VERITAS™

VERITAS WHITE PAPER

Storage Virtualization

VERITAS

Table of Contents

What Is Storage Virtualization?

The term *storage virtualization* has become a standard SAN industry term to mean the dramatic simplification of complex and heterogeneous storage configurations by using logical representations of physical resources. Its roots are in concepts such as virtual memory for computing, in which a mixture of computer memory and nonvolatile storage is used to provide a uniform virtual memory space larger than the actual physical computer memory available. Similarly, technologies like host-based logical volume managers have for years provided a logical storage mapping layer between the server application and large amounts of direct-attached storage, allowing for nondisruptive online disk management as administrators add and tune capacity.

For SANs, the dream of storage virtualization is that an application or end user on any server can use, request and change available storage on the basis solely of its required attributes and without any regard for vendor, location, physical organization or media type. Ideally, this should work in a completely heterogeneous environment, which includes multiple vendors and platforms for servers, software, network elements and storage devices. True storage virtualization requires new technologies and architectures in storage management.

Why Storage Virtualization?

The storage-virtualization concept is gaining mindshare primarily because the costs of deploying and managing storage continue to grow faster than the actual cost of storage procurement. This is happening at a time when corporations are streamlining operations using highly centralized intranet and Internet architectures, requiring an accelerating amount of storage capacity and needing to have these digital assets available continuously. At the same time, the price and scarcity of IT talent put an increasing burden on the amount of storage each IT specialist is responsible for maintaining.

When combined with SAN architectures and server/storage consolidation, storage virtualization may be a panacea to gain control of these exponentially growing operating costs for managing storage. True storage virtualization does require new technologies but also new ways of organizing IT staff. Increasingly, data-center managers are considering the job of managing storage as a task for both storage administrators and storage consumers.

A storage administrator is someone or some group with control over a large amount of centralized storage resources. These resources may include storage subsystems, the storage network and likely storage-management services such as backup or capacity management. A storage consumer is a group of applications, end users or even entire organizations that require storage with specific attributes (sometimes called Quality of Storage Service). These attributes include qualities such as application performance guarantees in terms of latency or throughput; "time-to-capacity," or how quickly new storage with specific attributes can be brought online or "provisioned," the "time-to-recovery" required for applications to regain access to data after a failure and data protection guarantees to provide assurances on how survivable the storage consumers' data is to failures at the server, data center or even site level.

This storage consumer may want a limited degree of administrative control over its storage resources, managed as virtual storage accounts, but is ultimately dependent on the storage administrator to deliver the Quality of Storage Service[1] required. And of course there is increasing accountability between these two organizations, requiring implementation of security and charge-back reporting to demonstrate a lower total cost of storage. This is especially true with the new model of outsourced storage service providers or where managed service providers are hired to manage on-site storage assets.

Storage virtualization is essentially a set of tools that storage administrators can use as a platform for building quality of storage services required by storage consumers.

Where Should Storage Virtualization Be Implemented?

With many vendors promoting and building storage-virtualization technologies, many different places have emerged to embed storage-virtualization functionality in different parts of the SAN architecture: in the host server, in the storage subsystem, in an "out-of-band" appliance or server or in an "in-band appliance" or server. In some cases, this functionality can be embedded into a host bus adapter or a SAN networking device, but the architectural concepts remain the same.

In the Host Server

Overview

Host-based virtualization is most often associated with logical volume managers, which have been commercially available on mainframe and UNIX server platforms for years and are now available on Windows platforms. Logical volume managers provide a virtualization layer that maps physical storage associated with device LUNs into logical diskgroups and logical volumes. The host server applications or file systems then mount the logical volume with no regard to the physical storage location or vendor type. Some volume managers also provide software RAID, allowing logical host volumes to span multiple storage subsystems with mirroring, striping and RAID5 and provide dynamic reconfiguration capabilities, permitting the administrator to add or delete physical storage without disrupting the online availability of the application.

Host-based storage virtualization can be extended beyond a single host view by providing SAN-enabling infrastructure that allows hosts to be aware of resources outside their immediate control—e.g., a pool of spare storage capacity. Alternatively, shared disk clustering technologies such as clustered volume managers can be used to enable multiple servers to have simultaneous read/write access to the same set of logical volumes. This approach usually requires a parallel database, such as OPS, or a clustered file system to coordinate lock management and to ensure data access coherency. This approach, which can also be described as symmetric storage virtualization, is generally available in homogeneous server environments because of incompatibilities in on-disk formats between server platforms.

Application Service Is "Bound" to a Logical Diskgroup

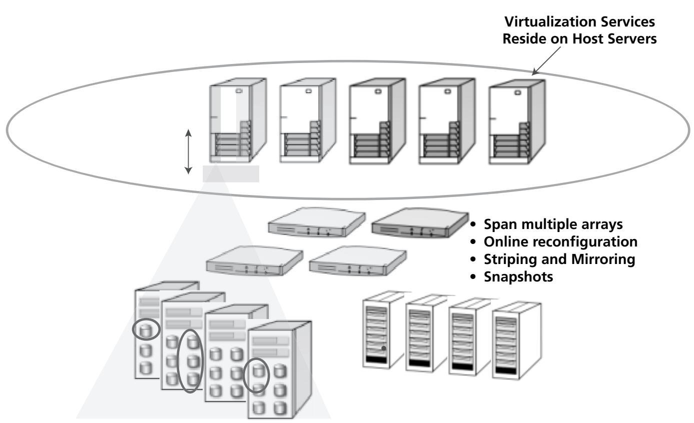

Figure 1

Benefits

- Storage virtualization is based on stable and well-tested technology, such as commercially available volume managers.
- Close coupling to host server file systems and databases allows capacity management of an application's storage without requiring the application to be brought offline.
- Storage virtualization can be extended to two or more servers using shared disk clustering technologies, such as clustered volume managers.

Limitations:

- Scope of control is "server-centric" and limited to logical "islands" of storage mapped to a specific server, unless extended with shared disk clustering.
- Storage administrator may have limited administrative authority over host servers—e.g., the administrator may not have root privileges.
- Although this type of architecture can be extended with shared disk clustering solutions, doing so is sometimes perceived as complex to set up and administer.

In the Storage Subsystem

Overview

This approach is most often associated with large-scale RAID subsystems that contain large amounts (many terabytes) of storage, many I/O channels to attach to SANs or directly to many hosts and intelligent controllers that provide services such as LUN access control on a host-specific basis, caching and advanced storage management capabilities like data replication.

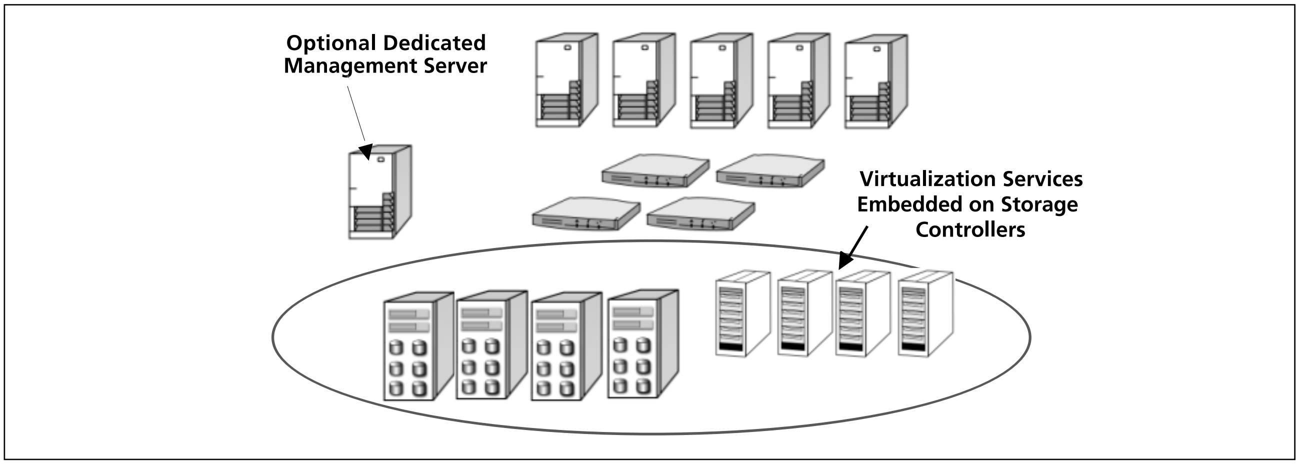

Figure 2

Benefits

- Storage administrator has all control at the device, which is sometimes seen as simpler and more secure because it is close to the physical data.
- By isolating the solution from the host server, this architecture facilitates heterogeneity across many server operating systems.
- Solution can be highly tuned to the hardware capabilities and organization of the physical storage subsystem.

