OUTDATED

Near-FT Support of OLTP Applications using a HA StarWind Virtual SAN and a Hyper-V Cluster

Cost Effective HA Storage for Critical Business Applications

FEBRUARY 2015

TECHNICAL PAPER BY DR. JACK FEGREUS

≈ StarWind

Trademarks

"StarWind", "StarWind Software" and the StarWind and the StarWind Software logos are registered trademarks of StarWind Software. "StarWind LSFS" is a trademark of StarWind Software which may be registered in some jurisdictions. All other trademarks are owned by their respective owners.

Changes

The material in this document is for information only and is subject to change without notice. While reasonable efforts have been made in the preparation of this document to assure its accuracy, StarWind Software assumes no liability resulting from errors or omissions in this document, or from the use of the information contained herein. StarWind Software reserves the right to make changes in the product design without reservation and without notification to its users.

Technical Support and Services

If you have questions about installing or using this software, check this and other documents first - you will find answers to most of your questions on the Technical Papers webpage or in StarWind Forum. If you need further assistance, please contact us.

In 2016, Gartner named StarWind "Cool Vendor for Compute Platforms".

Gartner does not endorse any vendor, product or service depicted in its research publications, and does not advise technology users to select only those vendors with the highest ratings or other designation. Gartner research publications consist of the opinions of Gartner's research organization and should not be construed as statements of fact. Gartner disclaims all warranties, expressed or implied, with respect to this research, including any warranties of merchantability or fitness for a particular purpose.

About StarWind

StarWind is a pioneer in virtualization and a company that participated in the development of this technology from its earliest days. Now the company is among the leading vendors of software and hardware hyper-converged solutions. The company's core product is the years-proven StarWind Virtual SAN, which allows SMB and ROBO to benefit from cost-efficient hyperconverged IT infrastructure. Having earned a reputation of reliability, StarWind created a hardware product line and is actively tapping into hyperconverged and storage appliances market. In 2016, Gartner named StarWind "Cool Vendor for Compute Platforms" following the success and popularity of StarWind HyperConverged Appliance. StarWind partners with world-known companies: Microsoft, VMware, Veeam, Intel, Dell, Mellanox, Citrix, Western Digital, etc.

Contents

What's In This Document

In this analysis, openBench Labs assesses the performance of a StarWind High Availability (HA) Virtual Storage Area Network (Virtual SAN) in a Microsoft® Hyper-V® failover cluster. This paper provides technically savvy IT decision makers with the detailed performance and resource configuration information needed to analyze the trade-offs involved in setting up an optimal HA configuration to support business continuity and any service level agreement (SLA) with line of business (LoB) executives.

For our test bed, we set up three Windows Hyper-V host servers: Two ran Windows 2012 R2, while the third ran Windows 2012. We installed StarWind's iSCSI Virtual SAN software directly on each Hyper-V host not in a Hyper-V virtual machine (VM). Consequently, we were able to directly leverage all local resources, including direct attached storage (DAS) devices, in creating a StarWind Virtual SAN. More importantly, all of the logical volumes imported into a Hyper-V cluster on Windows 2012 R2 hosts were fully compatible with all of the advanced CSV features for clusters.

Using the three Hyper-V hosts, we configured an iSCSI quorum disk in our 3-node HA Virtual SAN. Consequently, we were able to configure a Hyper-V failover cluster using the two hosts running Windows 2012 R2 and a quorum disk that would remain available if either or both of the Hyper-V failover cluster nodes failed. We then introduced a vSphere® 5.5 VM running StarWind to create a 4-node Virtual SAN to deploy shared storage resources for deploying near fault tolerant (FT) VMs on the Hyper-V cluster.

With this layered HA configuration, we focused our attention on the ability to scale-out the performance of HA storage in a common IT scenario: a SQL Server database running an OLTP business application. We analyzed our HA failover tests from both the perspective of LoB users, with respect to the time to restore full Exchange processing, and IT management, with respect to the time to restore full HA capabilities to the StarWind Virtual SAN, following a system crash in our virtual infrastructure (VI).

