

EMC[®] VNXe[®] HIGH AVAILABILITY

Overview

Abstract

This white paper discusses the high availability (HA) features in the EMC[®] VNXe[®] system and how you can configure a VNXe system to achieve your goals for data availability in your environment. Some concepts discussed in this paper do not apply to single-SP VNXe3100/VNXe3150 systems.

January 2013

Copyright © 2012 EMC Corporation. All rights reserved.
Published in the USA.

Published December 2012

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

The information in this publication is provided as is. EMC Corporation makes no representations or warranties of any kind with respect to the information in this publication, and specifically disclaims implied warranties of merchantability or fitness for a particular purpose. Use, copying, and distribution of any EMC software described in this publication requires an applicable software license.

EMC2, EMC, and the EMC logo are registered trademarks or trademarks of EMC Corporation in the United States and other countries. All other trademarks used herein are the property of their respective owners.

For the most up-to-date regulatory document for your product line, go to the technical documentation and advisories section on the EMC online support website.

Part Number h8276.4

Table of Contents

Executive summary	4
Audience.....	4
Terminology	4
High availability in the storage system	5
Storage processor	5
Failover example	6
Failback of storage resources	7
Failover of the management interface.....	8
Power Modules.....	8
Memory.....	9
Write cache	9
Cache protection module in the single-SP VNXe 3100/3150 systems.....	10
Data availability in your connectivity infrastructure	10
Fail-Safe Networking.....	11
Implementation for iSCSI storage	11
Implementation for NFS storage	16
High availability in disk storage	21
Number of backend buses.....	21
RAID configurations.....	21
RAID 5	22
RAID 6	23
RAID 10	23
Dual drive ownership.....	23
Hot sparing.....	23
Using replication for high availability	24
Continuous Availability	24
Conclusion	25
References	26

Executive summary

IT administrators are concerned constantly about the accessibility of their organization's business data. EMC® focuses on helping customers maintain the most highly available storage environment possible. Having a highly available environment involves many factors, including the environment's architecture and design, and how you configure your infrastructure connectivity, storage system, and disk storage. Adding replication to your environment also enhances availability.

A highly available storage system does not have a single point of failure; that is, if a component fails, the system maintains its basic functionality. For example, having two storage processors (SPs) in your VNXe® system allows you to configure alternate data paths at the storage system level. This allows the system to remain operational if a component in one path fails.

In many cases, a highly available system can withstand multiple failures, if the failures occur in different component sets. Once detected, failed components in the VNXe system can be replaced easily and brought online without affecting users or applications.

This document discusses different factors that affect availability. It also discusses how to design your VNXe system to achieve the level of availability that your environment requires.

Audience

This white paper is intended for EMC customers, partners, and employees who want to understand the features in the EMC VNXe product that can maximize data availability. Some details in this white paper may not be applicable to customers with single-SP VNXe3100/3150 systems.

Terminology

Fail-Safe Networking (FSN)—A high-availability feature that extends link failover into the network by providing port-level redundancy.

iSCSI Server—A VNXe server that uses the iSCSI protocol to manage Microsoft Exchange storage groups, generic storage virtual disks, Hyper-V datastores, and VMFS-based VMware datastores.

Link aggregation—A high-availability feature based on the IEEE 802.3ad Link Aggregation Control Protocol (LACP) standard that allows Ethernet ports with similar characteristics to connect to the same switch (physical or cross-stack). When aggregated in this way, the ports combine to make a single virtual link with a single MAC address. This virtual link can have multiple IP addresses.

Shared Folder Server—A VNXe server that uses either the CIFS or NFS protocol to catalog, organize, and transfer files within designated shares. A shared folder server is required to create shared folders that contain CIFS shares, NFS shares, or NFS VMware datastores.

Storage processor (SP)—A hardware component that performs VNXe storage operations, such as creating, managing, and monitoring storage resources.

High availability in the storage system

It is important to configure a storage system using a high-availability (HA) design to ensure that business-critical data is always accessible. VNXe offers N+1 redundant architecture, which provides data protection against any single component failure. With redundant components, including dual SPs and dual-ported disk drives, the VNXe system can overcome many different types of multiple component failure scenarios. Furthermore, VNXe system components, such as power supplies, battery backup units (BBU), memory, and disks, are all Customer Replaceable Units (CRU). This is vital when you need to quickly return a storage system back to its HA state.

Storage processor

In a dual-SP VNXe system, storage resources are distributed between the two SPs; however, a storage resource is assigned to only one SP at a time. For example, a shared folder created on SP A will not be associated with SP B unless an SP failover occurs.

An SP fails over when it reboots, experiences a hardware or software failure, or when a user places it in Service Mode. In this case, the storage resource fails over to the peer SP. The surviving SP assumes ownership and begins servicing host I/O requests. When one SP services all I/O requests, performance between hosts and the VNXe system can be degraded. [Table 1](#) describes the SP events that can cause a failover.

Table 1. Events that cause SP failover

Event	Response
SP rebooting	The system or a user rebooted the SP. If the SP is healthy when it comes back online, the storage servers will fail back to it, and the system will return to normal. Check the System Health page to ensure that the SP is operating normally.
SP in Service Mode	The system or a user placed the SP in Service Mode. An SP automatically enters Service Mode when it is unable to boot due to a hardware or system problem. Use the service actions on the Service System page to try to fix the problem. If the SP is healthy, you can reboot it to return it to Normal Mode.
SP powered down	A user powered down the SP.
SP failed	The SP failed and must be replaced.

Failover example

In [Figure 1](#), the hosts connect to shared folder storage and generic iSCSI storage through SP A and SP B. Assume that SP A experienced a hardware failure and the Shared Folder server and iSCSI server failed over from SP A to SP B. The hosts now access the storage servers through SP B. If the storage servers were hosted on SP B, and SP B experienced this hardware failure, the storage servers would have failed over to SP A.

