

OUTDATED

Exchange Failover in a StarWind Virtual SAN using a vSphere VI with Virtualized iSCSI Devices on DAS

Cost Effective HA Storage for Critical Business Applications

FEBRUARY 2015

TECHNICAL PAPER BY DR. JACK FEGREUS



Trademarks

"StarWind", "StarWind Software" and the StarWind and the StarWind Software logos are registered trademarks of StarWind Software. "StarWind LSF5" 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](#).

Copyright ©2009-2015 StarWind Software Inc.

No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior written consent of StarWind Software.

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

Snapshot of Findings	4
Driving HA Through Virtualization	5
<i>Capital Cost Avoidance</i>	5
<i>Building Exchange HA on Virtualized DAS Devices</i>	6
Analyzing HA Data Traffic Patterns.....	7
<i>Maximizing iSCSI Throughput via Path Switching</i>	7
<i>Moving to a 3-Node Virtual SAN</i>	8
iSCSI I/O Virtualization	9
Direct Virtual SAN Backup.....	10
Enterprise Email Stress Test	11
<i>Multi-Dimensional HA Recovery Dependencies</i>	11
StarWind HA Dynamics	12
<i>Crash!</i>	12
<i>Synchronize</i>	12
<i>Full Recovery</i>	12
<i>Fast Virtual SAN Synchronization</i>	13
<i>Full Virtual SAN Synchronization</i>	14
Application Continuity and Business Value	15

“Using StarWind Virtual SAN software, IT is able to transform internal direct attached storage (DAS) on a server into a shared storage resource, an essential prerequisite for providing continuous access to data, eliminating single points of failure, and enabling VM migration via vSphere vMotion®.”

Snapshot of Findings

This document is for advanced StarWind users who seek to deepen their practical skills and develop a keen understanding of the latest virtualization trends. It gives an insight into particular High Availability configurations and their balance between performance and cost-effectiveness. In addition, the technical paper is completely practical and every piece of information is derived from real-life application of the mentioned technology.

In this analysis, openBench Labs assesses the High Availability (HA) performance of a StarWind Virtual Storage Area Network (Virtual SAN or Virtual SAN) in a VMware® vSphere™ 5.5 HA Cluster. For our test bed, we configured an HA Virtual SAN using three StarWind virtual machines (VMs) as a foundation for a 2-node vSphere cluster in a VMware vCenter™ with automated vMotion® failover. With this layered HA configuration, we tested the ability to scale-out the performance of HA storage features using a common IT scenario: an email service configured with Exchange 2010 running on a VM server and a primary domain controller (PDC) supporting Active Directory (AD) for users of Exchange on another VM server.

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. Using two, three, or four StarWind servers, IT is able to configure storage platforms with differing levels of HA support. Furthermore, IT is able to provide a near fault tolerant (FT) environment for key applications by layering HA systems support on an HA-storage foundation.

To demonstrate performance scalability, we created 2,000 AD users for our Exchange service and utilized LoadGen to generate email traffic. We assigned each user a LoadGen profile that automated receiving 80 messages and sending 13 messages over an 8-hour workday. Using this load level, we established performance baselines for a data protection with StarWind Virtual SAN support for storage, vSphere HA support for VM execution, and direct SAN-based agentless VM backups.

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).

StarWind Virtual SAN Software		
HA Protection for a vSphere VI - DR Test Assessment		
HA Failure Simulation	LoB User Results	IT Administrator Results
StarWind VM Server Failure	No Interruption in Exchange Processing	<ul style="list-style-type: none"> StarWind VM Server Restored in 3 Minutes Full Virtual SAN Synchronization of 3 Datastores Restored in 5 Minutes via Fast Synchronization, while Processing 9 Outlook TPS
vSphere Host Failure	Exchange Processing Restored in 3 Minutes	<ul style="list-style-type: none"> ESX Host Restored in 3 Minutes Full Virtual SAN Synchronization of 3 Datastores Restored in 186 Minutes via Full Synchronization, while Processing 9 Outlook TPS

Driving HA Through Virtualization

Using StarWind Virtual SAN software, IT is able to transform internal direct attached storage (DAS) on a server into a shared storage resource, which is an essential prerequisite for providing continuous access to data, eliminating single points of failure, and enabling VM migration via vSphere vMotion®. By decoupling iSCSI targets from underlying devices, StarWind enables IT to meet explicit application needs by exporting datastores with different support features on virtual devices created with NTFS files on standard storage resources, including DAS.