Snapshot of Findings

StarWind Virtual SAN Software		
4-Node HA Virtual SAN Protection for a Hyper-V Cluster - DR Test Assessment		
HA Failure Simulation	**LoB Results**	**IT Results**
StarWind VM Server Failure	Full Processing Never Interrupted	• Full Virtual SAN Synchronization Restored in 73 Minutes while Running OLTP Queries at 160 TPS
Hyper-V Host Live Migration	Full Processing Never Interrupted	• Full Virtual SAN Synchronization Never Lost

*"A StarWind Virtual SAN also provides critical heartbeat and synchronization services for iSCSI
targets without special caching hardware, such as flash devices, which enables any IT site, including
remote office sites, to cost-effectively implement HA for both hosts and storage resources."*

Driving HA Through Virtual SAN

In a Hyper-V environment, shared storage plays an important role in implementing a failover cluster.
When a shared volume is imported into a failover cluster, as a cluster shared volume (CSV), only one
Hyper-V server can be the owner of the disk resource at a particular point in time; however, both
servers can access the disk. Consequently, there is no need to failover disk resources when failing over
a VM. More importantly, the VM failover process in a Hyper-V cluster will have little or no impact on a HA
StarWind Virtual SAN.

While shared storage has long been synonymous with an expensive hardware-dependent SAN;
however, StarWind is seeking to change that notion by providing an iSCSI-based Virtual SAN that can be
deployed and managed using standard Windows Server tools, such as the Hyper-V and the Failover
Cluster Manager. Building on a mature iSCSI software foundation, StarWind uses advanced virtualization
techniques to represent NTFS files as logical disk devices that can be decoupled from iSCSI targets to
enable device modification. In particular, StarWind provides IT administrators with a wizard-based
console for creating feature-rich logical device files located on any standard storage resource.

A StarWind Virtual SAN also provides critical heartbeat and synchronization services for iSCSI targets
without special caching hardware, such as flash devices, which enables any IT site, including remote
office sites, to cost-effectively implement HA for both hosts and storage resources. In this way, StarWind
enables IT to address the needs of line of business (LoB) users by exporting logical disks with explicit
features needed by applications. Moreover, a StarWind Virtual SAN requires no capital expenditures
and avoids disruptive internal struggles between IT storage and networking support groups with simple
single-pane-of-glass management.

Scaling HA Storage for OLTP

We began by configuring three host servers in our Hyper-V VI as a 3-node high-availability (HA) StarWind Virtual SAN. On each of these servers, the StarWind Virtual SAN service ran directly on the host server. As a result, we were able to directly access Windows Server resources on the Hyper-V host without passing data through the Hyper-V hypervisor.

Next, we configured the two host servers running Windows Server 2012 R2 as a Hyper-V failover cluster within our Hyper-V VI. To enable failover in a 2-node cluster, it was essential to include a quorum disk that would function as an on-line disk witness. Consequently, we needed to provide a shared quorum disk that could remain available on-line in the event of a cluster host failure. By creating a quorum disk backed by a 3-node HA Virtual SAN, our quorum disk would remain available in the event of a failure of one or both of our cluster hosts.

Finally, we added a fourth StarWind Virtual SAN node using a VM that was running Windows 2012 R2 in a private vSphere 5.5 cloud. We used our expanded 4-node HA StarWind Virtual SAN configuration as a foundation to test the effectiveness of StarWind at scaling out HA support for VMs running in our Hyper-V failover cluster.

"Our 4-node HA Virtual SAN configuration, which was processing about 160 TPS with an average access time of 16ms, never incurred a performance penalty—manifested as either a lower TPS rate or higher average response time—while resources were restored when a server returned online after a storage or a host computing disruption."

