depending on the switch and vendor, customers may be able to upgrade their existing switches.

Figure 4 presents a simplified look at the components of EMC Invista and how they could be deployed in a SAN.

Figure 4: Deployment of EMC Invista

Hosts can connect directly to the intelligent switches or connect through a standard Layer 2 (non-intelligent) front-end fabric for fan-in (as depicted). Arrays can be connected to the intelligent switches directly (as depicted) or via a back-end fabric. This implementation is designed to require minimal changes to most standard SAN configurations, making EMC Invista easy to deploy.

Competitive Approaches

Up until recently, most network-based storage virtualization has been based on inserting a virtualization device in the network data path (or “in-band”) between the hosts and the arrays. Some vendors implement network virtualization by inserting their intelligent array “front-end” into the network while opening up the “back-end” to third-party arrays. Other vendors implement network virtualization with a dedicated server running virtualization software installed on a standard operating system, typically Linux, and still others implement virtualization using a custom appliance.

One disadvantage of in-band virtualization is the addition of an extra hop and ensuing latency between the hosts and the physical storage. Some in-band appliances attempt to improve performance by employing caching within the appliance itself. However, as with storage systems, the performance benefits of caching depend largely on the application and on the efficiency of the virtualization engine’s cache. Also, caching within the network carries with it additional complications and risks to data integrity. In addition, for high availability environments that require the presence of dual redundant in-band virtualization appliances, preserving cache coherency between the dual appliances requires cache mirroring, which adds back some latencies.

Another disadvantage of in-band virtualization architectures is a limit to scalability. Since all I/O within the virtualization domain needs to go through the virtualization device, it becomes the bottleneck. One cannot increase performance by simply adding additional devices. Each device takes on responsibility for its own separate virtualization domain, and multiple devices result in needing to manage separate independent islands of virtualization.

But perhaps the biggest disadvantage of the in-band solutions is that they prevent the continued use of existing functionality (such as replication software) on the storage arrays they virtualize. This restriction forces users to replace their existing deployments with functionality provided by the new virtualization product. This renders customers existing investments in software, customization, and processes valueless. Adopters of these solutions must also ensure that the “new” functionality they will be using has to equivalent feature/functionality what they are abandoning or that they can live with reduced functionality.

EMC Invista’s out-of-band approach avoids these pitfalls. The implementation in the switch utilizes an open and standard approach. No “state” or version of the data is ever held in the network. Until the data is properly stored on the array, the application is not made to believe that the job has been completed. Integration at the switch ensures that it scales along with the SAN infrastructure. In addition, this approach to storage virtualization protects investments in value-added functionality already on the storage array. Only EMC Invista allows users to continue using their existing array-based software, preserving their existing investments.

Conclusion

Network-based storage virtualization promises to provide an unprecedented level of control over the infrastructure. It will help meet the growing need for continuous data availability, simplify and optimize the management of networked storage, and enhance the overall flexibility required to operate within today’s highly complex environments and meet changing business requirements. Implemented the right way, as is the case with EMC Invista, it is a key enabler for advanced information lifecycle management strategies in the enterprise.