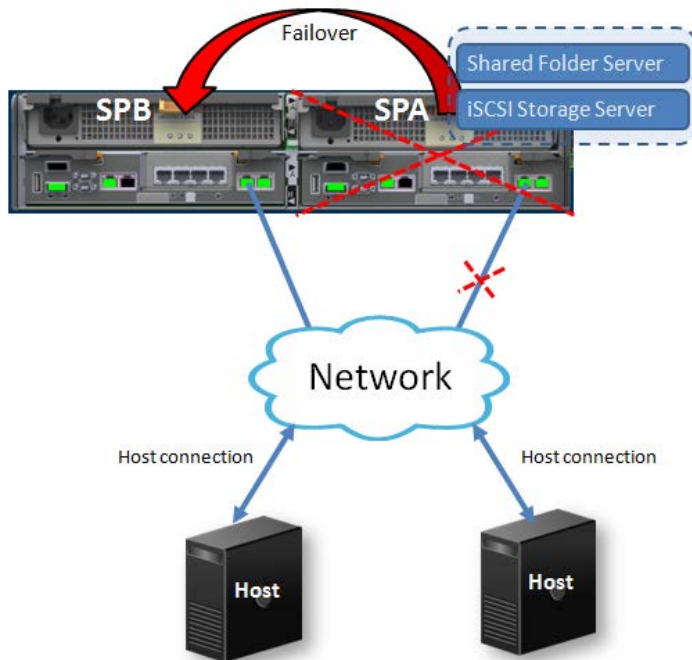


Figure 1. Failover from a hardware failure

In Figure 2, the System Health page shows that a VNXe3100's SP A faulted, causing the iSCSI server to fail over to SP B. You can view this page to determine if an SP or one of its components has an error.

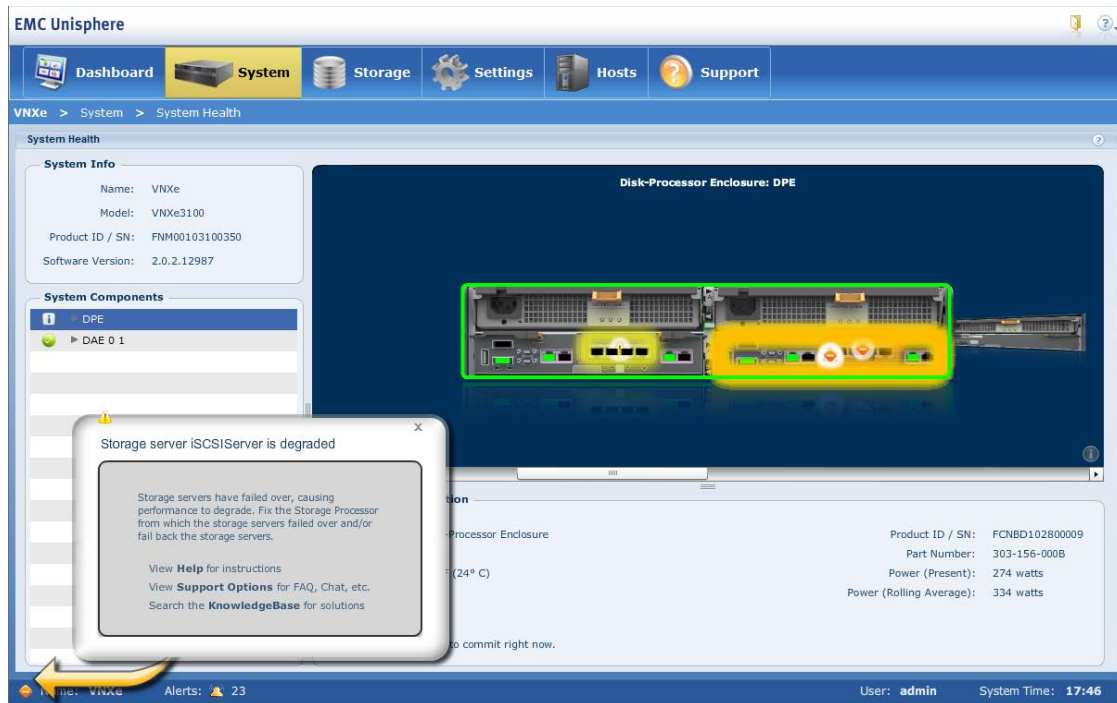


Figure 2. SPA failure indicated in the System Health page

After the storage resources are failed over, data traffic is rerouted to SP B. Hosts that were accessing the storage resource through SP A may experience a small pause in I/O servicing. Hosts should be configured to wait for the storage resources to become available on SP B. Note that different storage resources may behave differently when there is an SP failover. For example, CIFS users may have to reconnect to their shared storage.

Failback of storage resources

Failback is the reverse of failover. It involves moving all storage resources that have failed over back to their original SP. By default, storage resources automatically fail back when the original SP returns to a healthy state.

Administrators can disable automatic failback using the Failback Policy (**Settings > Management Settings**). If disabled, an administrator must perform a manual failback for the storage resources to return to the original SP (shown in [Figure 3](#)). Administrators should disable the Failback Policy only if they need to monitor the failback process or when the failback needs to occur at a specific time.



Figure 3. Manual failing back the SP

Failover of the management interface

The management services for the VNXe system run on one SP at a time, and it does not matter on which SP they run. In the event of an SP failure, the management server fails over to the peer SP, and the management stack starts on the peer SP. Assuming both SPs' management ports are cabled and on the same network, this process is not visible to the user, other than for a brief time when the management stack restarts on the peer SP. If Unisphere™ is open in a web browser at the time of the SP failure, you see a pop-up message indicating a loss of connection to the failed SP. Another pop-up message appears when the connection is reestablished with the peer SP. The management stack remains on this SP even after the failed SP returns to a healthy state.

Power Modules

VNXe systems have redundant power supplies, power cables, and battery backup units (BBUs)¹ that protect data in the event of internal or external power failures. The VNXe system employs dual-shared power supplies. If one power supply fails, the other one provides power to both SPs.

The BBU does not function like an uninterruptable power supply (UPS), because it is not designed to keep the storage system up and running for long periods in anticipation of power being restored. With the BBU, the power to the SP is maintained. This protects the data in the cache and dumps it to the internal SSD vault when there is power loss. The BBU is sized to support the connected SP and is required to maintain the write cache long enough for it to be stored to the vault.

¹ A single-SP VNXe3100/VNXe3150 only has one battery backup unit.

Memory

Each SP has its own dedicated system memory. The VNXe3100 and VNXe3150 have either 4 GB or 8 GB per SP, and the VNXe3300 has 12 GB per SP. This system memory is divided into SP memory and cache memory. Write cache memory is mirrored from one SP to its peer SP. Figure 4 shows a conceptual view of the VNXe memory.