Limitations

- Solution is proprietary to each subsystem vendor and often requires highly specialized expertise to maintain.
- Solution cannot effectively manage storage virtualization across multiple arrays without additional virtualization services on the host or a special appliance. Essentially each array must be discretely managed with vendor-specific tools, and the stored data is generally hard partitioned to a specific host server and operating system.
- Each subsystem has high initial costs and high total price per megabyte.

A Special-Purpose “In-Band Appliance”

Overview

This architecture is a new approach to storage virtualization in which special-purposed storage management devices are deployed off the host server but reside in the data flow of the SAN. In-band appliances can be implemented in specialized devices with a vendor-specific operating environment or a standard-server CPU platform that is transformed into easy-to-manage storage-access devices using enabling appliance software. This appliance software provides a logical volume management layer that enables management of direct or SAN-attached devices, combined with a data-access layer. The data-access layer can present either the logical storage as a file system for a NAS solution or a virtual disk in which logical volumes appear as device LUNs to an attached storage consumer. This type of appliance can be thought of as an aggregator of storage capacity and services because it effectively provides a virtual portal into the SAN. Often, these types of appliances come bundled with sophisticated management software to simplify the setup and maintenance of attached storage. In-band appliances can be packaged with storage as an alternative to vendor-specific RAID controllers, in the fabric devices for an integrated SAN solution or as a stand-alone device to provide data access across many devices in a SAN fabric.

Because all data to and from a storage consumer must flow "through" this type of appliance, this solution is typically equipped with a high availability capability, allowing another appliance to take over data access transparent to any clients on the network in case of an appliance failure. Also, whereas concern sometimes exists about potential added latency created by its location on the data path, cache and intelligent caching mechanisms can be added to enhance throughput and system response in all but the most write-intensive applications. Even in this case, in practice most applications and databases issue multiple I/Os "threads" at once. Because the CPU on the application server processes previously acknowledged I/Os threads rather than waiting for I/Os to complete sequentially, overall I/O performance in most cases can be enhanced with an in-band appliance equipped with a cache.

One other promise of in-band storage appliance software is scalability. By using advanced technologies such as clustered volume managers available today on commercial platforms, data access can be spread across many in-band appliances, overcoming a limitation inherent in vendor-specific in-band implementations.

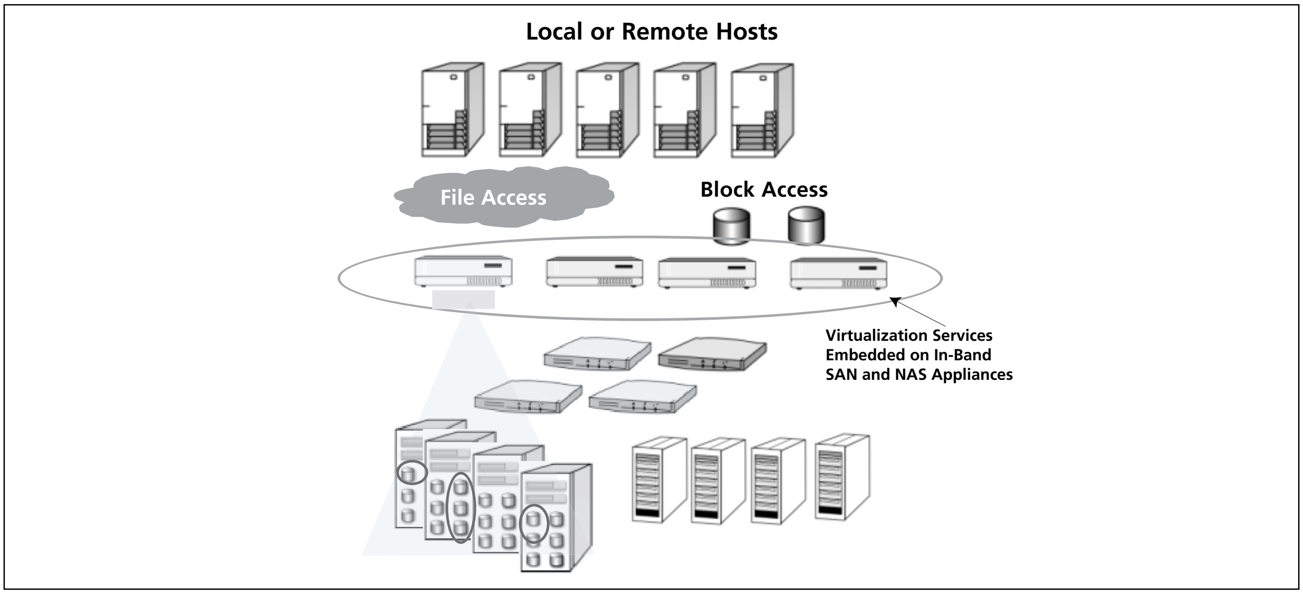

Figure 3

Benefits

- The storage consumer can be easily provided data access with either file or block-oriented protocols, regardless of the types of physical storage subsystems used.
- Security: The storage administrator has complete control of which resources are presented to the storage consumer, preventing accidental or malicious damage from any "untrusted" clients using standard access-control techniques.
- The storage consumer can be implemented on a variety of server CPU platforms for a range of price/performance.

Limitations

- Each appliance has a single point of control within the SAN, much like the host-centric model. Extending storage virtualization across multiple appliances requires advanced clustering technologies, which are sometimes complex to administer.
- High-availability clustering techniques across two or more appliances are necessary to protect data access against failures.
- An additional element in the data path may cause performance problems, requiring caching techniques in the appliance.

A Special-Purpose “Out-of-Band Appliance”

Overview

This architecture is another very new approach to storage virtualization in which standard server CPU platforms are transformed into special-purposed appliances using appliance software. Unlike the in-band appliance, however, this solution separates the flow of control and data. The out-of-band appliance is responsible purely for administrative and control tasks such as changing the storage configuration or access control or creating snapshots, while data I/O flows directly between host servers and the storage subsystems. This appliance can also be viewed as a virtualization server or an asymmetrical storage-virtualization architecture.

The virtualization server usually requires a logical volume management layer that manages the SAN attached devices, combined with a control layer that communicates with a virtualization client on the host servers or embedded “off-host” in an HBA or fabric device. The virtualization client gathers the physical storage information from the appliance during setup or reconfiguration. And all storage configuration requests are passed from the virtualization client to the virtualization server, providing a simplified and highly centralized administrative model for these tasks.

Virtualization clients are located “off-host” to isolate all storage management from the host operating system. This is sometimes attractive to manage security concerns better or because the storage administrator doesn’t have root access to the host server. A virtualization client is located in the host operating system as a lightweight process to minimize cost of ownership by maintaining hardware independence. Additionally, this method avoids additional latency due to an extra device in the data path or ensures tight coupling to application-level services such as application failover using high-availability clustering or application-performance management.

Benefits

- Storage administrator has completely centralized control of the virtualization services at the appliance, which avoids using host server CPU cycles for administrative tasks.
- While control information flows between the appliance and the server, all data I/O is performed directly through the SAN fabric, allowing for maximum throughput.
- This solution is very portable to heterogeneous server environments because the required host code for the virtualization client is relatively lightweight compared to a fully functional host-based volume manager.

