# Fibre Channel over Ethernet Fault Tolerance: Achieve High Availability in a Converged Network

## What You Will Learn

Critical requirements for any modern data center are resiliency and fault tolerance. Networks must be designed and built with superior reliability and uptime. These requirements are even more important in a converged environment in which storage and LAN traffic is sharing the same physical infrastructure.

This document describes the software and hardware features available on the Cisco Nexus® 7000 Series Switches and Cisco® MDS 9500 Series Multilayer Directors and their roles in addressing data center high-availability requirements. In particular, this document discusses how storage traffic resiliency is achieved in a Fibre Channel over Ethernet (FCoE) environment.

## FCoE Overview

Data centers are typically designed for multiple networks with different requirements for latency and resiliency: a separate loss-intolerant network dedicated to storage (SAN), and a traditional loss-tolerant Ethernet network (LAN). Managing multiple networks presents administrators with a challenging, inflexible environment with redundant network adapters, cabling, and switches.

FCoE is a standards-based protocol that natively maps Fibre Channel to Ethernet for transport in a loss-less Ethernet LAN. FCoE allows consolidation of I/O in the data center, which addresses many of the challenges of today's data centers. For a more detailed introduction to FCoE, please see Cisco FCoE Fundamentals.

## High Availability and Fault Tolerance

High availability in the data center refers not only to device-specific availability and uptime, but also to network design and features that prevent downtime in the case of a catastrophic event. Uptime in this context refers to availability of the switch to direct traffic. As more and more equipment is added to the data center network, the high availability of the network may be undermined. Network architects need to consider design best practices to reduce single points of failure and achieve network uptime goals in the data center.

The Cisco Nexus 7000 Series is designed to be highly available and to perform uninterrupted under any conditions. For more information about the Cisco Nexus 7000 Series features and technologies that bring high availability to the data center, please refer to Continuous Operations and High Availability.

Cisco MDS 9500 Series directors also provide exceptional high availability, as documented in detail in High Availability of SAN with Cisco MDS 9500.

These features are further discussed here in the context of storage-related traffic and their benefits in a converged network.

## Supervisor-Level High Availability

Cisco Nexus 7000 Series and Cisco MDS 9500 Series products provide director-class high availability. They include dual redundant power supplies, hot-swappable optical interface modules, and dual supervisors in active-standby mode. When the switch boots up, the active and standby roles are defined for the two available supervisors. The standby switch is in the hot-standby state, monitoring the active supervisor for failure. If a failure occurs, the standby

switch assumes the active role, and the active supervisor is reset. This switchover is performed using nonstop forwarding (NSF), which helps ensure the correct forwarding of traffic in transit.

In a switchover occurs, the FCoE forwarding data plane will not be affected, and storage traffic will continue to be forwarded by the switch until the control plane in the newly active supervisor is up and protocols have converged.

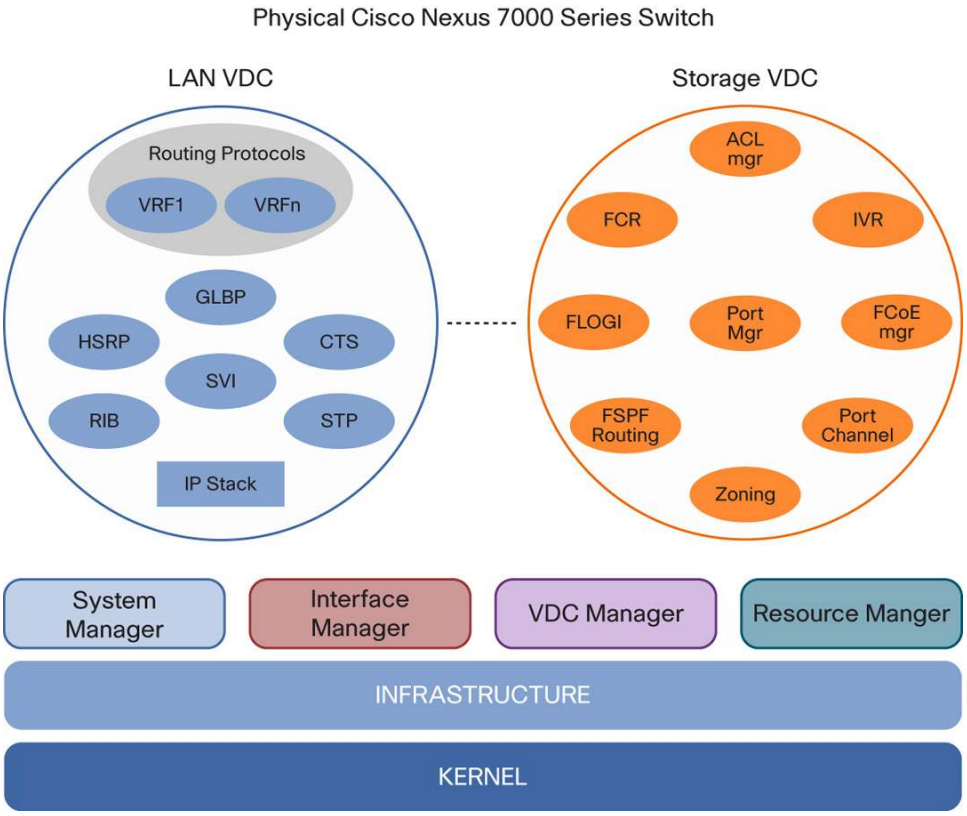
Process restart is another feature of the Cisco NX-OS Software operating system; a process failure causes the failing process to restart without causing a system wide failure.