HA in a Hyper-V Cluster
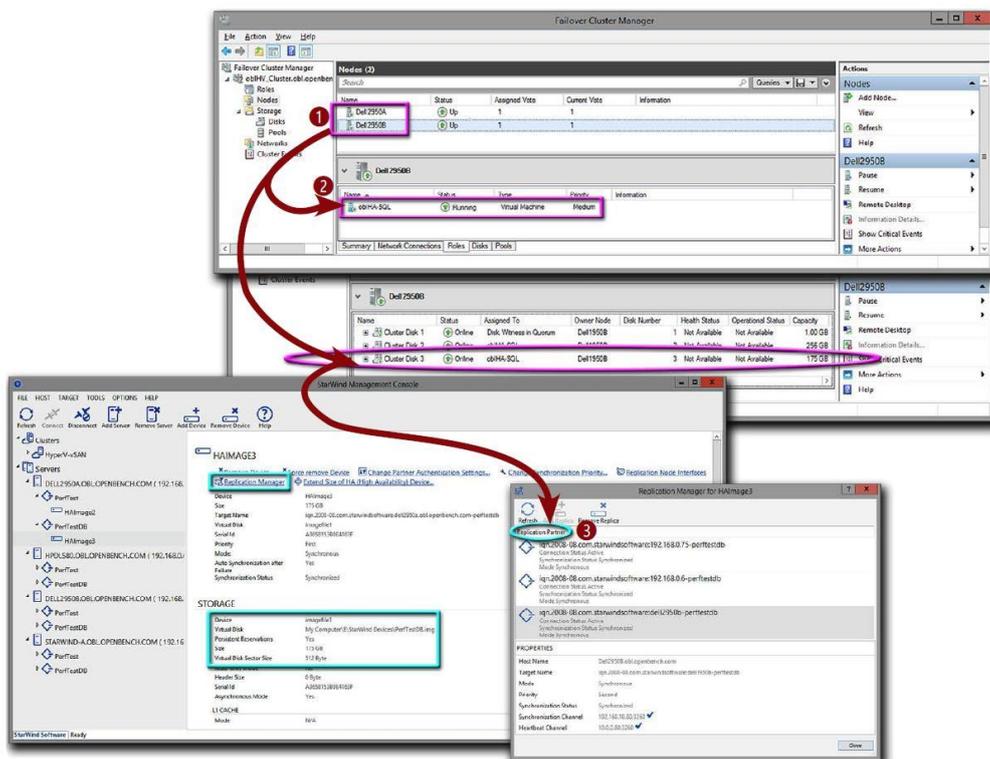
Virtual SAN & Cluster Configuration for OLTP

To support our Hyper-V test environment, we used two Dell PowerEdge R2950 servers running Windows Server 2012 R2 with Hyper-V along with an HP DL580 server running Windows 2012 and Hyper-V. We created a 2-node Hyper-V failover cluster using the two Dell PowerEdge R2950 servers❶ and a 3-node HA StarWind Virtual SAN to provide a shared quorum disk that would remain available on-line in the event of a cluster host failure or the need to synchronize the quorum disk on a cluster host. We then designated running of our test VM as a cluster role❷. For HA storage on the test VM running our OLTP application, we extended our HA Virtual SAN to four StarWind nodes, in which each virtual device file was synchronized with three partners❸ and then imported into the Hyper-V cluster and converted into CSV resources.

A 4-node configuration masks the impact of synchronizing a replica device from business users after a Virtual SAN device or node fails—much like a multi-spindle RAID-5 array can be rebuilt in a transparent background process after a disk crash. Our 4-node HA Virtual SAN configuration, which was processing about 160 TPS with an average access time of 16ms, incurred no performance penalty—manifested as either a lower TPS rate or higher average response time—while resources were restored when a server returned online after a storage or a host computing disruption.