Limitations

- A special driver or HBA on the host server is necessary to implement the virtualization client, which may not be acceptable if the storage administrator doesn’t have appropriate privileges.
- Operations such as relayout or snapshots may require extensive communication between virtualization clients and servers, resulting in performance problems.
- This very new approach needs to be tested for robustness with server applications and storage subsystem.
- With existing technologies, the out-of-band appliance must be connected via FC into every fabric topology to have appropriate level of administrative control. This leads to many appliances in an overall SAN environment, increasing management complexity.

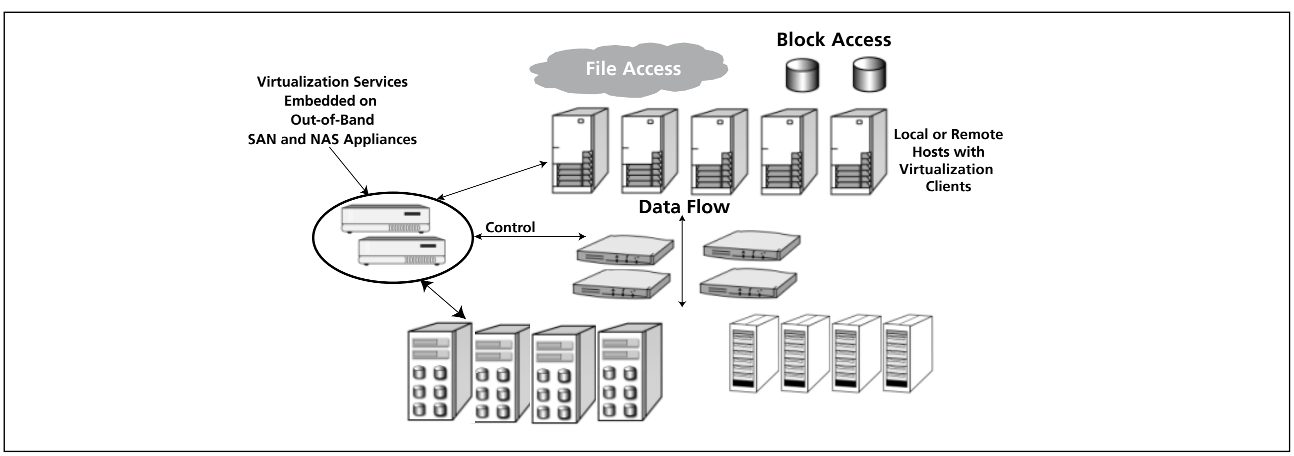

Figure 4

VERITAS Storage-Virtualization Strategy

The VERITAS storage-virtualization strategy is based on a few guiding principles with the goal of simplifying open, heterogeneous SANs while ensuring robust operation. These principles are:

1. Enable the SAN administrator to deploy storage-virtualization services at one or many places in the storage network, depending on what's best suited to the administrator's needs in terms of management and performance.
2. Embrace open interfaces and interoperability with third-party virtualization services to provide customer choice of vendor at each part of the value chain.
3. Implement storage virtualization in an evolutionary and modular manner by building on well-proven volume management technology.
4. Integrate management of applications to the storage-virtualization infrastructure to deliver better on Quality of Storage Service.

The key technologies of the VERITAS storage-virtualization strategy are the VERITAS Volume Manager™, widely regarded as the best-in-class volume manager available today, and the V^{3}™ SAN Architecture, a family of storage-virtualization technologies first announced in 1999. Several elements of the V^{3} SAN Initiative are now available as commercial end-user products or embedded in other products from VERITAS and third parties.

VERITAS Volume Manager™

The VERITAS Volume Manager is the cornerstone of the VERITAS storage-virtualization strategy. Available on both UNIX and Windows platforms[2], Volume Manager has for years provided an industry-leading and robust host-centric storage-virtualization layer. Volume Manager is tested with all major arrays and treats every device's LUNs as logical disks, regardless of vendor or drive technology. Some key fundamental storage-virtualization capabilities of Volume Manager are:

- It can build a diskgroup spanning device LUNs contained in multiple storage subsystems.
- Dynamic diskgroup import/export allows simplified, online configuration management of large amounts of storage.
- Online volume reconfiguration, such as volume resizing and dynamic storage relayout and optimization, allows the administrator to modify a server's storage configuration without requiring server application downtime.
- Software RAID for striping and mirroring allows the storage configuration to be optimized.
- Dynamic multipathing automates failover and load balancing across multiple paths from the host to a device.
- Volume Manager builds static point-in-time snapshots of stored data for use by other servers or applications and then quickly resynchronizes the snapshot back to the real-time state.

A key technology extension of Volume Manager, which is shipped today as a component of data-sharing solutions such as Oracle Parallel Server on Solaris, is the VERITAS Cluster Volume Manager™. Cluster Volume Manager provides symmetrical host-based data sharing by allowing a volume to be simultaneously mounted for use across multiple servers for both reads and writes. Both Volume Manager and Cluster Volume Manager technologies provide a stable and full-featured base to build on to create advanced storage-virtualization architectures.

VERITAS Cluster Server Moves Application Service and Re-Imports Diskgroups for Clusters of Up to 32 Nodes

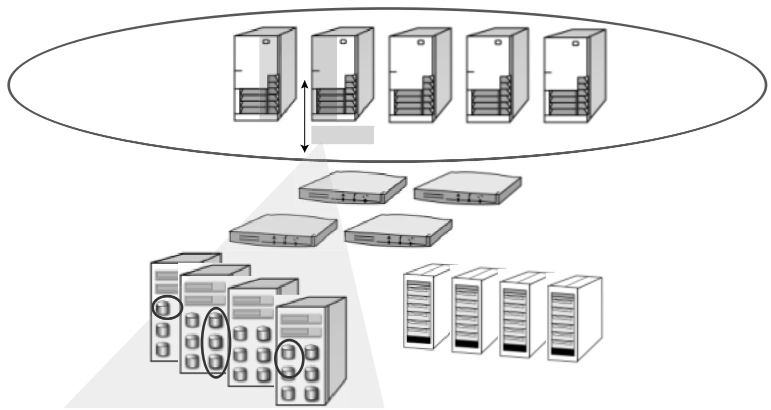

Figure 5

Building on Volume Manager for Time-to-Recovery

A QoSS characteristic that administrators require from SANs is improved time-to-recovery of server application. Improved time-to-recovery addresses the administrator's need to keep applications online in the event face of device, network, server or even application failures without a high degree of manual intervention. A key part of improving time-to-recovery is in enabling the application to reconnect quickly to its storage, and here is where storage virtualization can provide enhanced architectural platform.

An example of this in action today is the interaction of VERITAS Cluster Server™, a high-availability clustering solution, and Volume Manager. Cluster Server manages availability of application-service groups across 2 to 32 servers, consisting of an application and all its critical network resources, software processes and storage resources required to be online. In a SAN environment, Cluster Server requires all servers in a cluster to have connectivity to shared storage to allow any server to reconnect to another server's storage in the event of failure. Storage virtualization with Volume Manger is required for two critical reasons: to manage diskgroup ownership and to accelerate recovery process of the cluster.

Diskgroup ownership is an intrinsic feature of Volume Manager that allows multiple servers in a cluster to be connected to shared storage but only one server to be assigned ownership at any one time. The diskgroup ownership properties are shared among members of the cluster using the Cluster Server cluster communication. As long as a LUN in the SAN is mapped to a Volume Manager diskgroup in the cluster, no other server will mistakenly read or write to that disk space.

When a failure occurs and Cluster Server must move an application service group to another server in the cluster, the application service cannot restart until it reconnects into its storage on the SAN. The diskgroup "de-import/import" feature of Volume Manager is used in this case. Cluster Server first detaches or "de-imports" the diskgroup from the initial server, essentially releasing ownership of the storage. On the new server, Cluster Server then reattaches or "imports" the diskgroup to the new location of the application service group, thus reassigning ownership. The new ownership properties are once again distributed to all the servers on the cluster, ensuring data security.