A StarWind Virtual SAN also provides critical heartbeat and synchronization services to iSCSI targets without requiring special hardware, such as flash devices. Moreover, a StarWind Virtual SAN enables any IT site, including remote office sites, to cost- effectively implement HA for both the hosts and storage supporting VMs running business-critical applications.

Capital Cost Avoidance

StarWind is seeking to change the old notion that high-performance shared storage requires an expensive hardware- dependent SAN to provide enhanced HA capabilities. In particular, 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 within VMware vCenter™ as part of virtual machine creation and deployment.

Building on a mature iSCSI software foundation, StarWind uses advanced virtualization techniques to decouple iSCSI targets from the storage devices that they represent. By decoupling iSCSI targets from devices, StarWind virtualizes distinct device types, which enable IT administrators to configure ESX datastores that have specialized features required by applications on VMs provisioned with logical disks located on these datastores.

Building Exchange HA on Virtualized DAS Devices

By decoupling iSCSI targets from underlying devices, StarWind enables IT to meet explicit application needs by exporting datastores with different support features on virtual devices created with NTFS files on standard storage resources, including DAS. A StarWind Virtual SAN also provides critical heartbeat and synchronization services to iSCSI targets without requiring special hardware, such as flash devices. A StarWind Virtual SAN enables any IT site, including remote office sites, to cost-effectively implement HA for both the hosts and storage supporting VMs running business-critical applications.

Today, email handles over 90 percent of business communications, which is why email is no longer just a business-critical application. Today, email is legally classified as financial data and presents IT with unique data protection issues. A disaster recovery (DR) plan for an email server can involve numerous regulations, including the Federal Rules of Civil Procedures.

Given all of the LoB stakeholders involved with email, an SLA for Exchange will need to address a number of key issues. For business continuity, an Exchange SLA will require IT to provide data protection support that minimizes both down time and data loss in the event of a processing outage. As a result, any SLA for Exchange data protection will need to meet a minimum recovery time and recovery point objective (RTO and RPO). Moreover, frequent incremental backups to meet an aggressive RPO goal fits well into a near-FT environment built on a HA StarWind Virtual SAN, since reading data during a backup is transparent with respect to data synchronization.

“IT administrators are able to set up multiple HA datastore levels using StarWind configured with a common set of Virtual SAN servers in a mix of replication schemes.”

Analyzing HA Data Traffic Patterns

To gain an insight into the workings of an HA StarWind Virtual SAN cluster, we ran two I/O tests on a VM file server running Windows Server 2012R2 and represented by 105GB of VMFS data. For testing, we created two 256GB disk image devices on our StarWind VM dubbed StarWind-A. We designated one device as a baseline single-node datastore, and replicated the other device on the HA Virtual SAN servers dubbed StarWind-B and StarWind-C.

With StarWind devices decoupled from targets, replication in an HA Virtual SAN is device- rather than server-specific. Any device on a StarWind Virtual SAN server can be replicated on up to four HA StarWind Virtual SAN members. Consequently, IT administrators are able to support multiple HA datastore levels with StarWind using a common set of Virtual SAN servers configured with a mix of virtual iSCSI devices and replication schemes.

On our VI datacenter, we deployed two logical iSCSI datastores associated with devices on 3-node and 1-node Virtual SANs to support a file server VM. We also imported all of the Virtual SAN datastores on an independent Windows backup server to test LAN-free VM backup throughput. More importantly, we set the MPIO policy on all of the servers importing Virtual SAN iSCSI targets—both single and multi-node configurations—to round robin, which transmits a fixed amount of data while rotating device paths.