Baseline StarWind Device Restoration
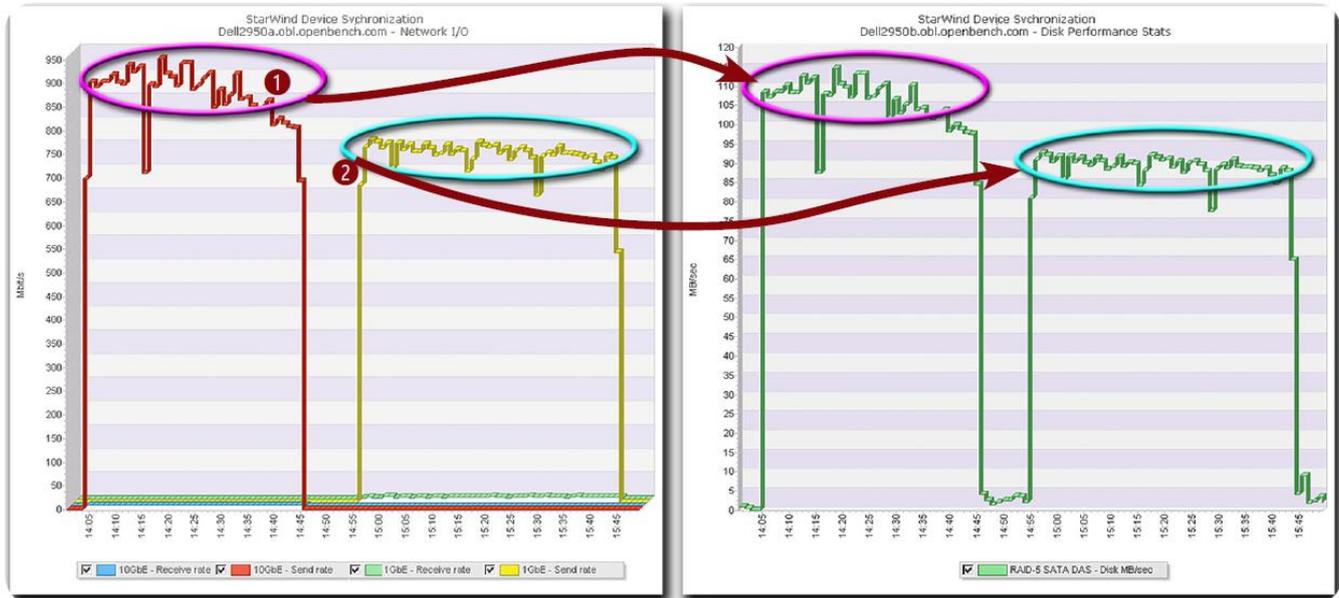
Virtual SAN Synchronization Performance

For users, HA technology is all about the handoff of devices during a failure; however, for IT operations, device handoff is just the start of a complex process with multiple dependencies. In terms of the time consumed and resources utilized, synchronization of a failed volume presents a significant potential for overhead degradation of business applications.

To maximize throughput of synchronization data, StarWind follows the best practices of backup packages by streaming data using large 256MB blocks. To leverage StarWind's large-block streaming, we used a key Hyper-V networking feature. For VM networking, Hyper-V utilizes virtual 10GbE switches, which can be linked to a physical host port and shared with the host. In that configuration, the host provisions a new virtual port based on the characteristics of the 10GbE virtual switch, rather than the characteristics of the physical NIC.

By leveraging a virtual 10GbE port to support the dedicated 1GbE NIC provisioned for synchronization traffic on each Hyper-V host, we were able to synchronize a 256GB device in about 40 minutes at 925 Mbps via a 10GbE network❶ and in just under 50 minutes at 750 Mbps via a 1GbE network❷. At the replication partner, synchronization data was being written to the local DAS RAID-5 array at averages of 110 MBps and 90 MBps respectively.

OLTP Stress Test

To model an OLTP application, we created a scenario based on the Transaction Processing Performance Council's TPC-E benchmark on a VM dubbed oblHA-SQL. In our business scenario, customers of a brokerage firm run SQL transactions to research market activity, check their account, and trade stocks, while brokers execute customer orders and update client accounts. To simulate external customers and internal traders, we used another VM to create and launch SQL queries based on ten templates, which were customized using a random number generator. We calculated an average execution rate in transactions per second (TPS) and measured an average response time for all queries.

For enhanced monitoring, we provisioned our test VM with two logical volumes which were located on independent StarWind virtual iSCSI drives. One volume contained Windows Server 2012 R2 and the SQL Server 2014 application, while the other volume contained the working 24GB database tables and logs.