Without a storage-virtualization technology like Volume Manager, the ownership properties of the shared storage can only be guaranteed using LUN-level operations such as "SCSI reservations." This technique can work for small clusters with small amounts of storage, but it does not scale to server farm running enterprise-scale applications because it requires Cluster Server to send frequent messages to the shared storage and to manage a complex list of LUNs. Similarly, the import process is dramatically simplified with

Volume Manager because VERITAS Cluster Server can just import a short list of relevant diskgroups rather than a long list of LUNS on the SAN. This synergy of Cluster Server and Volume Manager is in allowing real-world deployments of high-availability clusters of 8, 16 and even 32 server nodes over SAN today with time-to-recovery measured in minutes.

Building on Cluster Volume Manager with VERITAS SANPoint Foundation Suite™ HA

VERITAS Cluster Server Plus Cluster File System for "Scaling-Out" File System and Fast Time-to-Recovery

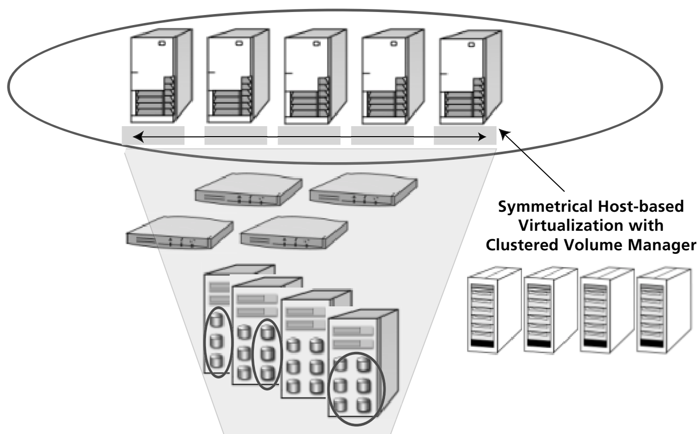

Figure 6

For true mission-critical applications such as e-commerce database servers or financial systems, every second of downtime affects the company's bottom line. This environment requires an even faster time-to-recovery storage-virtualization platform than just Volume Manager can provide. The symmetrical storage-virtualization capabilities of the Cluster Volume Manager provide this enhanced platform. The new VERITAS solution, SANPoint Foundation Suite™ HA, uses this technology to deliver fast recovery over a SAN.

Cluster Volume Manager is an extension of Volume Manager and so includes the same diskgroup-management capabilities mentioned above. In a clustered environment using Cluster Volume Manager, every host server can connect to and simultaneously own the same SAN attached storage using the same logical diskgroup name. When Cluster Server now moves an application service group built on Cluster Volume Manager to a new server in the cluster, Cluster Server does not need to take the step of importing the diskgroup because the new server already has this ownership. Eliminating this step for cluster configurations with large storage configurations can reduce the time-to-recovery to a minute or less.

With the data-sharing capabilities created with storage-virtualization technologies like Cluster Volume Manager, write access to the shared data must be managed to avoid data corruption or coherency problems when multiple servers write to the same data at the same time. This requires essentially a "write-locking" mechanism to ensure that only one server has write access to a piece of data at one time. The locking mechanism is best handled at the application level, where context of the data is best understood. An example of this is in building parallel databases such as Oracle Parallel Server on top of Cluster Volume Manager.

SANPoint Foundation Suite HA builds both Cluster Server and VERITAS Cluster File System™ on top of Cluster Volume Manager to implement this write locking at the file-system level. Cluster File System is an extension of the VERITAS File System™, which uses standard Cluster Server cluster communication protocols to implement a write-locking mechanism. Using Cluster File System, the same File System file system can be mounted simultaneously on all servers on a cluster. This allows a Cluster Server application service group that includes the file system to avoid the process of remounting the file system on the new server because this server already has this access. As an integrated solution including Cluster Server, Cluster File System and Cluster Volume Manager, SANPoint Foundation Suite HA provides fast recovery of file systems or applications using a file system by building on the symmetrical storage-virtualization capabilities of Cluster Volume Manager.

Time-to-Recovery for Disaster Recovery: VERITAS Volume Replicator™ with Volume Manager

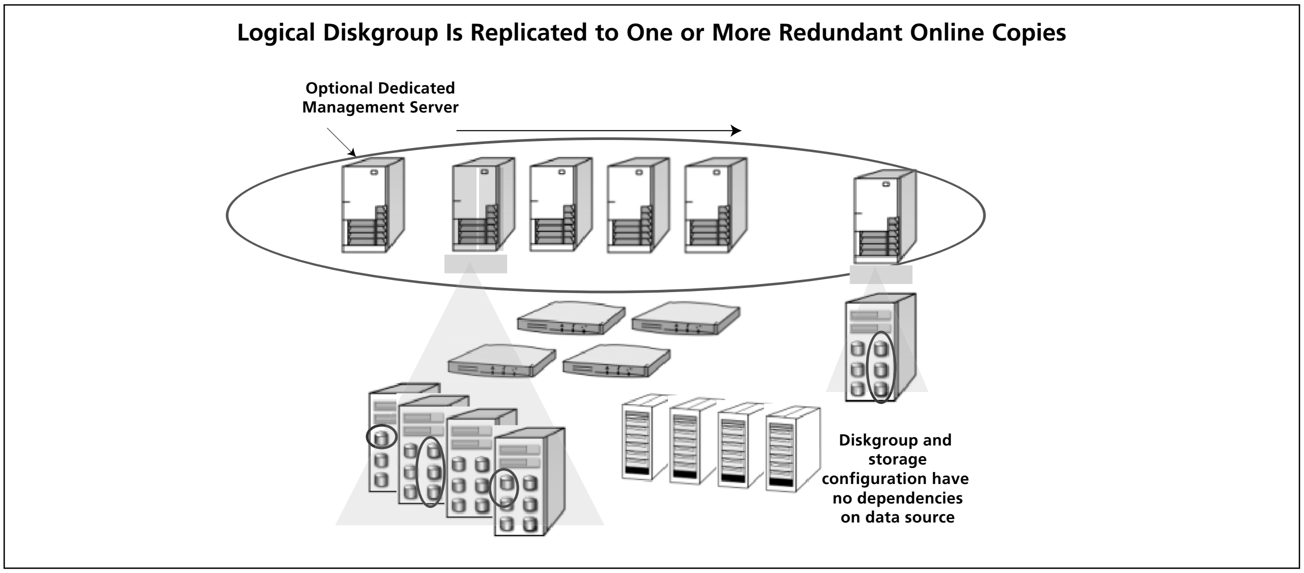

Figure 7

Disaster recovery is often thought of as ensuring that all mission-critical data is stored on tape at a remote site in preparation for a catastrophic failure such as data-center flooding or an earthquake that renders an entire campus nonoperational. For companies that build their business on data availability, the hours or days necessary to restore the server data from tape and bring the applications back fully online are unacceptable. They require a faster time-to-recovery solution that can ensure that applications can be brought online at a remote site in minutes with as little manual intervention as possible. Volume Replicator is a product that builds on the storage-virtualization capabilities of Volume Manager to accelerate time-to-recovery for disaster recovery.

To accelerate time-to-recovery for disaster-recovery plans, a redundant copy of the application's storage must be kept up-to-date at one or more remote sites on online storage. Having the redundant copy of online storage instead of tape eliminates one of the most time-consuming steps in disaster recovery: recovery from tape. Volume Replicator™ addresses these needs by replicating logical diskgroups managed by Volume Manager. A diskgroup at the primary site can be replicated to one or more remote servers over any IP network connecting the sites. After initialization, only changes to the primary sites diskgroup are transmitted to the secondary sites to optimize use of available bandwidth. Note that the diskgroup at the remote sites can be any type of software or hardware RAID configuration because the data replication is done at the logical Volume Manager diskgroup level with no requisite knowledge of the underlying storage or LUN configuration. Although the diskgroups at the secondary servers are not actively mounted in order to prevent data corruption, usually another identical instance of the primary application is running in idle mode waiting to be brought online.