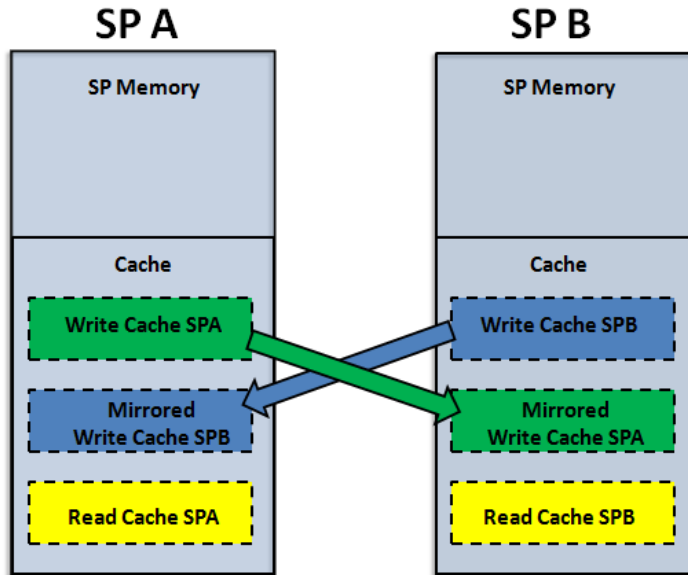


Figure 4. Mirroring write cache memory

The SP cache is used for read and write caching. Read cache is for data that is held in memory in anticipation of it being requested in a future read I/O. Write cache stores write request data waiting to be written to a drive.

In addition to an SP's own read and write cache, cache memory contains a mirror copy of its peer SP's write cache. This is an important availability feature. The majority of all cache memory is allocated to write cache, with the remainder allocated to read cache. The read cache is smaller because the I/O for a read operation is often already in cache, unless the read operation is performing non-sequential read I/O.

Write cache

The VNXe write cache is a mirrored write-back cache. For every write, the data is stored in cache and copied to the peer SP. Then, the request is acknowledged to the host. In this process, write cache is fully mirrored between the VNXe system's SPs to ensure data protection through redundancy. In addition, requests are acknowledged before they are written to disk.

When the VNXe system is shut down properly, the SP cache is flushed to backend drives and disabled. If an SP fails and then reboots, the cache is kept intact through all system resets.

If there is a power loss, each SP uses BBU battery power to write its copy of the write cache to its SSD (a flash-based hard drive), which does not need any power to retain

the data. Upon reboot, the SP cache contents are restored on both SPs. The two SPs then determine the validity of the contents. Normally, both copies are valid. In the event that one SP has a newer version of the cache contents (or if one of them is invalid), the SP with the latest valid copy synchronizes its contents with the peer SP before re-enabling the cache and allowing access to storage resources.

Cache protection module in the single-SP VNXe 3100/3150 systems

Although this white paper focuses on dual-SP VNXe systems, it is important to note that single-SP VNXe3100/3150 systems have a cache protection module in the peer SP slot. This component has a second active cache. It contains the same data as the cache of the single SP (just like the mirrored configuration used in a dual-SP VNXe system). When a failed SP is replaced, the data in the cache protection module is copied over to the SP's cache and then written to disk.

Data availability in your connectivity infrastructure

When designing an HA environment, it is important to carefully plan the connectivity infrastructure. A single point of failure at the host-level, switch-level, or storage system-level can result in data being unavailable for the host applications.

Fail-Safe Networking

All ports on the VNXe system are configured automatically with FSN. To take advantage of FSN, you must cable both SPs the same way. There is a primary physical port on one SP and a secondary physical port on its peer SP. If the primary physical port or link fails, FSN fails over to the corresponding port on the peer SP. The data then gets routed internally through the inter-SP communication link, from the corresponding port on the peer SP to the SP associated with the storage resource. For example, if a given storage resource is accessed through eth2 on SP A and this port fails, FSN will fail over traffic to eth2 port on SP B.

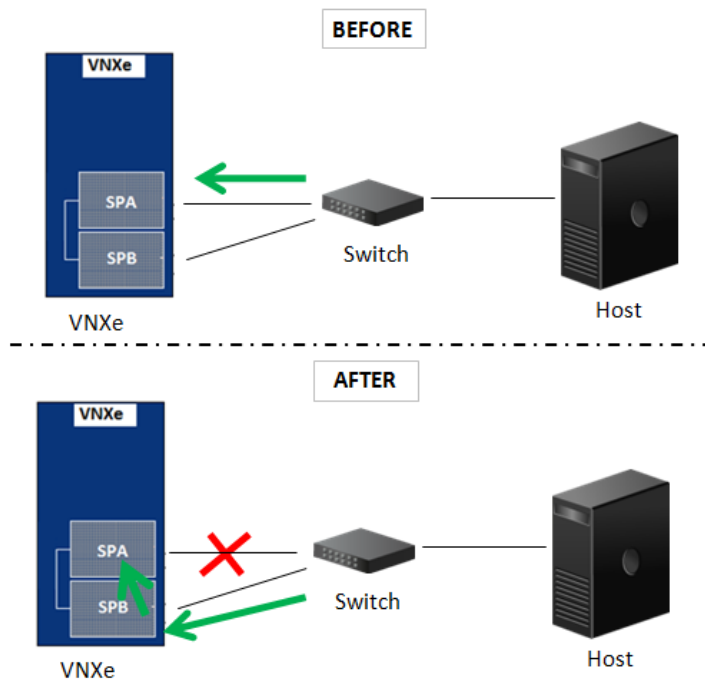


Figure 5. Fail-Safe Networking invoked

In Figure 5, a shared folder was established on SP A and is accessible through network port eth2. When the VNXe system detects that the network link failed on SP A, it reroutes the shared folder data path from SP A to network port eth2 on SP B. It then routes the data internally from SP B to SP A. Once the failed network link recovers, the data path reverts to its original route. Figure 5 shows the logical flow of the data, but does not show what would happen if multiple links per SP were aggregated.

Implementation for iSCSI storage

To ensure that there are redundant paths between the host and storage system, there must be a failover path in case the primary path fails. In an HA configuration, you should set up I/O connections from a host to more than one port on an SP and configure I/O connections between the host and peer SP as an additional safeguard. Having a host connect to more than one of the storage system's front-end ports is called *multipathing*.