HA Virtual SAN Crash Dynamics
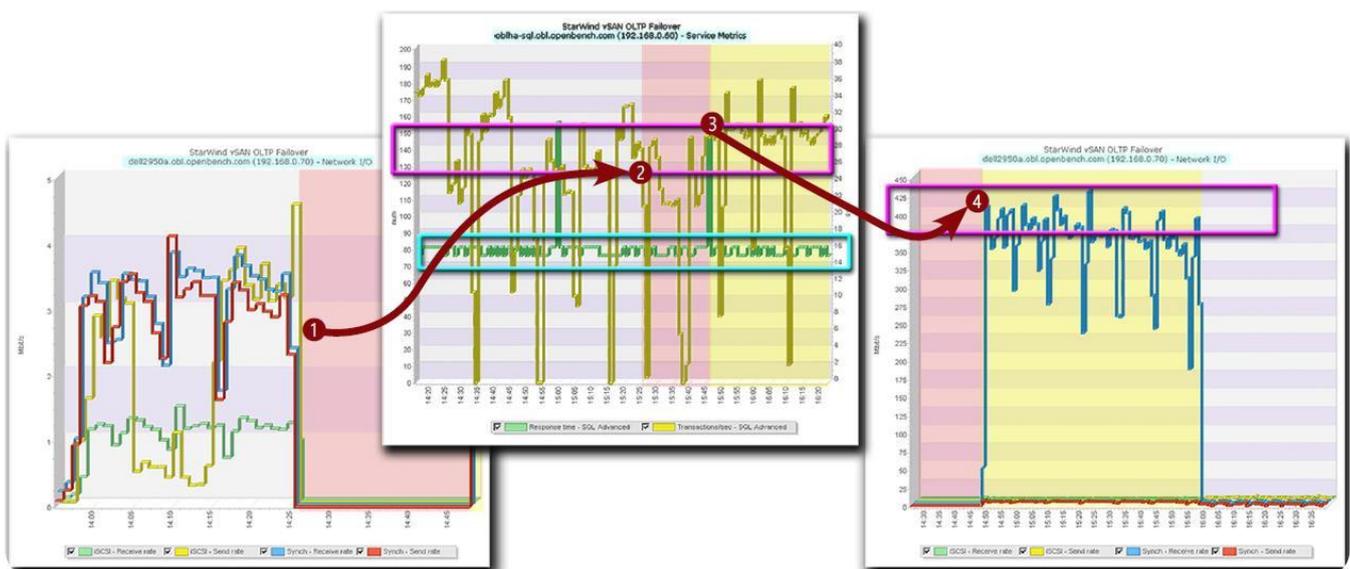
Anatomy of a Hyper-V Virtual SAN Node Crash

Crash!

It's 2:26 pm, and the Hyper-V server, Dell2950A, which exports the iSCSI devices that support the company's critical OLTP business application, just crashed❶. For LoB users and executives, however, business processing continues as normal: The SQL Server on oblHA-SQL maintains transaction processing❷ at approximately 145 TPS—a drop of 10 percent—and users continue to experience an average response time of 16ms.

Synchronize

With the cause of the server failure resolved, IT operations restarts the failed server❸ 21 minutes after the outage occurred. Since SQL Server transaction processing has continued over the period that the server was inactive, more data was processed than could be cached. Consequently, once the StarWind Virtual SAN service sensed the failed server had returned on-line, the service started a full synchronization❹ of the device images on Dell2950A at about 400Mbps and processed client traffic for the OLTP test load at 160 TPS—the processing rate before failure—during the synchronization process.

Full Recovery

Full synchronization, which took place in tandem with continued processing of the OLTP benchmark, took 73 minutes to complete, as soon as the failed server had been brought back on-line. At that point, all four HA Virtual SAN nodes were fully synchronized and servicing the VM under test.

More importantly, with a 4-node HA Virtual SAN, StarWind throttled synchronization processing by approximately 60 percent, to avoid negatively impacting application processing. This well-balanced processing occurred with the default synchronization setting, which puts a relatively high priority on synchronization.