When a catastrophic event occurs, Volume Replicator detaches the replication process at the remote site and mounts the diskgroup on secondary server. The remote application can then import this diskgroup for use in minutes and be brought back online. With its capability to replicate changes to Volume Manager diskgroups over IP networks of very long distances between any heterogeneous storage configurations, Volume Replicator can provide fast time-to-recovery if Volume Manager is used as a storage-virtualization platform for host-based, storage subsystems and in-band or out-of-band appliance solutions.

The V^3 SAN Architecture

The V^3 SAN Architecture, first announced and demonstrated at the Fall 1999 VERICON conference, is a set of technologies for extending Volume Manager beyond traditional host-attached storage and into SAN-wide virtualization. The elements of the V^3 SAN architecture[3] are:

V^3 SAN Infrastructure

Although the Volume Manager was designed to scale to many thousands of storage objects, like all existing host software it does not consider any of the new services and capabilities that SAN architectures bring. For example, there is no awareness that a device LUN may be several miles away from the host server to ensure an intelligent mirroring decision. Similarly, Volume Manager has no provision to understand complex SAN topologies or administer new services such as “Fabric Zoning” or “LUN masking.” To enable SAN capabilities for Volume Manager and other existing server applications, new SAN-enabling infrastructure is required that can bridge between the SAN configuration server applications via simple and consistent management APIs. By ensuring that these APIs are consistent across different devices, operating systems and applications, new SAN-aware server applications can be developed more quickly. This category of SAN-enabling technology is referred to as V^3 SAN Infrastructure and is available as technology components of a SAN application or product.

The first V^3 SAN Infrastructure component is the V^3™ SAN Access Layer™ (SAL), currently available as an embedded component of the VERITAS SAN management product SANPoint Control™. V^3 SAL focuses on the task of providing an intelligent “SAN Discovery Service and Management.” SAN Discovery Services can be defined as the process of detecting all SAN devices and services, correlating the information to ensure that it is correct and is kept up-to-date and building an object-oriented database that can be used by any SAN application. The V^3 SAL uses a variety of FC-based and LAN-based protocols to communicate with SAN devices and services, but provides a simple XML-based API for a SAN application to access the correlated SAN information in its database. For active management, the SAL manages the access-control mechanisms of the SAN. In particular, it provides an abstracted interface for control of fabric zoning and LUN management.

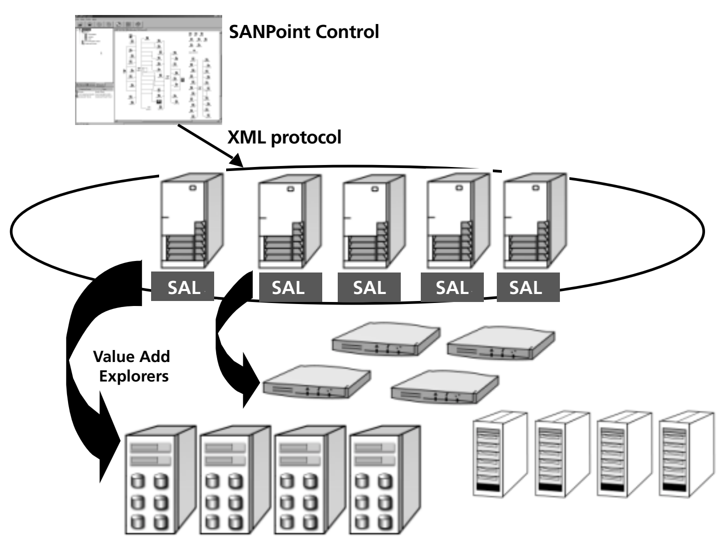

Figure 8

Because the V^3 SAL is resident on host servers in most early SAN installations, it is usually associated with “Host Based Virtualization.” However, it is based on distributed architecture with a V^3 SAL database component and V^3 SAL clients. The V^3 SAL database can reside on any server in the SAN, and all host-based applications have the same level of functionality and visibility. This distributed implementation allows the V^3 SAL to be used in other virtualization architectures—e.g., the V^3 SAL DB could be embedded on an out-of-band appliance with V^3 SAL clients integrated with other virtualization client software on host servers. Or the V^3 SAL DB and Clients could be located on a set of in-band SAN appliances allowing the appliances to become enabled with these advanced management capabilities.

V^3 Management Applications

The last category of the V^3 SAN Architecture is V^3 SAN management applications, available commercially today under the name SANPoint Control. V^3 SAN management applications provide the storage administrators with the tools to manage storage virtualization centrally across server platforms and across all SAN devices, bringing order to the chaotic experience of managing heterogeneous SANs. The V^3 SAN management tools provide services such as visualization of SAN resources in task-specific ways, administration of complex tasks such as zoning or storage allocation, typical management functions such as reporting and event handling and automation of tasks using policy management. Products in this category like SANPoint Control are designed to use virtualization services from VERITAS or third-party suppliers wherever they may be located on the SAN, enabling the administrator to benefit from centralized management regardless of which virtualization architecture the administrator prefers.

The V^3 SAN management applications build on the V^3 SAN Infrastructure to visualize, monitor and control the SAN. SANPoint Control, for example, today uses the information in V^3 SAL to provide discovery and management services. Because the V^3 SAL API also allows for vendor-independent control of SAN devices, it becomes much more straightforward to build tools to facilitate complex tasks such as zoning without having to write vendor-specific or architecture-specific applications. As new V^3 SAN Infrastructure technologies become available, it will be possible to build increasingly sophisticated SAN management applications (*see Improving Time-to-Capacity*).

For maximum visibility and control of the SAN, each SAN-attached server should have a V^3 SAL instantiation. Use of this host-based intelligence allows several key benefits that would be otherwise impossible, including:

- Discovery of key HBA attributes, including: model, firmware and device driver levels, as well as the ability to correlate multiple HBAs in a host as a SINGLE host, versus multiple HBAs.
- Discovery of key host attributes, including OS platform, IP address and OS handles, which provide the critical link associating logical devices to a host and allow SPC to visualize the virtual paths from host to device. Without this capability, the devices can only be associated to an HBA level with no context to the host server.
- Ability to discover LUN information by accessing array-management protocols, which reside on the host (for example, EMC Symmetrix SYMCLI).
- Ability to discover Volume Manager information, including diskgroup mapping to array level LUNs.

This new family of storage-management applications will also become increasingly more integrated with host-based VERITAS products such as Volume Manager, Cluster Server and NetBackup™, as well as other host server applications such as databases. In addition, they will integrate with the V^3 Storage Appliance to provide an integrated SAN-wide and "top-to-bottom" management suite for all day-to-day administrative and monitoring tasks.

V^3 Storage Appliances

The second component for the V^3 SAN Architecture is the V^3 Storage Appliance. This is essentially an in-band storage-virtualization architecture, as previously described, based on the VERITAS Volume Manager that can be used for both NAS or SAN solutions.

For the NAS solution, the data-access layer used is the existing VERITAS File System with its advanced journaling and resiliency features plus an extended SAMBA layer to provide client access in either CIFS or NFS access protocols. For the SAN solution, the data access layer is a new "Virtual Disk" layer (or target mode driver) that provides the mapping layer from a logical volume to a virtual disk that appears as a device LUN to any attached server. In both cases, sophisticated management tools are bundled into the appliance software that make the setup, configuration and maintenance much simpler than for a conventional host-server implementation.

By building on the Volume Manager platform, a host of virtualization services such as dynamic online reconfiguration can be easily integrated into the V^3 Storage Appliances. For ready-to-use high-availability solutions, a pair of redundant appliances can be deployed with Cluster Volume Manager extensions for Volume Manager and embedded Cluster Server for a fast failover solution. In addition, the V^3 Storage Appliance can also now become a platform to deploy advanced services such as volume replication between two or more appliances separated across a wide area network.

