



Executive Briefing

Nexsan Promotes Greater IT Infrastructure Flexibility with API Support for Symantec's Veritas Cluster Software for High Availability of Applications Across Multiple OS

Jack Fegreus
June 27, 2011

THE PRIVATE CLOUD: DISTRIBUTED COMPUTING 3.0

A 2010 Symantec data center survey that examined 1,780 IT sites pegged just 18% as operating without a virtual infrastructure either in place or planned. On the other hand, roughly half of the sites with a virtual operating environment (VOE) were planning to expand their use of virtual machines (VMs) and begin moving mission-critical applications into their VOE. This step in virtualization evolution is essential for IT to create an agile infrastructure that enables data centers to economically scale and operate a cloud architecture.

Nexsan Storage Value Proposition:

- 1) Leverage strict storage standards compliance to integrate tightly with multiple operating systems and hypervisor APIs to leverage productivity features developed by third-party vendors:** Through strict compliance with the SCSI-3 standard and extended integration Symantec's Veritas Storage Foundation (SF), Nexsan allows IT to tightly integrate Nexsan arrays with Veritas SF-HA software to configure a responsive high-availability Veritas cluster server.
- 2) Create a highly scalable storage cloud that can be managed as a single virtual entity:** The Nexsan management software integrates with the Windows Virtual Disk Service (VDS) to provide a complete single-pane-of-glass storage management interface for Nexsan and Microsoft storage management tools, including Server Manager and Storage Manager for SANs, to provide full access to all Nexsan configuration data and directly manage any Nexsan device.
- 3) Automate I/O Performance:** Through strict compliance with advanced storage standards, Nexsan storage arrays offer sophisticated dynamic MPIO options, including Asymmetric LUN Access (ALUA) load balancing, SCSI storage reservations that persist over bus resets, and biased FC/iSCSI path utilization.
- 3) Maximize Density and Reliability with Hierarchical Storage:** Nexsan devices support a mix of SSD, SAS, and SATA drives to support a full range of SLA storage requirements with respect to capacity, throughput, access, and reliability.

With the rapid expansion of resource virtualization, the IT distributed architecture paradigm is approaching a critical tipping point as the notion of distributed computing is generalized into what is now dubbed cloud architecture. The pivotal resource for this tipping point is a SAN-based storage infrastructure that scales out in both capacity and performance.

The importance of storage comes from the need for IT to meet the data access and throughput requirements of systems running a hypervisor, which concentrate the I/O streams of multiple Virtual Machines (VMs) to create an absolute requirement for low I/O latency. As a result, storage resources are inextricably linked to the capital and operational expenses that IT must restructure to maximize the return on investment (ROI) of a virtualization initiative.

To meet IT storage infrastructure needs, Nexsan's storage architecture is designed

around strict compliance with storage standards in order to integrate tightly with multiple operating system and hypervisor APIs. As a result, Nexsan arrays allow IT to maximally leverage the productivity features developed by third-party vendors. In particular, Nexsan's SCSI-3 implementation has led to Nexsan becoming a Symantec Technology Enabled Partner that is certified with all Veritas Storage Foundation (SF) products. This is especially important for IT sites adopting cloud-based infrastructure featuring hosted applications that must be optimized for cloud computing.

STORAGE CLOUD CONFIGURATION MANAGEMENT

To simplify server provisioning in a scale-out storage infrastructure, Nexsan embeds a WEB-enabled Graphical User Interface (GUI) in each storage array. Through this single Nexsan GUI, IT administrators can access all of the Nexsan devices in a storage cloud and provision every server with a hierarchy of logical volumes built from RAID arrays using SSD, SAS, or SATA drives. This interface also makes it easy for IT administrators to leverage Nexsan's MultiSAN I/O architecture for high availability with respect to data access and parallel I/O throughput over multiple SAN fabric paths.

IT administrators typically use this capability to share logical volumes among servers via FC and iSCSI paths. More importantly administrators use this utility to expose or mask Nexsan's sophisticated multi-path I/O (MPIO) options on a SAN depending on the capabilities of host servers. For example, hosts can be restricted to accessing volumes on a Nexsan array using a simple non-redundant controller setup. Alternatively, host servers can be allowed to access all array ports and all LUNs (APAL mode) that have been associated with a logical storage volume.

“Through this single Nexsan GUI, IT administrators can access all of the Nexsan devices in a storage cloud and provision every server with a hierarchy of logical volumes built from RAID arrays using SSD, SAS, or SATA drives.”