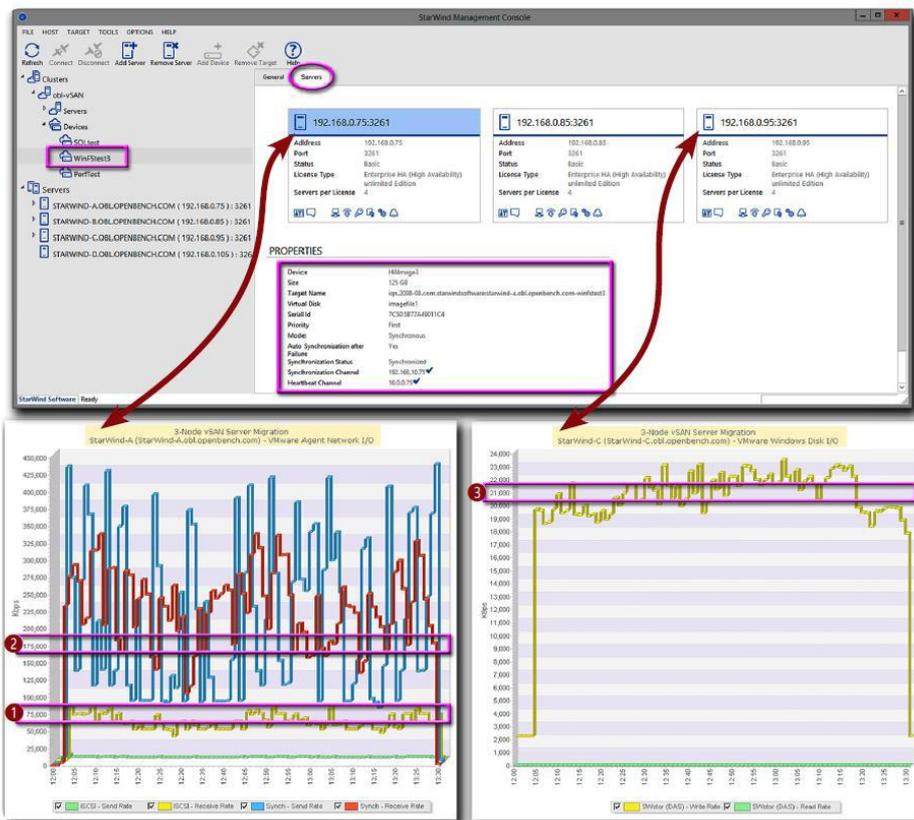
Maximizing iSCSI Throughput via Path Switching

In a typical high-bandwidth 8Gbps Fibre Channel SAN, maximizing I/O throughput is an HBA issue rather than a path bandwidth issue, as it is in a 1GbE iSCSI SAN. Consequently, the simple notion of SAN device paths as either active or passive has been replaced with optimized active, unoptimized active, and passive paths and asymmetric logical unit access (ALUA) MPIO policies have been introduced to maintain SAN connectivity over optimized device paths using a Most Recently Used (MRU) algorithm.

Since a StarWind Virtual SAN typically runs in a 1GbE environment with multiple iSCSI paths, network load balancing, rather than HBA switching overhead, is the critical success factor. Consequently, StarWind publishes every device path as optimized to maximize the availability of all iSCSI device paths. While StarWind maximizes device path availability, the ALUA MPIO policy on every server importing Virtual SAN devices controls how those devices are accessed. Even with all StarWind paths characterized as optimized, an ALUA MPIO policy based on an MRU algorithm is equivalent to a static fixed path policy. An MPIO policy that is biased towards path switching, such as round robin, is essential for leveraging all of the possible paths to datastores exported by a HA StarWind Virtual SAN.

“StarWind minimizes the impact of HA Virtual SAN scale-out on VM applications by isolating iSCSI and synchronization traffic, holding iSCSI data traffic constant, and varying only the volume of synchronization data—no matter the number of Virtual SAN nodes.”