Improving "Time-to-Capacity"

One of the keys to Quality of Storage Service (QoSS) that well-managed SANs must provide is better time-to-capacity. As SAN configurations become larger, holding many tens or hundreds of terabytes of various types of storage arrays and with many types of attached applications or users, the SAN administrator must be able to allocate storage to applications efficiently without significant manual work or reliance on many disparate tools. This process, called storage provisioning, requires key steps by the SAN administrator:

- Characterize the needs required by a storage consumer on the SAN, including the type of RAID level, initial capacity required, how quickly new storage must be available when needed and a maximum quota for anticipated capacity needs.
- Identify attributes of all storage on the SAN, and classify this storage in terms of application QoSS parameters.
- Allocate the initial storage capacity to the storage consumer.
- Monitor each storage consumer's allocated capacity to determine if and when configuration changes are required.
- Change the SAN configuration to provision new storage, remove storage or perhaps tune it for optimal performance or improved reliability.

Ideally, a well-designed storage-virtualization architecture should dramatically simplify the steps above, especially in an environment with a large number of storage consumers with different requirements.

With SANPoint Control 1.0 in combination with the VERITAS Volume Manager, the basics of managing time-to-capacity were put in place. For the first time, the SAN administrator was able to flexibly visualize and monitor the overall SAN environment from a single place and initiate fabric zone configuration changes regardless of the vendors used in the SAN. Much of this capability is possible because SANPoint Control is built on top of the V^3 SAN Access Layer. Fabric zoning administration allows specific device LUNs to be made visible to the host server by bringing them into the host's "visible" fabric zone. To complete the operation, the device LUN usually has to then be configured using the array vendor's management tool to be mapped or "bound" to a specific host server. Last, Volume Manager can be launched in context from the SANPoint Control console to grow an existing application or file system volume to bring the new device online.

Although this base level of storage provisioning helps to address some of today's SAN management challenges, much more opportunity exists for simplifying and scaling time-to-capacity with increasing levels of capabilities.

Level 1: SANPoint Control 2.0 for Enhanced Storage Allocation

SANPoint Control 2.0 introduces the concept of storage accounts to allow the SAN administrator to move beyond the artificial administrative limitation of fabric zones and LUN masks. A storage account is a logical set of storage consumers and the storage provisioned to them. The storage account can be a basis for implementing access control to SAN resources, for setting quota management policies and thresholds and for charge-back accounting. This administrative concept allows the SAN administrator to deal better with large numbers of storage consumers and storage capacity by moving management up a level of abstraction. Using a storage account model for administration, the new release of SANPoint Control 2.04 adds four critical time-to-capacity features:

- **Enhanced Storage Provisioning Control**
  In addition to broadening the number of fabric vendors supported for fabric zoning capability, this release adds the capability both to visualize accurately and to control the LUN access control, often referred to as "LUN Masking," for enterprise RAID controllers via the V^3 SAN Access Layer. Building on this capability, SANPoint Control now provides the user a vendor-independent wizard to provision LUNs anywhere on the SAN to a specific host server port without requiring the use of device-specific configuration tools for either the fabrics or RAID controllers. This tool can also be configured to notify the host operating system automatically of a configuration change to signal the need for an operation to mount the new storage. In addition, quick visualization techniques such as automatically highlighting LUNs or their parent enclosures when a mapped server is selected on the topology map now allow the administrator to diagnose mapping configuration problems quickly.

- **Management of Storage Groups and Storage Accounts**
  All available storage on the SAN is classified into user-defined storage groups based on a similar set of attributes, including level of redundancy, physical location, capacity, performance and vendor type. Classification of available capacity into storage groups and storage accounts can be done automatically through the intelligent discovery provided by the V^3 SAN Access Layer or using a

wizard-based manual method. Once organized in storage groups, wizards guide the SAN administrator in creating storage accounts or assist them in identifying storage with targeted attributes by querying the SAL database for appropriate storage groups and provisioning the free LUNs into the server application.

- **Capacity Reporting**
  A facility to generate formatted real-time and historical reports detailing available capacity for each storage group and storage account on the SAN is now included. This provides the SAN administrator with the tools to quickly assess when to add new spare capacity of the appropriate type to the SAN before a crisis develops.

- **Improved Integration with Volume Manager**
  As a further step to improve the time-to-capacity for the server application, the Volume Manager can be automatically notified when new capacity has been added to the server. This alerts the administrator to perform a dynamic volume grow function for a specific diskgroup, seamlessly adding the new capacity into an application's usable storage.

Level 2: SANPoint Control 2.5 with the V³ Intelligent Provisioning Service

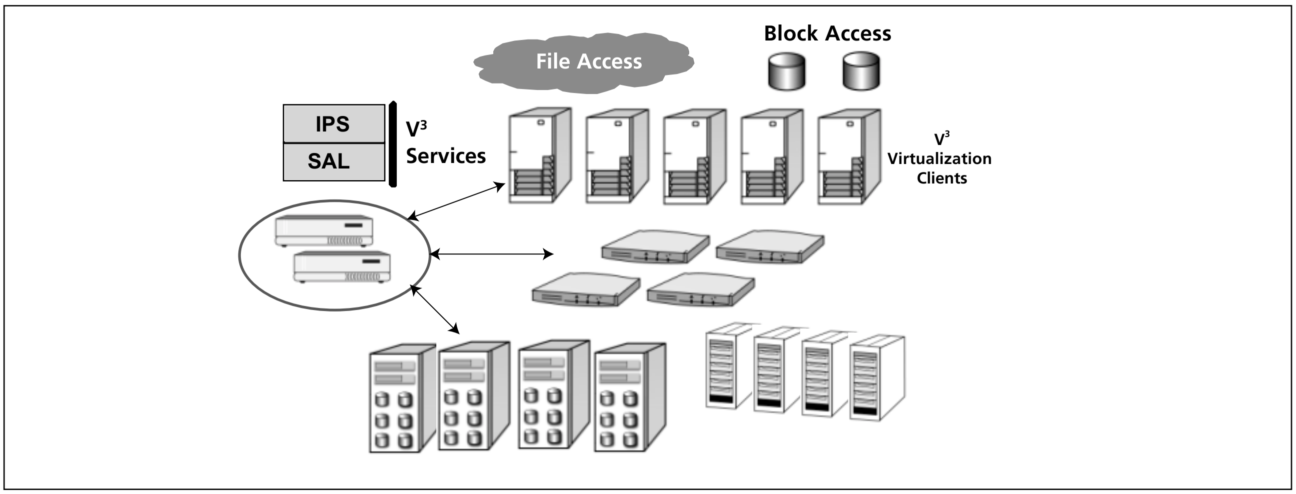

Figure 9

In addition to adding improved heterogeneity and discovery information, the SANPoint Control 2.5[5] adds two key new capabilities and an optional add-in application based on a new V³ SAN Infrastructure element, the V³ Intelligent Provisioning Service (IPS). The combination of these capabilities provides a comprehensive tool to improve management of time-to-capacity significantly with the dynamic provisioning of storage LUNs on a SAN. The two new features planned are:

- **Storage Pool Management**
  SANPoint Control 2.5 will make available LUNs on the SAN identifiable by their allocation state, which determines what provisioning actions can be performed on a specific LUN. With storage now identifiable in terms of both storage groups and allocation state, SAN storage can be managed as virtual storage pools of similar storage states and groups. The SAN administrator can easily query SANPoint Control to quickly visualize or request management operations on a specific virtual storage pool.

The allocation state identifies each LUN as:

HOLD: LUN is discovered but not yet ready to be provisioned. For example, the storage array may still be in process of being fully configured.

FREE: The LUN is available to be provisioned but not yet owned.

RESERVED: The LUN has been reserved for use by a specific storage account, but not yet provisioned. This state may have a time-limited policy applied that allows the LUN to degrade back to a FREE state.

ASSIGNED: The LUN has been assigned to a specific storage account.

- **Increased Integration with Host Applications**
  To increase the integration with host applications, the storage consumer can now be a Volume Manager group, file system or specific database instance. This allows a finer level of administration in SAN provisioning, allowing capacity management on an application basis. Second, an event-automation mechanism is provided to allow a Volume Manager

group, file system or DB to be triggered to automatically grow its capacity, allowing the application to use the newly provisioned storage automatically. Conversely, the DB, file system or Volume Manager group can use threshold triggers to indicate to SANPoint Control that it is close to maximum capacity use, allowing the SAN administrator to provision new storage of the appropriate class proactively.