When configured in APAL mode, a Nexsan array provides data access to logical volumes via both the controller handling the physical drives that make up the logical volume and the other controller in the array. This configuration provides hosts with high availability with respect to logical storage volumes; however, this configuration also exposes a performance penalty. When a host makes an I/O request for a logical volume through a controller not servicing the physical drives underlying the volume, a trespass occurs as the I/O request is automatically redirected to the other controller.

To avoid this performance issue, Nexsan complies with the Asymmetric Logical Unit Access (ALUA) method specified in the SCSI-3 standard when exporting target volumes in APAL mode. The Nexsan array distinguishes paths that connect to an FC port on the controller servicing the device—dubbed active and optimized—from paths connecting to an FC port on the other controller—active but are not optimized. Nonetheless, for this sophisticated array-based MPIO mechanism to work, it must be fully recognized by the operating system of the host accessing the Nexsan array.

Recent versions of the VMware hypervisor, ESX 4.1 and ESXi 4.1, and the Dynamic Multi-pathing (DMP) feature of Veritas SF 5.0 recognize and work with Nexsan's ALUA MPIO configuration. To demonstrate the value of this SCSI-3 method for multi-pathing, we upgraded a VMware host running ESX 4.0, which is not optimized for ALUA, to ESX 4.1, which is optimized for ALUA. This upgrade automatically boosted I/O throughput for VMs resident on a datastore imported from a Nexsan array in APAL mode by 20%.

ENTERPRISE CLUSTER ENABLEMENT

More importantly, Nexsan's clean implementation of the SCSI-3 standard is critical for successful integration of Nexsan arrays with Symantec's Veritas SF, which allows IT administrators at sites with large heterogeneous SANs to fully virtualize connectivity from hosts to storage arrays. DMP in VSF enhances I/O performance by distributing I/O requests across all available paths according to pre-defined load balancing policies. Furthermore, DMP provides automatic re-routing of I/O requests over alternate paths transparently when SAN paths fail, which enhances data and application availability.

Veritas Storage Foundation-HA (SF-HA), which provides the software infrastructure for a Veritas Cluster Server (VCS), utilizes the SCSI-3 constructs of registration and reservation as the foundation for

high-availability services that protect shared data from being corrupted in a sophisticated SAN environment. In particular, strict implementation of the SCSI-3 specification ensures the ability to make persistent reservations across SCSI bus resets, to discover multiple paths from a host to a disk, to access devices safely from multiple nodes, and to block access from nodes—a process called I/O fencing.

By leveraging the SCSI-3 specification, VCS provides a global locking mechanism to explicitly determine the nodes that remain in a cluster during a membership change and block systems not acknowledged as cluster members from accessing shared storage. In practice, systems join a cluster by registering a key on a SCSI-3 volume. Next, cluster members are able to ensure that only registered systems are able to write data by setting reservations to Write Exclusive Registrants Only. Furthermore, failed systems can be removed from a cluster by simply removing that system's key.

For IT administrators, strict end-to-end SCSI-3 compliance between Veritas SF-HA and Nexsan storage arrays establishes the foundation of a rock-solid cluster configuration. Moreover, Nexsan has worked with Symantec to include specific Nexsan array plugins for all Nexsan arrays in all versions of Veritas SF. These plugins come in the form of an array support library (ASL) and an array policy module (APM).

To enable VCS to leverage all of the features of a Nexsan array, the Nexsan ASL contains explicit instructions to discover device attributes, claim devices, and correlate paths to the same device. The Nexsan APM compliments the ASL with a set of commands to efficiently manage multiple paths when performing such tasks as path

“Strict end-to-end SCSI-3 compliance between Veritas SF-HA and Nexsan storage arrays establishes the foundation of a rock-solid cluster configuration.”

selection, path failover, SCSI reservation, and SCSI release.

STACKING UP A SCSI-3 CLUSTER

To assess the effectiveness of a Veritas cluster environment built on Nexsan storage arrays, openBench Labs set up two SunFire T2000 servers running Solaris 10 on SunOS 5.1 foundation. To provide storage for each host server, dubbed by Symantec as a System Under Test (SUT), we set up a Nexsan SATABeast2 storage array. The SATABeast2 featured two controllers with two 4Gbps FC ports on each controller. To reflect our Nexsan array configuration, each SunFire server was provisioned with two dual-port 4Gbps QLogic QLE2462 Fibre Channel HBAs.

We then loaded version 5.1 of Veritas SF-HA on each SunFire server in order to configure and test a Veritas cluster. From the Nexsan array, we exported 96 LUNs configured on four internal RAID-5 arrays to each of the members of the Veritas cluster. To support internal cluster operations we used three LUNs to configure a disk group, dubbed vxencoord, for I/O fencing.