Moving to a 3-Node Virtual SAN



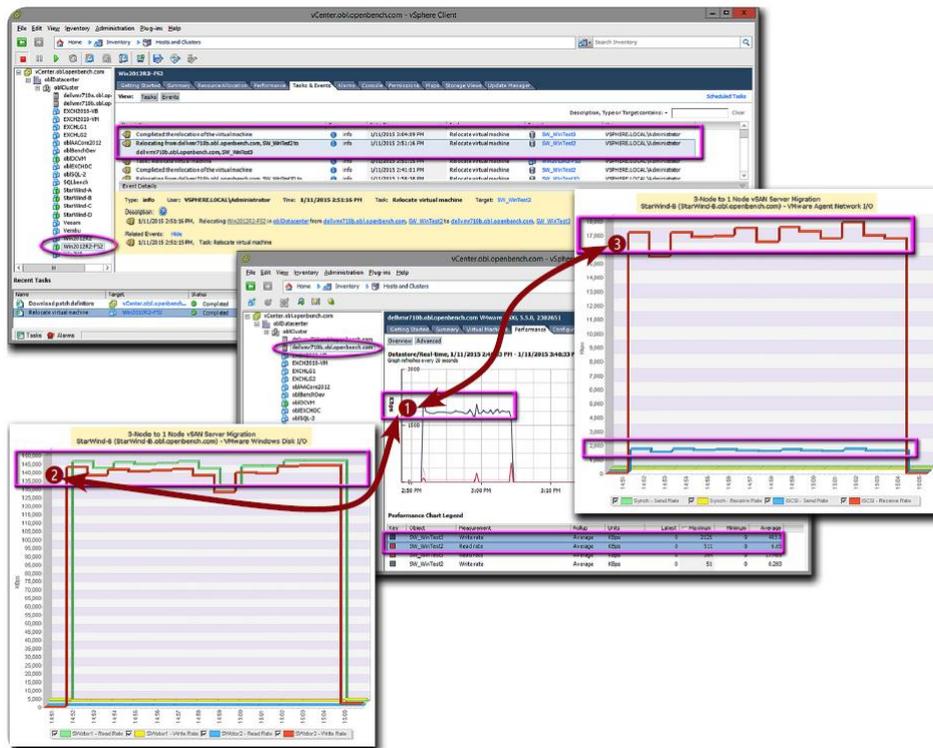
Virtual SAN Migration Throughput

We used storage vMotion to deploy the file server VM on the 3-node HA Virtual SAN datastore from an FC SAN. While migrating the VM file server, each StarWind server in our 3-node HA Virtual SAN received approximately 35GB of the 105GB of iSCSI data¹ directed at the logical destination datastore by the ESX host, which was transmitting data at 72 megabits per second (Mbps). During this process, each HA Virtual SAN node sent a copy of its iSCSI data to each of its two partners as synchronization data² traffic at 150 Mbps.

On each node, the StarWind service combined iSCSI and synchronization data to write a full copy of the VM file server data³ at 21 megabytes per second (MBps) on the device file representing the datastore. Moreover, StarWind minimizes the impact of HA Virtual SAN scale-out on VM applications by isolating iSCSI and synchronization traffic, holding iSCSI data traffic constant, and varying only the volume of synchronization data—no matter the number of Virtual SAN nodes.

"A StarWind service is able virtualize and accelerate iSCSI traffic between datastores that share a common StarWind Virtual SAN server."