When implementing an HA network for iSCSI storage, keep in mind that:

- An iSCSI storage resource on a VNXe system is presented on only one SP at a time.
- You can configure up to four IP interfaces for an iSCSI storage server. These IP interfaces are associated with two separate physical interfaces on the same SP.
- Network switches should be configured on separate subnets. Servers cannot be attached directly to a VNXe system.

Figure 6 depicts an HA configuration for a VMware ESX environment. The VMFS datastore to be accessed resides on SP A. Switch A and Switch B are configured for separate subnets (Switch A–10.245.15.1/22; Switch B–10.245.20.1/22). Each ESX host can access the VMFS datastore via separate NICs. For each ESX host, these NICs are configured for different subnets. For this example, the iSCSI server was configured with two IP interfaces with different IPv4 addresses.

The links in each of the following diagrams are denoted by integers [1-8] and are referenced when explaining traffic routing.

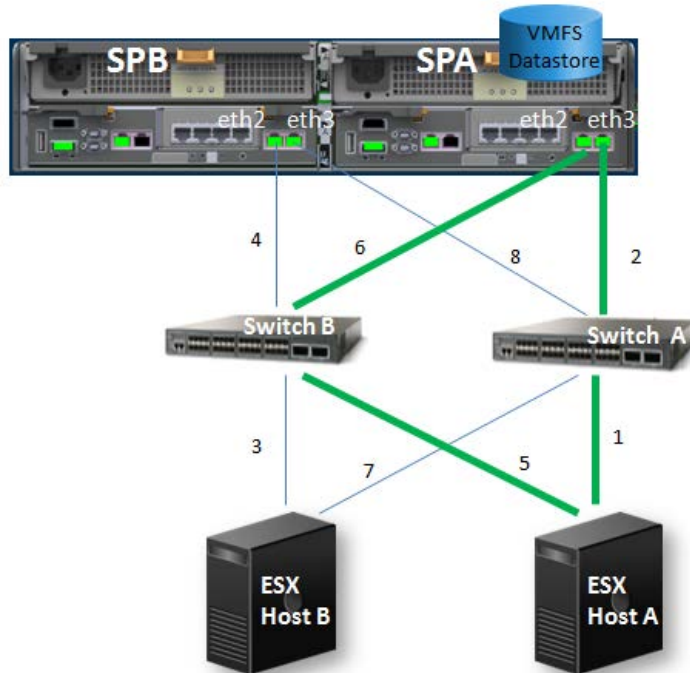


Figure 6. HA configuration for iSCSI storage

When provisioning an iSCSI storage server, specify the network information for one of the network interfaces, as shown in [Figure 7](#).

Figure 7. Provisioning an iSCSI storage server

After creating the iSCSI storage server, assign the second IP address to this server by navigating to the iSCSI storage server details, as shown in [Figure 8](#).

IP Address	Subnet Mask/Prefix Length	Gateway	External Port/MAC Address	VLAN ID
10.245.21.24	255.255.252.0	10.245.20.1	8:0:1b:b5:62:8	0
10.245.20.24	255.255.252.0	10.245.20.1	8:0:1b:83:cf:1d	0

Figure 8. Provisioning the second IP address for an iSCSI storage server

[Table 2](#) shows the storage resource configuration on the VNXe system. There are two IP addresses associated with each iSCSI storage server, and each of these IP addresses is bound to a separate, physical port on the given SP. For HA purposes, configure each iSCSI server with two IP addresses, regardless of whether the SP storage resources were created in SP A or SP B. In this example, the storage resources have been configured on SP A and SP B.

Table 2. Multiple IP addresses assigned to a given iSCSI Server

Component	Storage Server	Port	IP Address
SPA	iSCSI_serverA	eth2	10.245.21.24
		eth3	10.245.20.24
SPB	iSCSI_serverB	eth2	10.245.21.25
		eth3	10.245.20.25

In a fully HA network, ESX Host A has multiple paths to access the VMFS storage on SPA. In this example, the primary path ESX Host A takes to access the VMFS datastore is through Switch A (link 1 -> link2). There is an alternative path through Switch B (link 5 -> link 6) in case the primary path is down.

If a port or link fails, ESX Host A can still access the VMFS storage via an alternative path (link 1 -> link 8 -> inter-SP communication link), as shown in Figure 9, even though the VMFS storage resource still resides in SPA. When FSN is initiated for port eth3 on SP A, traffic is routed through the inter-SP communications link.

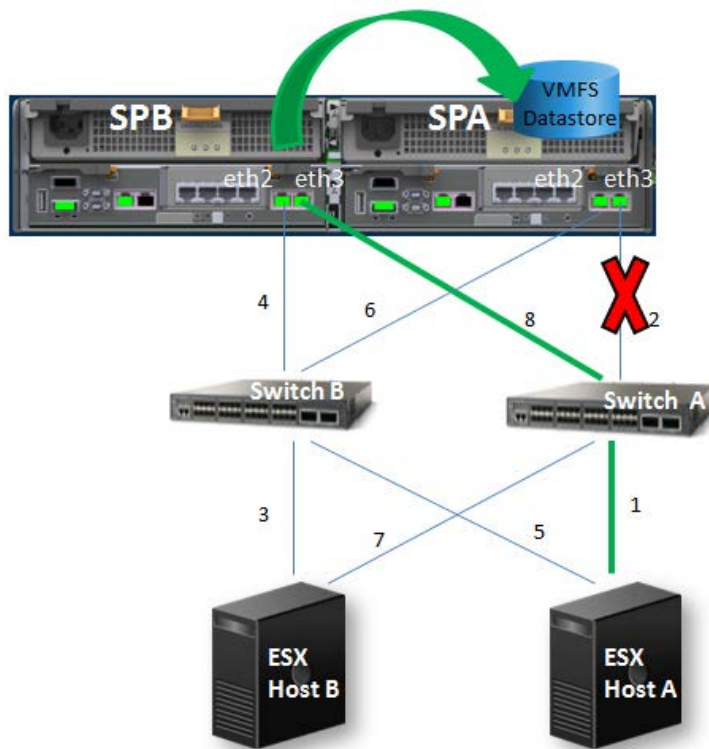


Figure 9. Alternative path remains after SP port of link failure

If Switch A fails, ESX Host A can still access the VMFS storage via the alternative path (link 5 -> link 6), as shown in [Figure 10](#). The resulting data path is the same as in [Figure 9](#).