To test the functionality of cluster's I/O features for clients, we used Veritas Volume Manager (VxVM) to create four striped volumes, each of which contained 23 LUNs exported from the Nexsan SATABeast2 array. These VxVM volumes were placed in two disk groups and each disk group was assigned to a default host server.

To coordinate the compliance testing of our Veritas SF-HA cluster, we configured a VM running RedHat Enterprise Linux v5 on an ESX4.1 host server. All test scripts associated with the Veritas SA-HA compliance suite were configured and launched from this host, which the compliance suite dubs the Test Master (TM). These test scripts fell into three categories:

- 1) Infrastructure configuration, including ASL and APM software validation;
- 2) DPM host and array functionality;
- 3) VCS failover.

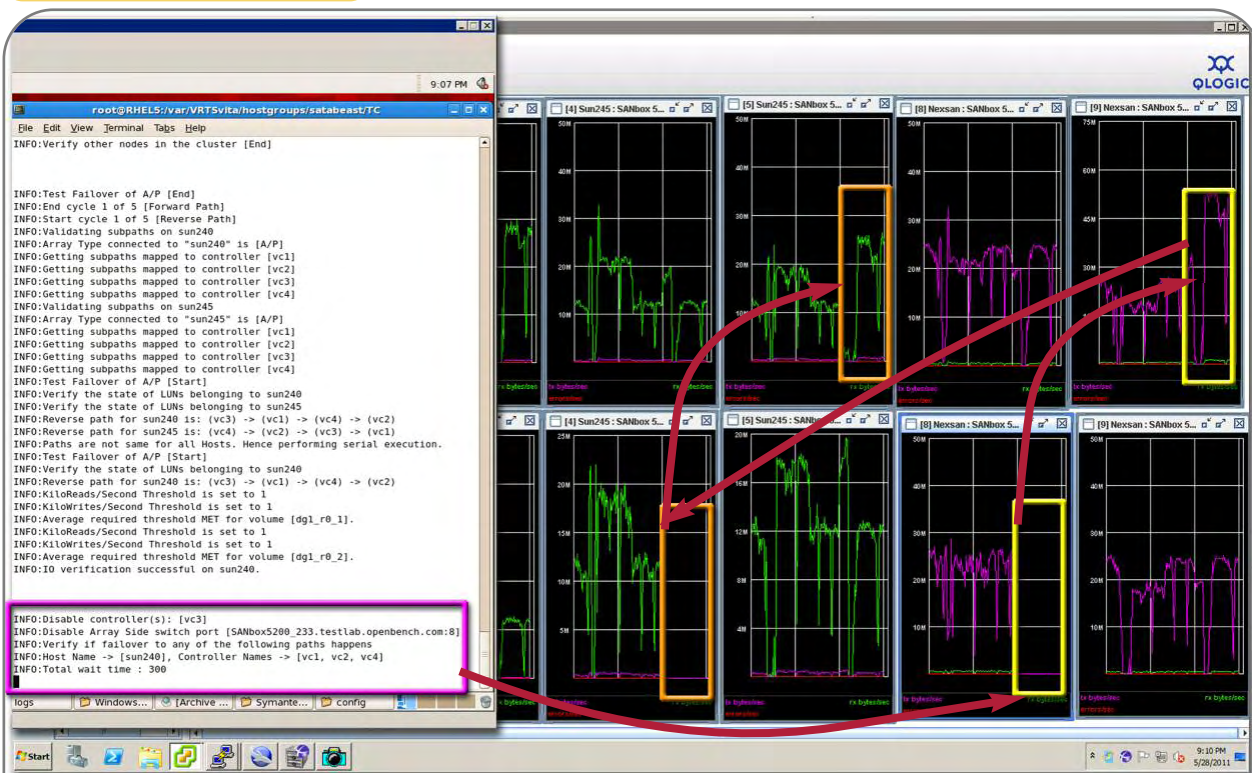
To help validate test results, all DMP and VCS test scripts, which disrupted SAN fabric paths and VCS member nodes, were run in conjunction with a background stress test. Each stress test continuously streamed I/O requests—80% writes and 20% reads—as the main test script brought either FC host ports, FC array ports, or VCS member nodes on- or off-line. As a result, the ability of the Nexsan-VCS environment to deal with disruptions to the SAN fabric and cluster could be evaluated by measuring latency and recovery levels in the corresponding I/O throughput stream.