Moreover, with a 4-node SAN, the loss of one node does not trigger automatic changes in the Virtual SAN topology. In a 2-node Virtual SAN, which is the smallest configuration for HA support, any write-back device cache—the default cache configuration— is switched immediately to a write-through cache and a mandatory cache flush is processed.

What's more, StarWind further minimizes the impact of synchronization after a failure by first attempting to perform a cache-based fast synchronization. Fast synchronization utilizes an internal cache to track the iSCSI data changes after a device fails. If the failed device returns on-line before the cache is flushed, StarWind limits data synchronization to just the data changed while the device was off-line.

HA Hyper-V Cluster Dynamics

Using Hyper-V live migration, IT administrators are able to move VMs between hosts in a Hyper-V cluster with minimal service interruptions. In particular, an IT administrator can invoke live migration via the Failover Cluster Manager, PowerShell, or the Microsoft System Center Virtual Machine Manager.

The live migration process begins with the source server creating a connection with the destination server and transferring VM configuration data to the destination server. Using the configuration data, a skeleton VM is set up on the destination server and memory is allocated to the skeleton VM.

If there are sufficient memory resources on the destination host, utilized memory pages are transferred from the source node to the destination. During this process, the Hyper-V service tracks memory pages on the source for page changes and then iterates the transfer of modified pages. This iterative memory transfer is a key determining factor in the time to complete the complete live migration.

With the memory pages transferred, the source server transfers the VM's CPU state, device states, and ownership of storage resources required by the VM to the new host. Once the control of storage resources is complete, the destination host boots the VM and cleans up network settings by changing the MAC address for port on the virtual switch used by the migrated VM to access the network.

Leveraging Live Migration
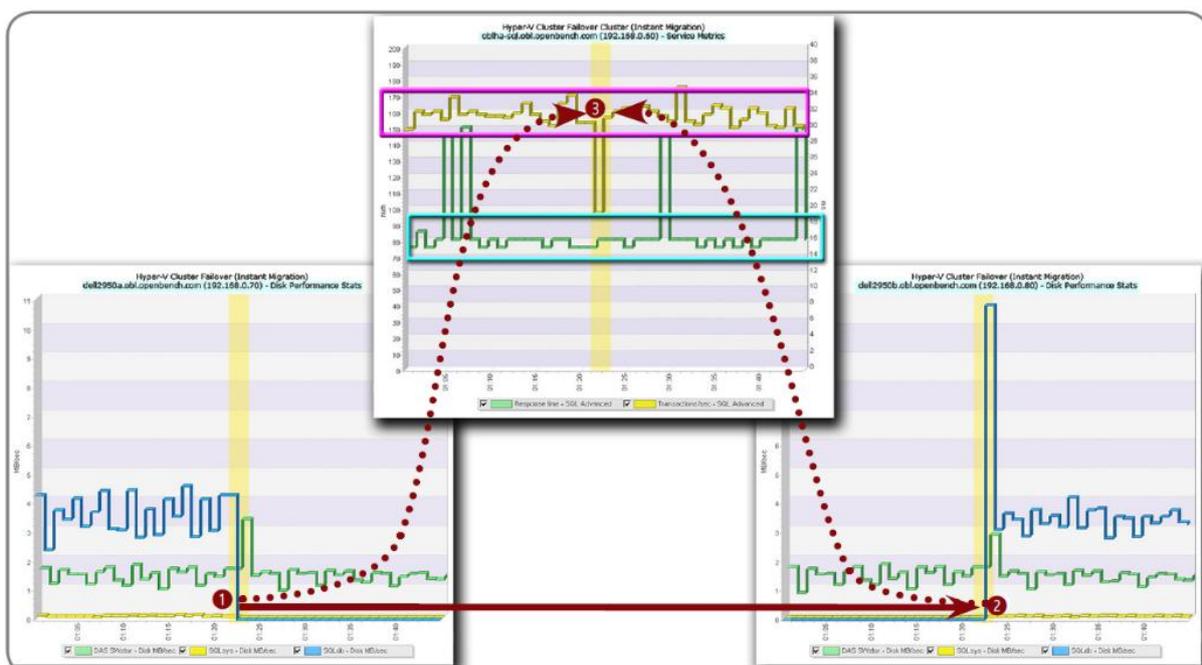
HA Virtual SAN Performance with Live Migration