Introducing the V³ Intelligent Provisioning Service (IPS) for Automation of Time-to-Capacity

A new element of V³ SAN Infrastructure is introduced as an optional element of this next release of SANPoint Control to automate the provisioning of storage to storage accounts intelligently. This technology is a rules-based engine that analyzes the SAN storage configuration and makes a storage provisioning recommendation based on the needs of a storage consumer. It is built on the information available from the V³ SAL and the administrative constructs in SANPoint Control for storage pools and storage accounts described above. This recommendation can be presented to the SAN Administrator, simplifying the Administrator's task of selecting the most appropriate storage, or can trigger a series of events to provision the recommended storage into the requesting storage consumer automatically.

The V³ IPS will be available as an add-in option to SANPoint Control 2.5 and provide some extensions to SANPoint Control GUI, but could also be used bundled into other future SAN management applications. Consistent with the V³ SAL strategy, VERITAS plans to make the APIs available for this service, allowing it to be used and integrated with third-party storage management applications. The V³ IPS in most cases will be implemented as an asymmetrical storage-virtualization service located on a management server or special function appliance and likely co-located with the V³ SAL DB.

Level 3: The V³ SAN Volume Manager

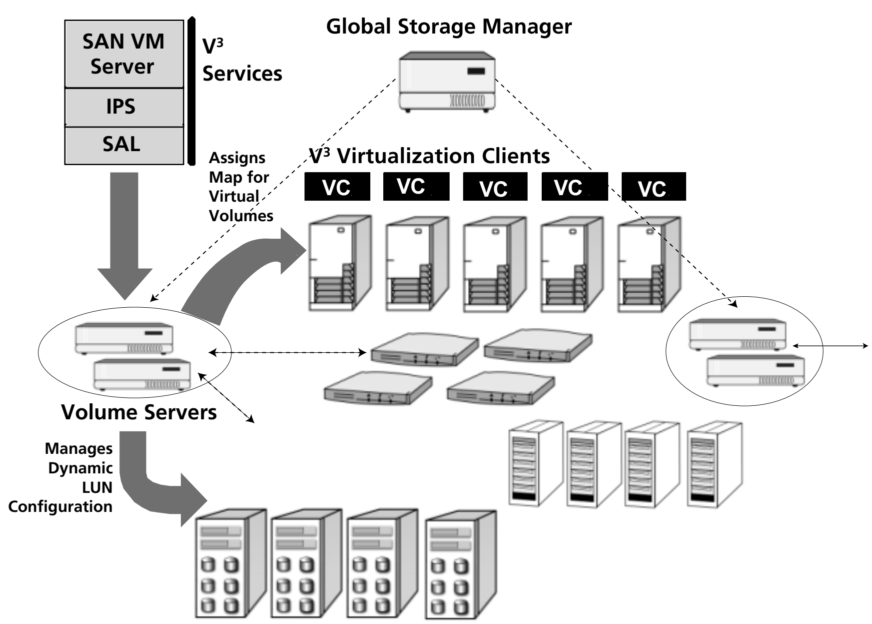

Figure 10

The levels of time-to-capacity described above provide a comprehensive suite of tools to manage storage provisioning of large heterogeneous SANs at the LUN level. As SAN capacity and the number of storage accounts become ever larger, it becomes necessary to look for ways to perform SAN administration at higher level of abstraction, a logical SAN volume. Managing logical SAN volumes promises to provide a dynamic volume provisioning service, allowing the SAN administrator to manage logical volumes SAN as easily as one can today in a single-sever environment with VERITAS Volume Manager. This new service, the V³ SAN Volume Manager, moves up a level of abstraction from LUN-level SAN management and uses an asymmetrical storage-virtualization architecture as discussed previously to provide maximum scalability and heterogeneity. The volume server can be implemented using either an out-of-band appliance or a standard host server. V³ SAN Volume Manager will be backward compatible with SANPoint

Control product line and host-based Volume Manager, but will also enable broader operating system heterogeneity with using lightweight volume clients. Most important, V^3 SAN Volume Manager builds incrementally on well-proven Volume Manager technology.

A logical SAN volume uses a single-volume object to logically represent a storage pool as described above—e.g., a set of storage within the same storage group and storage state and within a particular storage account. The volume server maintains and essentially owns the mapping information from the LUNs on the SAN to a SAN volume. Once a logical SAN volume is defined, the administrator may want to apply traditional volume-management operations such as snapshots, data replication and dynamic configuration changes onto the logical SAN volume with a single operation aggregating many tasks into one logical operation and centralizing administrative functions onto the dedicated virtualization server or appliance.

Using a SAN management application such as a future version of SANPoint Control, a logical SAN volume is allocated to a single virtualization client or even a group of virtualization clients to facilitate data sharing. These virtualization clients, each of which is mapped to a storage account, do not have any awareness of how LUNs on the SAN are being used but are merely provided with a map from the volume server of the disk locations that are part of the SAN volume assigned to that client. Any updates made by the volume server because of storage configurations or failure recovery are provided to the volume client as they occur. This decoupling is to simplify the host server operations required to provision and reconfigure SAN attached storage.

The V^3 SAN Volume Manager architecture also addresses one common pitfall of out-of-band storage-virtualization: scalability. To provide the appropriate level of manageability, each volume server must be attached directly into each SAN fabric topology. Because data-center-class SAN fabrics are often constructed using several SAN fabrics operating in "parallel" but not directly connected to each other, this restriction can demand complex connectivity from a single-volume server to each SAN topology. V^3 SAN Volume Manager removes this obstacle by allowing co-coordinating across multiple virtualization servers using a global storage server that acts as a "manager of managers" for all the virtualization servers. This global storage server connects into each volume server using standard network connections and assigns storage accounts to each volume server as connectivity allows.

V^3 SAN Volume Manager builds on the V^3 SAN Infrastructure services provided by SAL and IPS. The V^3 SAL provides the SAN discovery and control for fabrics and LUN for V^3 SAN Volume Manager in a vendor-independent manner. The V^3 IPS option allows the V^3 SAN Volume Manager to quickly and intelligently map available LUNs into SAN volumes that are appropriate to each storage account. Also, virtualization clients can be bundled in with the VERITAS in-band V^3 Storage Appliance for highly scalable NAS or SAN storage firewalls.

Last, one dramatic benefit of V^3 SAN Volume Manager is the ease in which data can be dynamically shared between virtualization clients of different operating systems. Heterogeneous data sharing is usually restricted because of differences in on-disk format for each operating system. The volume server for V^3 SAN Volume Manager can compensate for these differences before it passes the disk location map to each virtualization client. For file or application services on the virtualization clients that use similar application data formats, this new virtualization layer can then be used as the basis for sharing data across platforms.

Although revolutionary in concept, V^3 SAN Volume Manager can greatly improve time-to-capacity and minimize risk of adoption common to asymmetric storage-virtualization services by building incrementally on the Volume Manager technology. In addition, a very powerful and modular time-to-capacity implementation is possible by integrating volume clients into host-based Volume Manager products and V^3 Storage Appliance, the intelligent services of V^3 SAN Infrastructure (SAL and IPS) and the data-center-class SAN-management capabilities of SANPoint Control.

A Secure "Storage Service Provider" Model

The new concept of a split storage administrator/consumer model is being applied in three distinct ways: outsourced storage service providers, who host and manage storage resources at an off-site facility; outsourced managed service providers, who procure and manage storage resources hosted at the customer's facility; and an "in-sourced" specialized department within the bounds of a corporation that is accountable for managing centralized storage resources. For all three models, one of the most important Quality of Storage Service metrics is in managing security of access to storage. When departments or companies cede control over their business-critical data, they typically demand commitments that their data cannot be deliberately or accidentally stolen by another storage consumer using this centralized storage service. In addition, storage consumers may request access control to these centralized resources in either file access protocols, in which data is presented as a virtual file server, or block access protocols, in which data is presented as a virtual disk device. In all three cases, the providers can use storage-virtualization architectures to deliver this secured storage access.