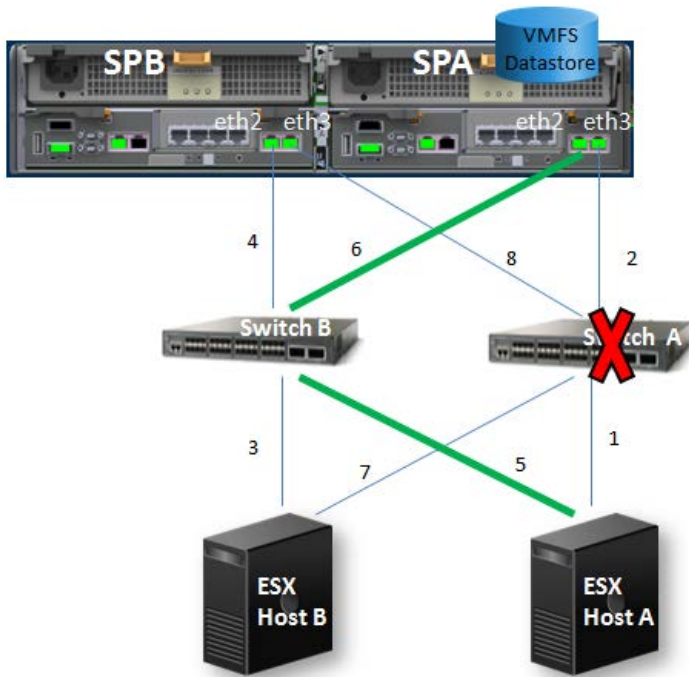


Figure 10. Alternative path remains after switch failure

In the scenario where SP A is faulted and cannot service requests from any hosts, the VMFS storage resource fails over to the peer SP B, as shown in [Figure 11](#).

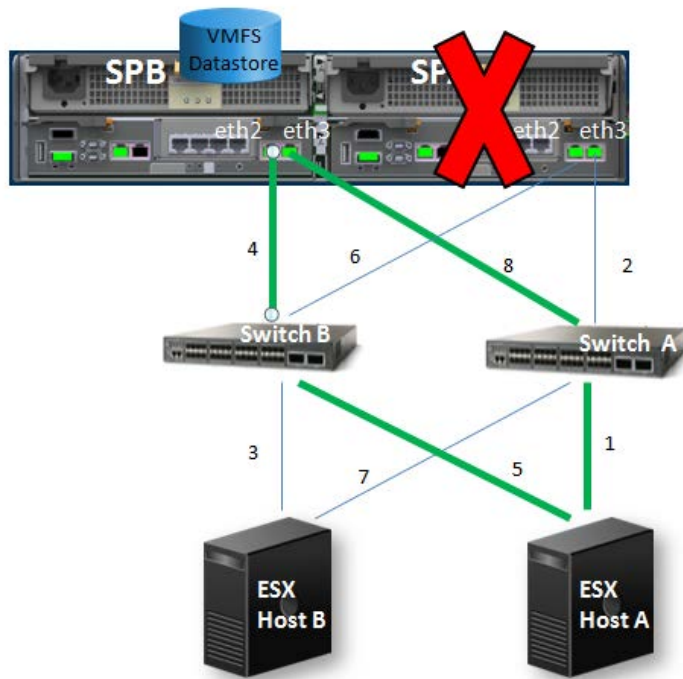


Figure 11. Alternative paths when an SP fails

The I/O traffic is now routed through different paths; either through primary Switch A (link 1 -> link 8) or secondary Switch B (link 5 -> link 4).

When implementing an HA network for iSCSI, you can leverage VMware ESXi's failover software. Native Multipathing Plug-in (NMP) contains policies for Fixed, Round Robin, and Most Recently Used (MRU) device paths. Once port bindings are added to the software iSCSI adapter, you can configure a single vSwitch with two NICs so that each NIC is bound to one VMkernel port. In addition, you can explicitly associate the VMkernel interfaces with the VMware iSCSI software initiator instance.

Implementation for NFS storage

To create a robust HA network for NFS, you can leverage the link aggregation feature on the VNXe system.

LACP is designed primarily to provide load balancing from a server to a large number of clients. Traffic is routed to links based on a hash of the source and destination MAC addresses. Link aggregation allows an administrator to bind multiple physical ports to a single logical port. For example, ports eth3, eth4, and eth5 can be aggregated to port eth2. Link aggregation provides the following features:

- PBasic load balancing across linked ports, since the network traffic is distributed between the aggregated physical ports. This is valid only when a large number of client connections are made to the storage system.

When there is a single client connection, only a single connection within the aggregated link is used. If there is more than one connection, but the number of connections is small, it is less likely that the load will be distributed evenly across the aggregated links.

- Redundant ports—As long as one port in the aggregated group is active, the VNXe system does not lose connectivity.

When implementing an HA network for NFS storage, keep in mind that:

- An NFS storage resource on a VNXe system is presented on only one SP at a given time.
- Two or more eth ports can be aggregated for a shared folder or VMware (NFS datastore) storage server.
- Aggregated ports combine to make a single virtual link with a single MAC address. Depending on how many storage servers use the aggregated ports, this virtual link may have multiple IP addresses.
- Using stacked switches enhances the level of protection in this HA environment.²You can create LACP groups that span physical switches (often referred to as cross-stack LACP) with stacked switches.

The diagram in [Figure 12](#) depicts an HA configuration for a VMware ESX environment. The NFS datastore to be accessed resides on SP A. Switch A and Switch B are set up for a stacked-switch configuration. Each ESX host can access the NFS datastore via separate, individual NICs. In addition, each of the NICs on an ESX host connects to different switches.

The links in each of the following diagrams are denoted by integers [1-8], and are referenced when explaining traffic routing.

When implementing an HA network for NFS, NIC Teaming on the ESXi hosts provides fault tolerance in case of a NIC port failure. Teaming and Failover policies (found in the **Properties** page of the vSwitch) help determine how network traffic is distributed between adapters and how to route I/O traffic in the event of an adapter failure. For more information, refer to the *ESXi Configuration Guide* (for v4.X) and *vSphere Networking Guide* (for v5.X) on the VMware technical resource page (<http://www.vmware.com/resources/techresources/>).