To test HA storage with Hyper-V Live Motion, we moved our test VM, oblHA-SQL, between two Hyper-V cluster hosts, Dell2950A and Dell2950B, while our OLTP business application was running. Using live migration with memory page compression and our StarWind logical volumes imported into our cluster as CSV resources, oblHA-SQL was incapable of receiving queries over the network during the transition from the host Dell2950A❶ to the host Dell2950B❷ for less than one second.

As we moved our test VM, oblHA-SQL, the SQL Server database continued to process around 160 TPS❸, only one transaction generated on the VM executing our SQL query test script triggered a query exception error. More importantly, neither of the StarWind Virtual SAN servers involved in the Live Motion transition needed to invoke a fast synchronization after the transition process completed.

For IT, unleashing the real power of live migration requires utilization of System Center Virtual Machine Manager to enable:

- Dynamic Optimization to load balance virtual machines across hosts within a Hyper-V cluster. It dynamically places virtual machines on the host with the most suitable available resources;

- Performance and Resource Optimization to drain VMs from a Hyper-V host before performing host maintenance and move VMs on clustered hosts after detecting performance or resource issues.

To facilitate the use of System Center, StarWind provides an integration module for managing an HA Virtual SAN with the System Center console.

Application Continuity and Business Value

For CIOs, the top-of-mind issue is how to reduce the cost of IT operations. With storage volume the biggest cost driver for IT, storage management is directly in the IT spotlight.

HA StarWind Virtual SAN Feature Benefits	
1	**StarWind Reduces Capital Expenses associated With Storage Provisioning:** Starwind Virtual SAN service enables SATA and SAS DAS storage to be shared in an iSCSI Virtual SAN without requiring the use of Flash Storage as a Level-1 cache.
2	**StarWind Reduces Capital Expenses associated With Server Provisioning:** light-weight StarWind Virtual SAN service runs directly on Windows host servers, which enables IT to retask older servers to scale out storage HA features.
3	**StarWind Reduces VI Management Overhead:** By decoupling iSCSI devices and targets, StarWind enables device reconfiguration without impacting iSCSI target sessions. In particular, HA support levels can be reconfigured without involving a Hyper-V host.
4	**Simplifies Virtual SAN Scale Out:** By virtualizing HA support as a device-specific function, we were able to scale out multiple levels of HA support as a datastore-specific feature.

At most small-to-medium business (SMB) sites today, IT has multiple vendor storage arrays with similar functions and unique management requirements. From an IT operations perspective, multiple arrays with multiple management systems forces IT administrators to develop many sets of unrelated skills.

Worse yet, automating IT operations based on proprietary functions, may leave IT unable to move data from system to system. Tying data to the functions on one storage system simplifies management at the cost of reduced purchase options and vendor choices.

By building on multiple virtualization constructs a StarWind Virtual SAN service is able to take full control over physical storage devices. In doing so, StarWind provides storage administrators with all of the tools needed to automate key storage functions, including as thin provisioning, data de-duplication and disk replication, in addition to providing multiple levels of HA functionality to support business continuity.

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

Jack Fegreus

openBench Labs

Web-site: www.openbench.com   Tel: (508) 329-1645
E-mail: jack.fegreus@openbench.com   Fax: (508) 329-1640

StarWind

US Headquarters	EMEA and APAC
+1-617-4497717 +1-617-507-5845	+44 20 3769 1857 (UK) +49 302 1788 849 (Germany) +33 097 7197 857 (France) +7 495 975 94 39 (Russian Federation and CIS) 1-866-790-2646

Support Portal: https://www.starwind.com/support   Sales: sales@starwind.com
Support Forum: https://www.starwind.com/forums   General Information: info@starwind.com

≈ **StarWind**

StarWind Software, Inc. 35 Village Rd., Suite 100, Middleton, MA 01949
©2015, StarWind Software Inc. All rights reserved.