iSCSI I/O Virtualization



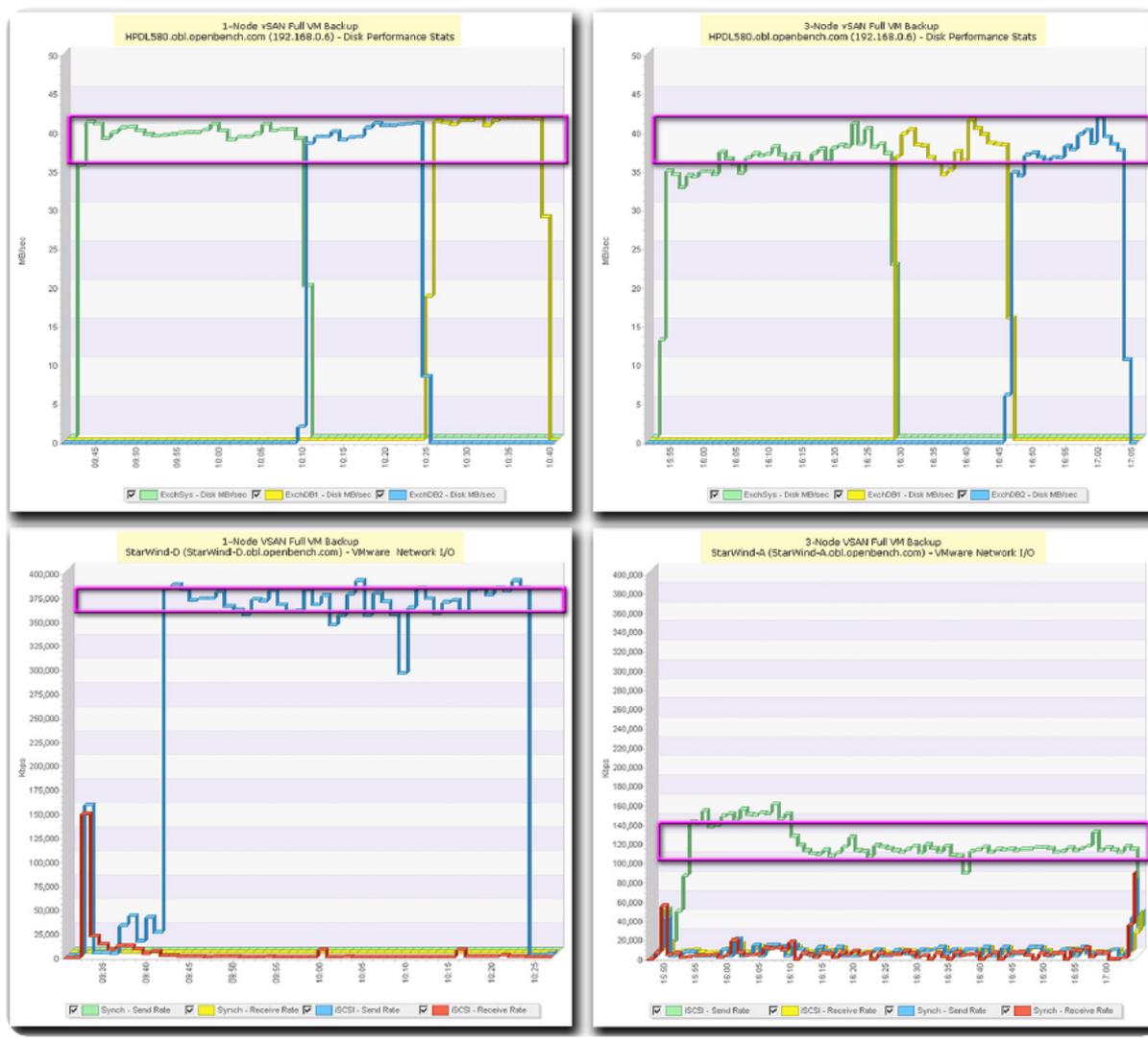
Virtualization of iSCSI Network Traffic

To efficiently read and write VM files a StarWind service accesses VM files in a logical device image, much like a VM backup package accesses VM files in a backup container. As a result, a StarWind service is able virtualize and accelerate iSCSI traffic between datastores that share a common StarWind Virtual SAN server for particular vSphere functions.

To assess iSCSI virtualization, we used vMotion to move the file server VM from the datastore on the 3-node HA Virtual SAN to a datastore on a 1-node Virtual SAN, which was exported by StarWind-B, a member of the 3-node HA Virtual SAN. Consequently, the StarWind service on StarWind-B had direct access to the virtual device files representing both the originating and destination datastores.