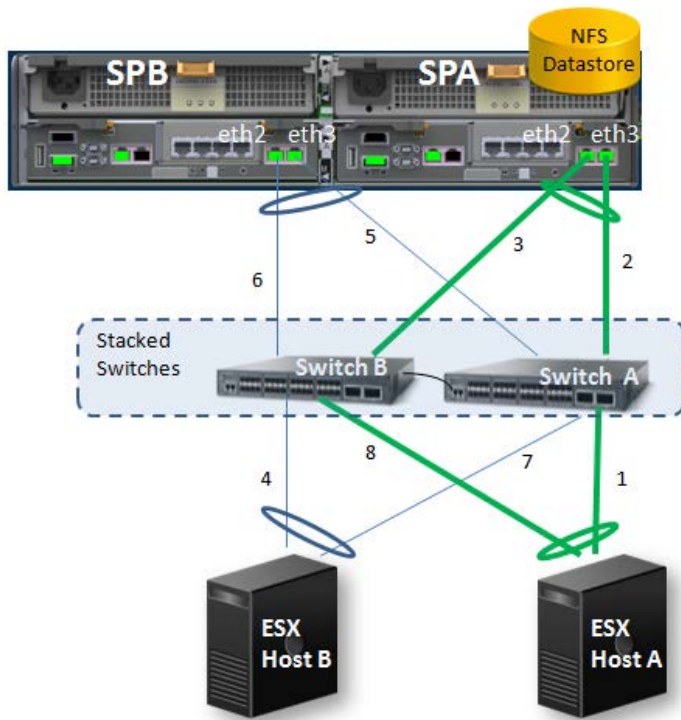


Figure 12. HA configuration for NFS storage

With link aggregation, all used ports must be cabled to the same switch, and the cabling on SP A must be identical to the cabling on SP B. Also, the switch must support and be configured for LACP. Figure 13 shows Unisphere's Advanced Configuration page for the VNXe3150 (Settings > More configuration... > Advanced Configuration). In this example, network ports eth2 and eth3 are bound to create one logical port.

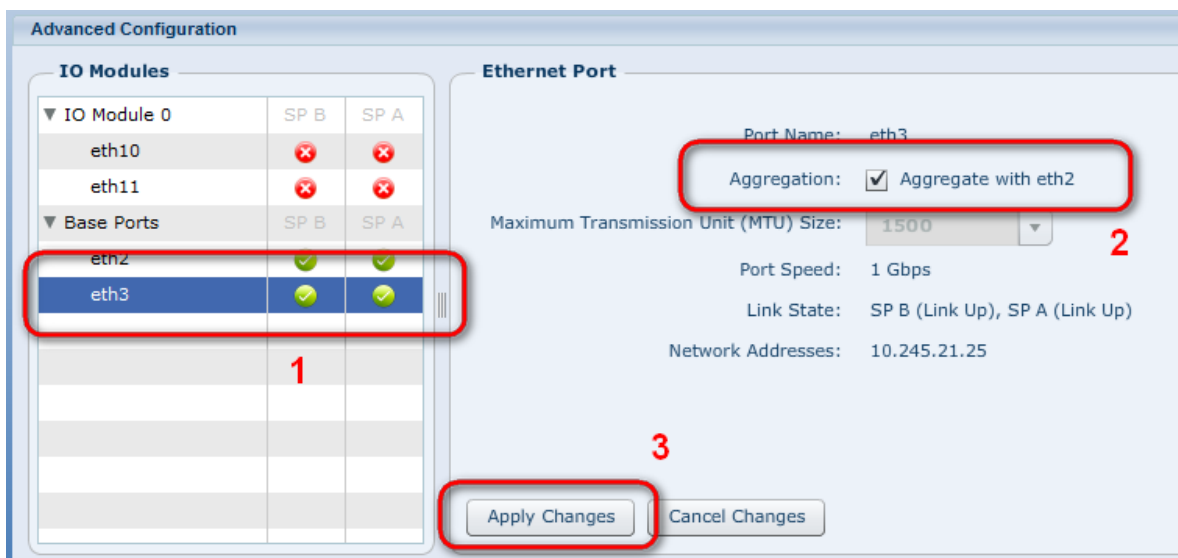


Figure 13. Configuring LACP in Unisphere

Table 3 shows the storage resource configuration on the VNXe system. Port eth3 is aggregated to eth2 on the respective SP, and an IP address is assigned to each port group. For HA purposes, associate each shared folder server with link-aggregated ports, regardless of whether the SP storage resources were created in SP A or SP B. In this example, link aggregation is configured on SP A and SP B.

Table 3. Aggregating together ports eth2 and eth3

Component	Storage Server	LACP Ports	IP Address
SPA	NFS_serverA	eth2/eth3	10.245.21.24
SPB	NFS_serverB	eth2/eth3	10.245.21.25

In a fully HA environment, ESX Host A can access the NFS storage on SPA through Switch A (link 1 -> link 2) and Switch B (link 8 -> link 3).

If a port or link fails, ESX Host A can still access the NFS storage through Switch A (link 1 -> link 3) or Switch B (link 8 -> link 3), as shown in Figure 14, even though the NFS storage resource still resides in SPA.