During each of the eight DMP tests, which examined the effects of SAN fabric failures, the stress test generated a local I/O request stream directly on each of the Sun servers. This local I/O request stream resulted in an I/O throughput rate that averaged around 120MB per second. On the other hand, the three VCS tests dealt with the effects of cluster node failures. As a result, it was necessary to run the I/O stress tests externally on client systems that utilized LAN-based data-sharing.

For data sharing, we used the cluster-aware Veritas NFS service. In particular, we associated the logical volumes in each of our two cluster disk groups with a virtual IP address owned by the cluster. Next, we setup a second VM running RedHat Enterprise Linux v5 as an NFS client and imported the NFS volumes. Finally, we directed the stress test run in conjunction with the VCS failover tests to run on this VM and the volumes imported via NFS.

As with DMP path-failure scenarios, VCS node failures were readily handled by the Nexsan-VCS environment. As soon as the lack of a heartbeat from a cluster member was confirmed, the remaining node utilized the SCSI-3 reservation mechanism to assume control of the failed node's disks. With full control of all of the LUNs exported to the cluster from the Nexsan array, the remaining node resumed exporting the NFS volumes using the cluster's virtual IP address.

NEXSAN-VCS I/O RESILIENCE

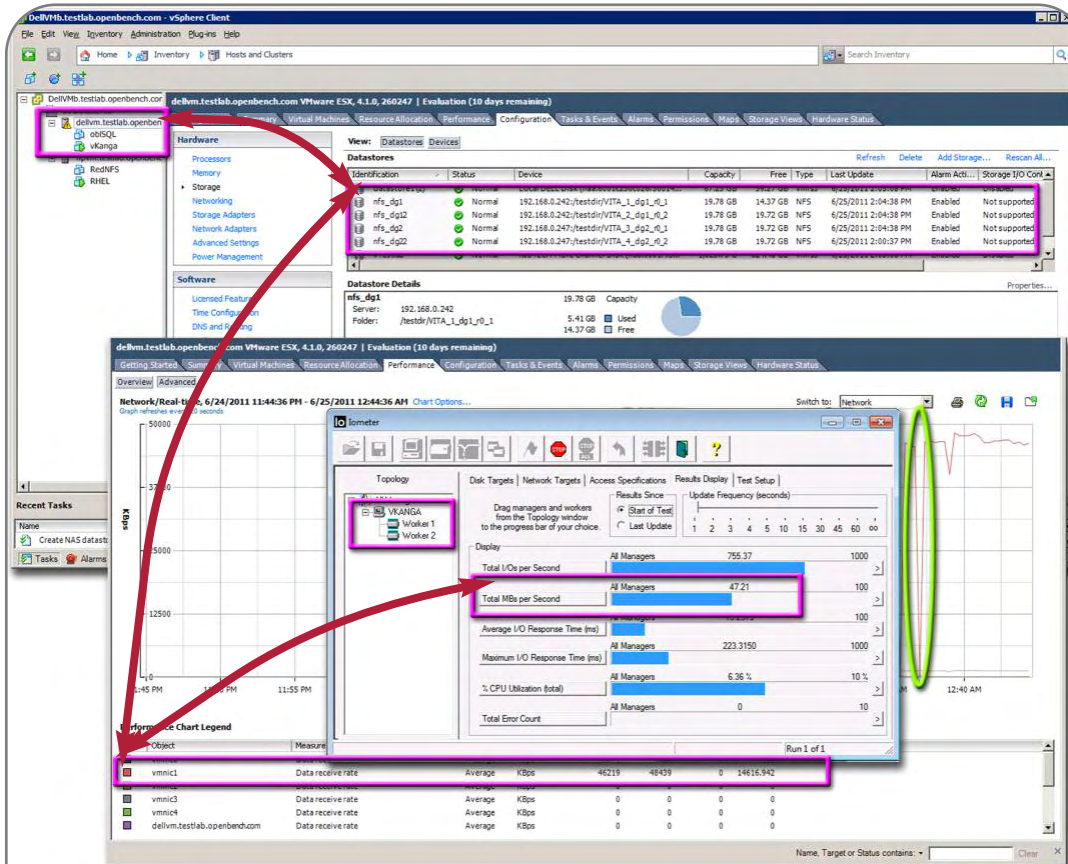


The lion's share of the Symantec validation suite tests DMP path and VCS host failures. The DMP tests are either host- or array-centric as each test fails sequences of FC ports and measures the response of both the host and the array to the port failure. In this DMP test the third controller port on the array is disabled. This causes a failover within the Nexsan array to another FC port. In turn, the server ports zoned with the failed array port drop all I/O while the server ports zoned with the array port to which the Nexsan array failed over I/O processing record a corresponding increase in I/O traffic.