To virtualize vMotion data flow, the StarWind service on StarWind-B transferred 105 GB of data **1** in 13 minutes by copying VM files from the device file representing the origin datastore to the device file representing the destination datastore at 140 MBps **2**, compared to 20 MBps for a typical VM migration. To satisfy the VM migration process, StarWind-B streamed token iSCSI data to the ESX host at 16.5 Mbps **3**, which vCenter reported as 2 MBps datastore throughput.

Direct Virtual SAN Backup



LAN-Free Virtual SAN Backup Performance

To test backup throughput for our Exchange VM server, we ran LAN-free backups of the Exchange VM on datastores provisioned on a 1-node and a 3-node Virtual SAN. For an IT administrator focused solely on running a backup, the migration of the file server VM from a basic 1-node Virtual SAN with no HA protection to a 3-node HA Virtual SAN was completely transparent with respect to the backup process. In both cases, LAN-free backup throughput on the backup server averaged 37 MBps.

From the perspective of a network administrator, however, iSCSI data traffic during the backup process was significantly different. The default ESX MPIO policy of round robin access automatically triggered the sending of fixed data traffic amounts over each I/O path in a rapid succession. Consequently, each of the three StarWind Virtual SAN servers streamed iSCSI data to the backup server around 120 Mbps. Moreover, since our backup involved data reads with no writes, the StarWind service sent no synchronization data to any HA Virtual SAN node.

“Network load balancing provided by StarWind Virtual SAN path optimization and a round robin client MPIO policy directly benefits application I/O responsiveness by spreading I/O operations over multiple network paths and datastores residing on independent storage arrays.”

Enterprise Email Stress Test

To test the ability of HA StarWind Virtual SAN to provide near-FT protection for a business-critical application, we configured an Exchange email service using two VMs: one VM ran Exchange 2010 sp3, while the other acted as a Primary Domain Controller (PDC) handling 2,000 user accounts in Active Directory (AD).

To generate an email transaction load for the Exchange service that would stress the underlying StarWind Virtual SAN, we added two VMs to the VI datacenter in order to run the LoadGen benchmark. On these servers, we generated nine Outlook transactions per second (TPS) for the 2,000 user accounts created in AD for email performance testing. This scenario executed about 32,700 Outlook transactions on the Exchange service every hour.

To improve disk I/O performance when executing this transaction processing load, we configured the Exchange vSphere VM with three logical disks: one volume for system files and two volumes for mailbox database and log files. To leverage this storage topology, we created two mailbox databases with 1,000 accounts per database.

Next, we assigned each user account with a work profile calling for the user to invoke an Outlook client to generate 131 email transactions over an 8-hour workday. These transactions fell into four major frequency-based categories: each user received and read 80 messages, composed and sent 13 messages, browsed calendars 12 times, and browsed contact lists 9 times, throughout the work day. Furthermore, with respect to IT backup operations, this scenario resulted in approximately 5GB of Exchange block data being flagged as changed as data was added or modified in the database and log files of the Exchange server every 15 minutes.

Multi-Dimensional HA Recovery Dependencies

To determine an optimal HA StarWind Virtual SAN configuration strategy for a business-critical application on a VM, a number of critical factors come into play. Key factors include the volume of data transactions that write or modify data, the amount of device cache available, and the size of the logical volumes provisioned for a protected VM. All of these factors are critical for the most operationally costly phase of an HA failover: the recovery of a protected VM to full production status via complete data synchronization on the StarWind Virtual SAN.

Our backup tests demonstrated, application transactions that simply retrieve data do not trigger data synchronization functions. In terms of StarWind processing, performing I/O reads incurs negligible costs.

Nonetheless, this does not imply that there is no benefit garnered for read transactions from the HA StarWind Virtual SAN optimization scheme. Network load balancing provided by StarWind Virtual SAN path optimization and a round robin client MPIO policy directly benefits application I/O responsiveness by spreading I/O operations over multiple network paths and datastores residing on independent storage arrays.