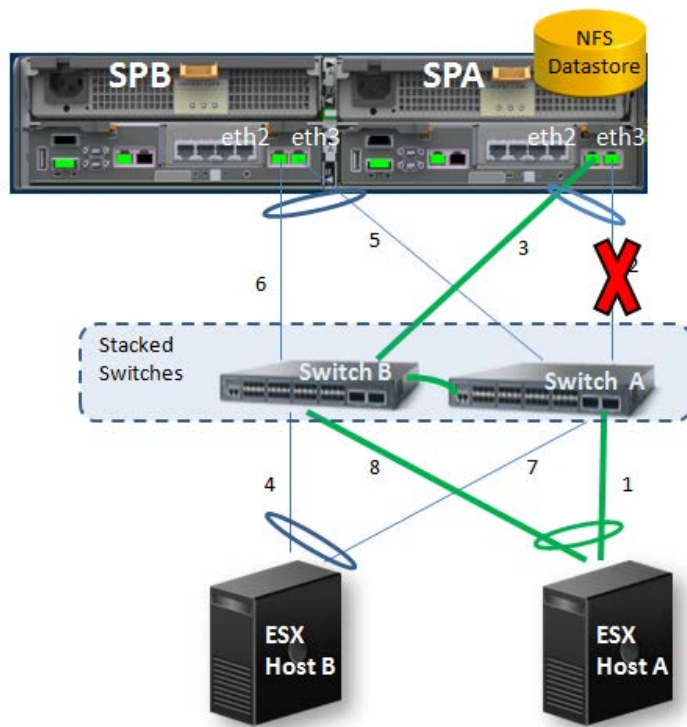
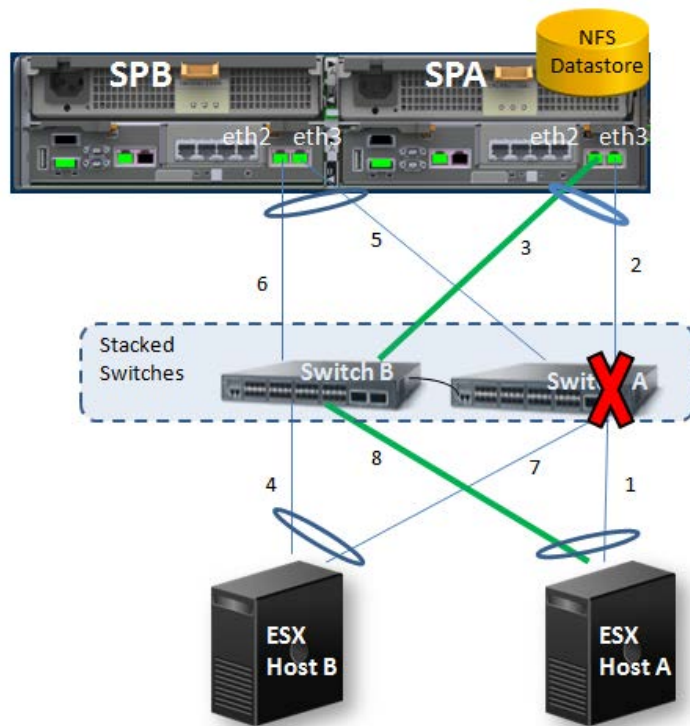


Figure 14. Link remains active after SP port or link failure

If Switch A fails, ESX Host A can still access the NFS storage through Switch B (link 8 -> link 3), as shown in [Figure 15](#).



[Figure 15](#). Link remains active after switch failure

It is important to note that if independent switches with LACP trunks to each switch are used, a switch failure forces all data traffic over the VNXe system's inter-SP communication link via FSN. If stacked switches are used, you can create an LACP trunk with links to each switch in the stack. If a switch fails, the links to that switch fail, but traffic fails over to the surviving links in the LACP group to the other switch. No FSN failover occurs. Traffic continues at full bandwidth over the primary network ports on the SP.

If SP A is faulted and cannot service requests from any hosts, the NFS storage resource fails over to the peer SP B.

The I/O traffic is now routed through a different path; through Switch A (link 1 -> link 5) or through Switch B (link 8 -> link 6) as shown in [Figure 16](#).

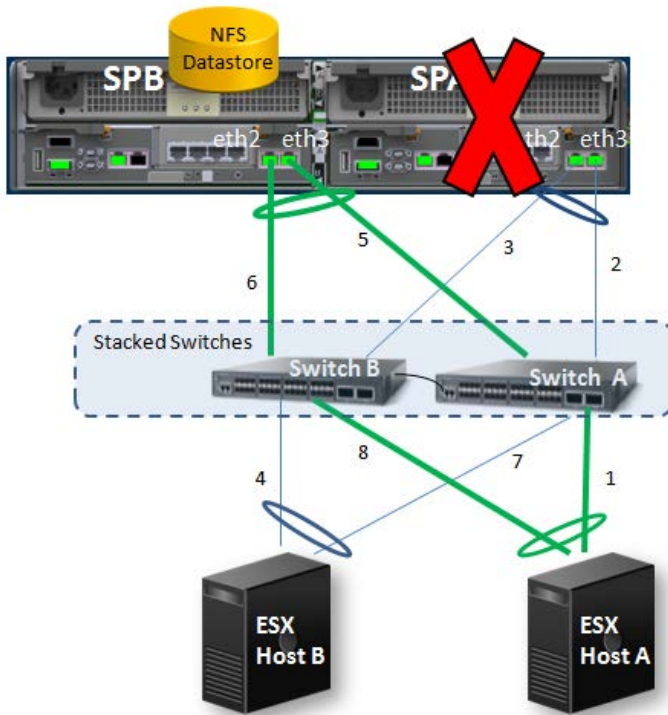


Figure 16. Alternate path when an SP fails

High availability in disk storage

Number of backend buses

Each SP in the VNXe system has a dedicated backend bus, which connects to additional DAEs for storage expansion. In a dual-SP VNXe system, if one bus fails, both SPs can still access the DAE through the other bus.

RAID configurations

To access data on the disk drives, you must create storage pools. A storage pool consists of disks that are bounded into one or more *RAID groups*. (End users do not have to create separate RAID groups manually.) All disks in a RAID group must be the same type (SAS, near-line SAS, or flash), and have the same speed and capacity. This ensures the highest usable capacity and reliability, and consistent performance.

VNXe offers RAID 5, RAID 6, and RAID 10. Different configurations offer different kinds of protection against disk failures, and are suitable for different applications. [Table 4](#) provides a brief summary of the RAID levels supported by VNXe. The following sections provide more information about each type of RAID group.

Table 4. RAID level details and configurations

RAID Type	Protects	Recommended for	Disk Configuration VNXe3100/3150	Disk Configuration VNXe3300
RAID 5* Striped, distributed, parity-protected	Against single disk failure.	Transaction processing; is often used for general purpose storage, relational databases, and enterprise resource systems.	(4+1), (6+1), (10+1) A minimum of 5 disks must be allocated each time you allocate to a pool.	(4+1), (6+1), (10+1) A minimum of 5 disks for (4+1), 7 disks for (6+1) or 11 disks for (10+1) must be allocated each time you allocate to a pool.
RAID 6** Striped, distributed, double-parity-protected.	Against double disk failure.	Same uses as RAID 5, only where increased fault tolerance is required.	(4+2), (10+2) A minimum of 6 disks for (4+2) or 12 disks for (10+2) must be allocated each time you allocate to a pool.	(4+2), (10+2) A minimum of 6 disks for (4+2) or 12 disks for (10+2) must be allocated each time you allocate to a pool.
RAID 10 Mirror-protected.	Against multiple disk failures, as long as the disk failures do not occur in the same mirrored pair.	RAID 10 may be more appropriate for applications with fast or high processing requirements, such as enterprise servers and moderate-sized database systems.	(3+3) A minimum of 6 disks must be allocated each time you allocate to a pool.	(3+3) A minimum of 6 disks must be allocated each time you allocate to a pool.