An extreme model of providing secured access is to physically separate the storage devices and storage networks on a consumer-by-consumer basis. Although this option provides absolute guarantees, it does not allow the storage administrator to leverage the economies of scale in deploying high-capacity storage networks and storage subsystems. Thus, this model will be the most expensive model for the storage consumer.

As described earlier, using SAN management solutions like SANPoint Control to centrally manage LUN Masking and fabric zoning allows the SAN administrator to allocate storage capacity dynamically to individual storage consumers. To make this environment secure, the SAN management product must provide roles-based administration and authentication to ensure that only trusted administrators can modify the SAN configuration. However, this solution must also ensure that a storage consumer can never "see" beyond its allocated portion of the SAN. And few security protocols are built into SAN standards or products that allow the level of enforcement typical of TCP/IP-based networks.

One approach is to deploy an out-of-band storage-virtualization appliance and require all storage consumers to gain access to storage resources via this service. Although this deployment can be done for both file and block access, this model requires all storage consumers to use a secure client, using either a special-purpose HBA or software agent on every storage consumer. This secure client would be responsible for ensuring that malicious or accidental sources could not work around the security access protocols. The obvious drawback of this model is that it can lead to high maintenance to ensure that all storage consumers are using a secure client.

A more manageable and flexible approach is to use both NAS and SAN forms of in-band storage-virtualization appliances, such as VERITAS V^3 Storage Appliances, essentially as a storage firewall (see diagram below). In this model, all traffic into the centralized storage service must pass through the storage firewall. The appliances have access to the entire SAN resources behind them and can present the logical storage within its "island of control" as either a file system or virtual disks. Existing file access control protocols can be used for storage consumers using file access, and the standard LUN masking techniques described earlier are implemented in the appliance to provide secured access for storage consumers using block access.

By embedding V^3 SAN Infrastructure into the V^3 Storage Appliances such as the V^3 SAL, management products like SANPoint Control can now be used to provide the time-to-capacity capabilities for allocating the storage behind the storage firewall. By incrementally adding the V^3 IPS service and eventually the dynamic volume provisioning services on in-band appliances behind the storage firewall, the storage administrator can build up a powerful and secure storage services solution with no requirement for specialized security clients on the storage consumers. As advanced services, such as data replication using Volume Replicator, become available on the V^3 Storage Appliance platform, service providers can even provide flexible disaster recovery and data-migration services. And they can do so using a multivendor SAN infrastructure, providing them a range of choice to help them manage to profitability.

VERITAS is working to integrate its V^3 SAN Infrastructure, SANPoint Control and V^3 Storage Appliance products tightly in order to provide all types of service providers with the tools they need to implement secure and well-managed storage services.

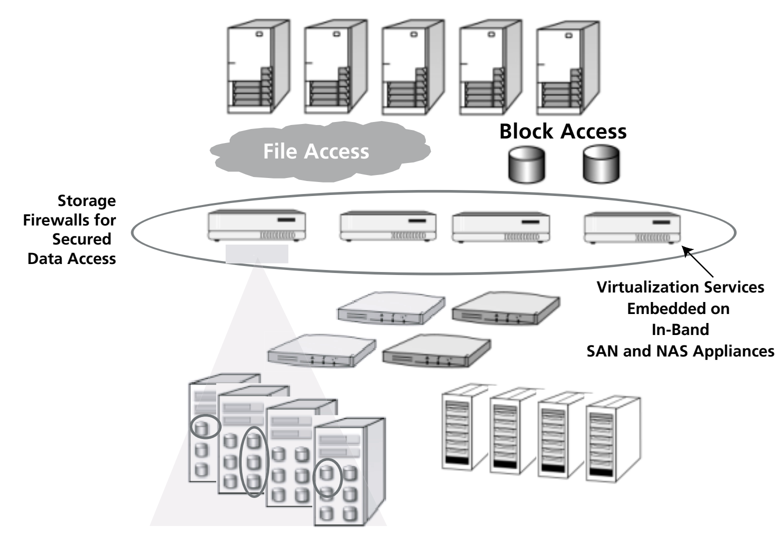

Figure 11

Summary and VERITAS Future Directions

A QUICK GUIDE TO TODAY'S STORAGE-VIRTUALIZATION PRODUCTS FROM VERITAS

Virtualization Architecture	VERITAS Product (Futures indicated by shading)	Description
Host Based	Volume Manager	Full-featured logical volume manager.
	SANPoint Foundation Suite using the Cluster Volume Manager	High-performance data sharing using a clustered file system on top of the symmetric storage-virtualization capabilities of Cluster Volume Manager. Operates today in a homogeneous Solaris environment.
	SPC using the V^3 SAL deployed on host server	Heterogeneous SAN management application for visualization, administration of zoning and storage allocation, event handling and reporting.
Storage Subsystem	V^3 Storage Appliance for NAS and SAN	Use V^3 Storage Appliance software as an alternative to proprietary filers or array controllers.
	SPC interoperates with leading third-party array controllers with embedded vendor-specific virtualization via V^3 SAL.	SPC integrates with LUN management in HDS and EMC arrays using the V^3 SAL.
In-Band Virtualization Appliance	V^3 Storage Appliance for NAS and SAN	V^3 Storage Appliances can be used in any SAN fabrics and storage subsystems tested with Volume Manager.
Out-of-Band Virtualization Appliance	{ V^3 Intelligent Provisioning Service (IPS) }	Automated LUN-level storage allocation
	{ V^3 SAN Volume Manager }	SAN-wide dynamic volume provisioning
	SPC using V^3 SAL DB on a dedicated server	Special-purposed appliance for heterogeneous SAN management

Futures indicated by shading

Virtualization Futures

VERITAS is investigating three future areas of storage virtualization to deliver better solutions for the storage administrator:

- **Extending Symmetrical Data Sharing to Heterogeneous Servers.**
  A key barrier to expanding data sharing beyond homogeneous operating systems is the need to ensure a common on-disk format for stored data. VERITAS is developing common disk-format technologies that will extend existing data sharing products based on Cluster Volume Managers, such as SANPoint Foundation Suite, into heterogeneous data sharing solutions. It is expected that these technologies will be used initially for sequential or nonconcurrent data sharing in which servers can import data from another operating system but don't need real-time concurrent write access.

- **Develop "Quality of Storage Service" protocols**
  It is likely that more than the storage-virtualization architectures discussed will survive and be built by different vendors. To ensure that these products can not only interoperate but also cooperate to deliver well-managed storage services, VERITAS is defining protocols with its partners that will allow the various components of a storage-virtualization solution to communicate with each other and to use each other's services.

- **Integrating Offline and Online Technologies for True Virtual Disk Solutions**
  The mainframe storage-virtualization environment allowed administrators to move data seamlessly between offline and online media using manual provisioning and automated techniques. This allowed the administrator to optimize the use of typically more expensive online storage subsystems for critical datasets and take advantage of faster online technologies to accelerate backup and archiving process using staged data migration. As open-system SANs grow to enterprise scale and require the level of accountability typical of legacy mainframe systems, VERITAS believes that this tighter integration of online and offline technologies is necessary to deliver a truly complete virtual storage environment.

1 Refer to QoSS Brief for more details.
2 The VERITAS Volume Manager is available on Solaris, HP/UX, Windows NT and Windows 2000. AIX and Linux versions are expected to be available by mid-2002.
3 See architecture briefs/appendix for V^3 SAL and Storage Appliance for more details.
4 SANPoint Control 2.0 ships in the first half of 2001, with support for Solaris and Windows servers.
5 SANPoint Control 2.5 is expected to be available in the second half of 2001 on Solaris, HP/UX and Windows platforms.

VERITAS WHITE PAPER

VERITAS Software Corporation
Corporate Headquarters
1600 Plymouth Street
Mountain View, CA 94043
650-527-8000 or 800-327-2232

For additional information about VERITAS, its products, or the location of an office near you, please call our corporate headquarters or visit our Web site at
www.veritas.com

VER05-STORVERWPR-0000 • 90-01136-399