StarWind HA Dynamics

Crash!



It's 11:23 am. For 2,000 active clients, a storage array supporting the mailbox database for their account on an Exchange server with just crashed. What happens next?

To assess the resilience of a HA StarWind Virtual SAN, we needed to test performance at a point of failure. While processing Outlook transactions at 9 TPS, our standard production load, we forced the StarWind-C VM, to reboot.

What happened next:

- From the perspective of 2,000 Exchange users, nothing happened. Exchange continued to run with no change in the way email services were running.
- From the perspective of IT an backup operator, nothing happened. The mailbox database backup that started at 11:20 completed successfully.

Synchronize



For users, HA technology is all about the handoff of devices at the time of failure. For IT administrators, device handoff represents only the start of a complex process.

Within two minutes, StarWind-C had rebooted. By 11:26, the iSCSI SAN service on StarWind-C automatically initiated a fast synchronization, which limits synchronization to the data changed since the VM was last in a synchronized state.

Full Recovery



The final and least predictable stage in an HA failure scenario is the restoration of all StarWind Virtual SAN servers to full operating status. Full recovery occurs once all of the StarWind Virtual SAN servers are synchronized and exclusively processing current data changes.

For IT administrators, the major questions are:
How long does synchronization take?
How much overhead does synchronization introduce?

As StarWind-C came on-line, fast synchronization took less than three minutes to update the Mailbox database and log files stored on device images maintained by StarWind-C with changes created by about 1,600 Outlook transactions executed while StarWind-C was off-line.

Fast Virtual SAN Synchronization



HA Fast Synchronization

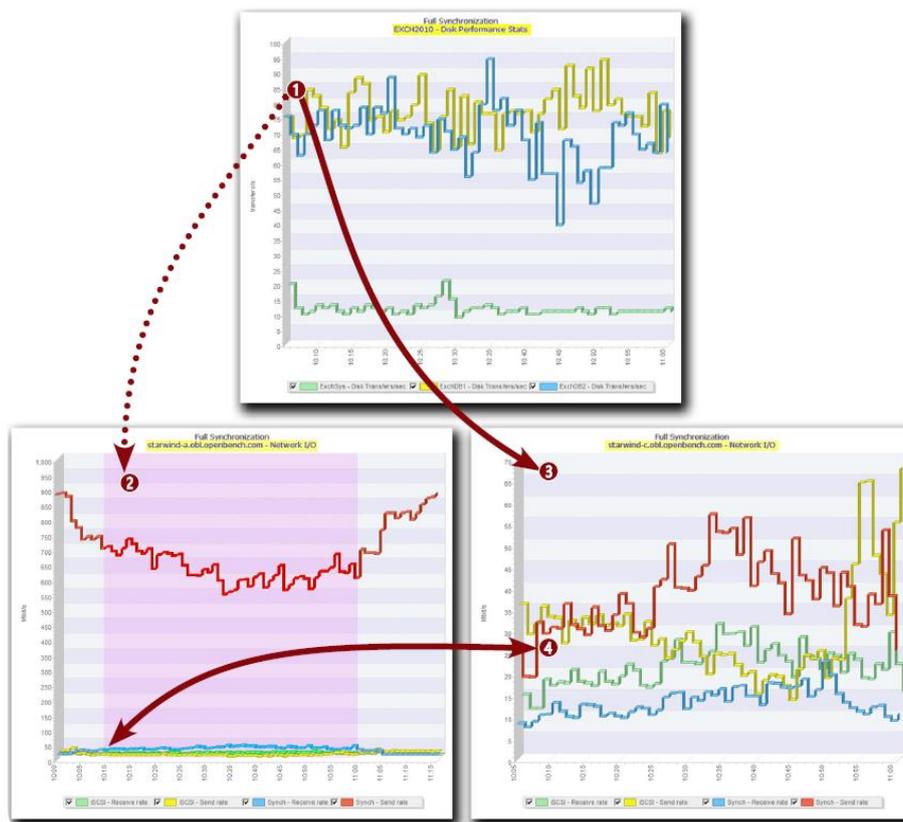
The loss of an HA Virtual SAN node triggers a number of background configuration changes. In particular, StarWind begins tracking the percentage of data blocks changed on each device. When the percentage reaches an internal parameter, StarWind forces a full rather than a fast synchronization. Consequently, the probability of recovering with fast synchronization is biased towards large devices.