* Flash drives as configured in RAID 5 (4+1) are only available for VNXe3150 and VNXe3300 systems

** RAID 6 is used by default for all NL-SAS drives in the VNXe systems. Also, note that SAS drives cannot be configured using RAID 6 (10+2).

RAID 5

RAID 5 stripes data at a block level across several disks and distributes parity among the disks. With RAID 5, no single disk is devoted to parity. This distributed parity protects data if a single disk fails. Failure of a single disk reduces storage performance, so you should replace a failed disk immediately.

The failure of two disks in a RAID 5 disk group causes data loss and renders any storage in the RAID group unavailable. The failed disks must be replaced and the data then restored from a disk-based backup or accessed via a manual failover to a replicated system. For this reason, you should replace failed disks immediately.

RAID 6

RAID 6 is similar to RAID 5; however, it uses a double-parity scheme that is distributed across different disks. This offers extremely high fault tolerance and disk-failure tolerance. This configuration provides block-level striping with parity data distributed across all disks. Arrays can continue to operate with up to two failed disks. Double parity gives time to rebuild the array without the data being at risk, if a single additional disk fails before the rebuild is complete.

The failure of three disks in a RAID 6 disk group causes data loss and renders any storage in the RAID group unavailable. As with RAID 5, the failed disks must be replaced and the data then restored from a disk-based backup or accessed via a manual failover to a replicated system. RAID 6 provides high performance and reliability at medium cost, while providing lower capacity per disk.

RAID 10

This configuration requires a minimum of six physical disks to implement in VNXe systems, where three mirrored sets in a striped set together provide fault tolerance. Although mirroring provides fault tolerance, failed disks must be replaced immediately and the array rebuilt.

A minimum of six disks can be allocated at a time to a pool, with three used strictly for mirroring. In other words, to provide redundancy, three disks out of every six are duplicates of the other three disks. This configuration is used in custom pools that are created using SAS disks.

Dual drive ownership

A dual-SP VNXe system supports dual ownership of hard drives by its two SPs. All hard drives are dual ported and can accept I/O from both SPs at the same time.

Hot sparing

Hot spares are spare disks that you can allocate when or after you configure your RAID groups. A hot spare replaces a failed disk in the event of a disk failure. When a disk drive fails, the VNXe system rebuilds data to the hot spare disk, using the remaining disks in the RAID group. When the failed disk is replaced, the VNXe system copies the data and parity from the hot spare to the new drive. This process is called *equalization*. After equalization is completed, the hot spare returns to its default status and becomes ready for any future disk failure event. You can use Unisphere to check the status of a hot spare.

If the administrator chooses to configure pools using the default options, a hot spare is assigned for the first 30 disks of the same disk type. Then, another spare disk is assigned for each additional group of 30 disks of the same disk type. For NL-SAS disks, you must configure hot spares manually.

The administrator is responsible for configuring additional hot spares. For ease of management, it is recommended that the administrator configures a hot spare on the

last drive slot on a DPE or DAE. The hot spare, however, may be configured anywhere in the system, except in the slots for system drives.

When planning hot spares for the system, the spares should be as large as or larger than the drives they may be required to replace. The hot spares must be the same type of disk as the disks they may replace.

Using replication for high availability

Replication allows the IT administrator to automatically maintain a complete second copy of the storage resource on the local or remote system.

File-level replication, such as CIFS or NFS shared folders, is managed entirely within the VNXe environment through the Unisphere interface. Once a storage resource has been created, IT administrators can create a replication session either to its peer SP (local replication) or to a remote system (remote replication), which can be a VNXe, VNX®, or Celerra system. IT administrators can control the frequency at which the source and destination systems are synchronized.

Block-level replication, such as Hyper-V and Exchange, leverage EMC Replication Manager to create, schedule, and manage iSCSI replications. Replication Manager provides a graphical user interface for managing the replication of iSCSI LUNs. Before creating a replica, Replication Manager ensures that applications are in a quiescent state and that the cache is flushed, so that the replica is consistent from the point of view of client applications.

For more information on replication, refer to the *EMC VNXe Data Protection* white paper on the EMC Online Support (<https://support.emc.com>) > VNXe Product Page.

Continuous Availability

SMB 3.0 protocol support is available with Microsoft Windows 8 and Microsoft Windows Server 2012 systems and has significant improvements over the previous SMB versions. As of VNXe Operating Environment (OE) version 2.4, the SMB 3.0 protocol is enabled by default. Continuous Availability (CA) is one of the enhancements introduced with this protocol.

CA enables applications to be less impacted during a Shared Folder server failure or recovery of application. In the situation where a storage processor is faulted or placed in service mode, storage resources are failed over to the peer storage processor. With CA, application access to all open files that were present prior to the failover is re-established and the failover is transparent to the end users. In addition, CA increases performance by making synchronous writes to the backend storage.

For more information on CA, refer to the *Introduction to SMB 3.0 Support* white paper on the EMC Online Support (<https://support.emc.com>) > VNXe Product Page.

Conclusion

Designing an organization's IT environment with high availability (HA) in mind is a top concern for IT administrators. Setting solid HA measures assures business continuity and lessens time and cost in recovery efforts, should there be technical difficulties. VNXe systems have these measures built in and can ensure that data is readily accessible to the customer.

This white paper describes the key HA features that VNXe offers at the network and storage levels. Configuring multiple paths to storage resources allows business critical data to remain accessible in situations of network failure. VNXe has an N+1 redundant architecture, which provides data protection against any single component failure. Choosing the appropriate disk configuration alleviates performance impacts and disruptions caused from disk failure. Replicating production data to a remote location provides an additional level of data availability, should a disaster occur at the primary site.

References

The following documents can be found on Powerlink or on the EMC Online Support (<https://support.emc.com>) › VNXe Product Page:

- *EMC VNXe Series Storage Systems – A Detailed Review*
- *EMC VNXe Data Protection – A Detailed Review*
- *Replicating VNXe iSCSI to VNX using Replication Manager 5.3.2 – Technical Notes*
- *Replicating VNXe CIFS/NFS Shared Folders to VNX – Technical Notes*
- *Introduction to SMB 3.0 Support*