As a final test of our Nexsan-VCS infrastructure, we used vSphere to import the four NFS volumes from the cluster's NFS service to a VMware host and configure the volumes as VMFS datastores. We utilized one datastore to provision a work disk for a VM

running Windows 7. We then utilized Iometer on this VM to act as an ersatz Veritas stress test writing data to the NFS-backed datastore at roughly 47MB per second.

NFS-BASED VMWARE PERFORMANCE



Using vSphere, we imported the four shared-disk volumes as VMFS datastores on a VMware host. After provisioning a work disk for a VM running Windows 7 using one of the NFS-based datastores, we launched Iometer in the VM and initiated a stream of large-block writes, which generated an average throughput rate of 47MB per second. With Iometer writing data at a steady state, we rebooted the Sun server controlling the shared disk backing the Iometer test. After a pause in processing on the VM, the remaining Sun server gained control of the shared disk, resumed the NFS service, and I/O processing automatically resumed on the VM.

As soon as the Iometer write stream reached a steady state, we booted the Sun server that owned the shared disk used to back the VM's work disk targeted by Iometer. Using SCSI-3 fencing the second Sun server easily

took control of the 23 SATABeast LUNs used to create the shared disk on the cluster. At the same time, the Veritas NFS service failed over to the second server, allowing that server to export the shared disk from the same IP address. As a result, the VMware host reconnected to the datastore and the Iometer application automatically resumed processing.

What's more, by integrating its storage firmware with the Windows Virtual Disk Service (VDS), Nexsan provides real-time SAN-wide visibility into the Nexsan storage cloud. At a base level, logical disks created on Nexsan devices have extended property sheets that include the name of the specific Nexsan device on which the logical disk was created, the array used within that device, and the internal name of the array partition that was mapped

as a LUN. As a result, IT administrators gain immediate access to essential information for SAN management that greatly simplifies all tasks associated with resolving root SAN performance issues. With the majority of IT administrators at many enterprise sites relying solely on out-of-date spread sheets mapping logical disks on hosts to array partitions on storage devices, this deep real time storage mapping is an essential resource for IT.

AUTOMATING PERFORMANCE OPTIMIZATION

While IT overhead is the primary factor in operating cost reduction, CIOs also have to worry about meeting Service Level Agreements (SLAs) made with Line of Business executives. These executives think in terms of business process parameters, so SLA objectives focus around two distinct notions: process availability and process performance. With storage at the foundation of any IT SLA guarantee for either process availability or performance, simply building a fast low-latency device is not sufficient to ensure SLA compliance.

The robust Nexsan hardware architecture provides IT with a storage platform that can satisfy a wide range of performance metrics with respect to access (IOPS), throughput (MB per second), or capacity (price per GB). In particular, Nexsan storage devices support multiple internal controllers that connect to multiple FC and iSCSI ports. All of these ports promote comprehensive redundancy to ensure availability.

“Nexsan provides automation for high-performance scale-out storage architecture via seamless integration with Veritas Cluster Server, as a Symantec Technology Enabled Partner.”

What's more, Nexsan provides automation for high-performance scale-out storage architecture via seamless integration with Veritas Cluster Server, as a Symantec Technology Enabled Partner. When failures occur, VCS can automatically choose the least utilized server. In addition, advanced failover logic that utilizes SCSI-3 storage standards ensures application uptime is maximized and server resources are utilized as efficiently as possible, without imposing costly homogeneous hardware restrictions.

The bottom line for IT sites struggling with implementing VOE and cloud-based computing is Nexsan's ability to provide a highly manageable multi-fabric SAN infrastructure that is capable of supporting a wide range of performance characteristics. By complimenting a high performance foundation with sophisticated management software, along with tight integration with host servers and applications, Nexsan empowers IT to use any advanced automation features to provision storage, commission and decommission applications, and failover applications running among various hosts.

Jack Fegreus is Managing Director of openBench Labs and consults through Ridgetop Research. He also contributes to InfoStor, Virtual Strategy Magazine, and Open Magazine, and serves as CTO of Strategic Communications. Previously he was Editor in Chief of Open Magazine, Data Storage, BackOffice CTO, Client/Server Today, and Digital Review. Jack also served as a consultant to Demax Software and was IT Director at Riley Stoker Corp. Jack holds a Ph.D. in Mathematics and worked on the application of computers to symbolic logic.