When we crashed StarWind-C while running our LoadGen transaction stream **1**, client MPIO policies redirected iSCSI connections to the remaining two nodes supporting the Exchange VM, which continued processing Outlook transactions without interruption. Within three minutes of rebooting, StarWind-C initiated a fast synchronization **2** to restore missed data transactions. Fast synchronization took 1 minute and 58 seconds. Consequently, our Virtual SAN recovered from a reboot of StarWind-C within five minutes, while the Exchange VM continued processing Outlook email transactions.

Fast synchronization, which limits synchronization to known changed blocks, is the most visible benefit of using a 3-node or 4-node HA Virtual SAN. In a 2-node HA Virtual SAN, the loss of a node causes StarWind to issue a mandatory cache flush and switch all write-back device caches to write-through. Consequently a full synchronization will be required.

“For an Exchange SLA, reconfiguration and full synchronization of a 3-node StarWind Virtual SAN was entirely transparent to users following the crash of an ESX cluster host.”

Full Virtual SAN Synchronization



HA Full Synchronization of a Datastore with a Mailbox Database

Next, we crashed the vSphere cluster host that was running the Exchange and StarWind-B VMs. The remaining cluster host rebooted the Exchange VM in about out three minutes and begin processing Outlook transactions using the Virtual SAN devices on the remaining two StarWind VM servers. Until we recovered the failed cluster host, we could not restart StarWind-B, which was dependent on the failed host's DAS resources. The time elapsed restarting the host and StarWind-B led to full synchronization of all three Exchange datastores, which took three hours and six minutes.

To enable the Exchange VM **1** to continue running 9 Outlook TPS during full synchronization of StarWind-B, the StarWind Virtual SAN reconfigured device ALUA properties making iSCSI targets inactive on StarWind-B and available only for failover on StarWind-A. Consequently, StarWind-A handled StarWind-B synchronization **2** while StarWind-C handled processing of client data **3**. StarWind-A also incorporated synchronization data from StarWind-C **4** while synchronizing StarWind-B to account for on-going Exchange processing. More importantly, for an Exchange SLA, reconfiguration and full synchronization of a 3-node StarWind Virtual SAN was entirely transparent to users following the crash of an ESX cluster host.

“For any business continuity SLA for Exchange, reconfiguration and full synchronization of a 3-node StarWind Virtual SAN was entirely transparent to LoB users.”

Application Continuity and Business Value

HA StarWind Virtual SAN Feature Benefits	
1	Reduces Capital Expenses associated With Storage Provisioning: A StarWind Virtual SAN 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	Minimizes Capital Expenses associated With Server Provisioning: Light-weighted StarWind Virtual SAN can be run on older ESX hosts to scale out storage HA features.
3	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 vCenter host.
4	Automates Network Load Balancing: By configuring access to all targets in an HA Virtual SAN as optimized, a StarWind Virtual SAN promotes path switching to balance traffic during normal operations and modifies access when recovering logical devices via synchronization following a server failure.
5	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.

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.

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

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.

openBench Labs

Web-site: www.openbench.com

Tel: (508) 329-1645

E-mail: jack.fegreus@openbench.com

Fax: (508) 329-1640

StarWind Software

US Headquarters

 +1-617-4497717
 +1-617-507-5845

EMEA and APAC

 +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: <http://www.starwind.com/support>

Sales: sales@starwind.com

Support Forum: <http://www.starwind.com/forums>

General Information: info@starwind.com



www.starwind.com **StarWind Software, Inc.** 35 Village Rd. Suite 100 Middleton, MA 01949 USA
©2015, StarWind Software Inc. All rights reserved.