### Process-Level Availability Through Virtual Device Contexts

Storage network designers have traditionally kept storage traffic separate from other LAN traffic to reduce collateral damage that may occur to traffic in transit. The virtual device context (VDC) feature of the Cisco Nexus 7000 Series Switches allows the switch to be virtualized into multiple logical devices with their own separate processes. Each local entity operates as an independent switch.

Each VDC can contain its own unique and independent set of VLANs and Virtual Route Forwarding (VRF) instances. Multiple interfaces can belong to a single VDC, which allows the forwarding data plane to be independent and virtualized. In addition, each VDC has a separate configuration, routing processes, and process memory. Security and management policies can be established on a per-VDC basis. The Cisco Nexus 7000 Series introduces the concept of a storage VDC that allows storage traffic to be segregated from the rest and monitored accordingly (Figure 1). This model gives enormous power to the storage administrators: when a VDC is used, all debugging, management, and configuration is applied to that VDC. In a converged environment, storage administrators can be in charge of storage VDCs and network operators can be responsible for LAN VDCs, all residing on the same infrastructure.

**Figure 1.** Self-Contained Processes in a VDC



On a Cisco Nexus 7000 Series Switch, each VDC is isolated from the others. This isolation allows the storage traffic to be unaffected by faults in other parts of the system (or other VDCs). Restarting a LAN VDC (either manually or through a software fault) will not affect a storage VDC. In addition, a separate management plane for configuration and debugging enhances the storage network administrator's capability to control network resources unaffected by actions taken on the LAN network, despite sharing the same physical resource. Debugging and troubleshooting is independent on each VDC. The high level of redundancy and management capabilities available on the Cisco Nexus 7000 Series is beneficial to any converged network.

In addition to providing process-level isolation and therefore resiliency, VDCs can be configured so that, if a failure occurs, NSF switchover is performed. In this case, a new control plane is instantiated, while the data plane forwarding is unaffected. This mechanism therefore provides switch-level redundancy, helping ensure the integrity of the storage traffic. The action performed in the event of a failure is defined by the modes in Table 1. Note that this mode can be configured only by the super user and applies to the whole switch and therefore all VDCs.

In the case of a switch with dual supervisors, the storage administrator can use the Reset mode and in the event of a failure perform a NSF switchover to the standby supervisor. This approach will not cause traffic loss and will help ensure the integrity of the storage traffic traversing the Ethernet network.

It is a best practice recommendation for the switches in a converged network to have dual supervisors. However, in situations in which the switch in a converged network has a single supervisor, the Restart mode should be set on the VDCs so that in the event of a failure, the storage VDC will restart itself and be available.

**Table 1.** Actions Performed Based on the Policy Set

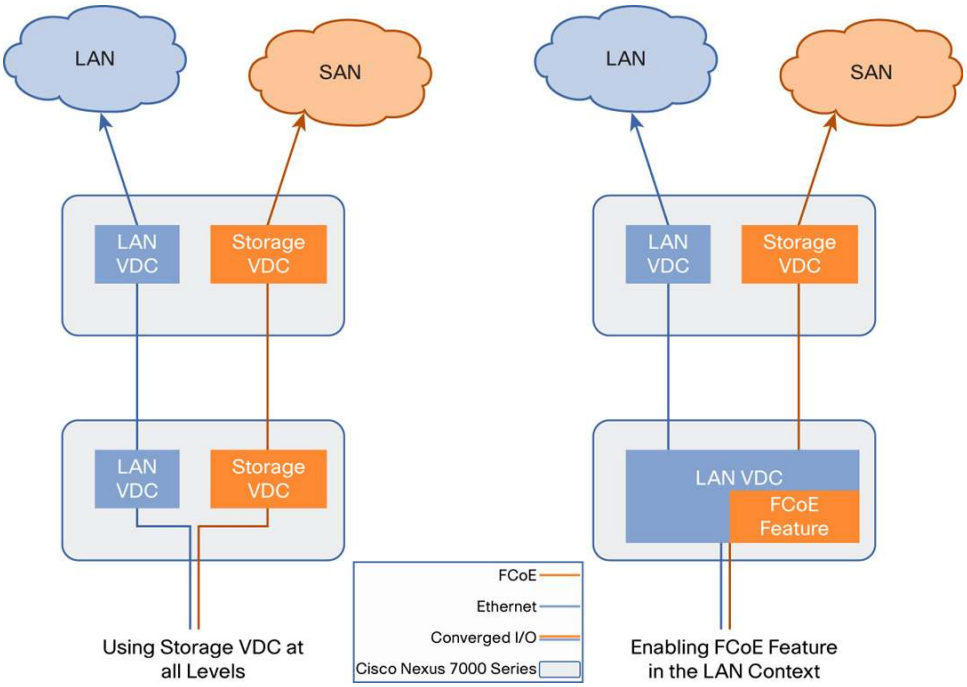
| VDC Policy | Action |
|---|---|
| **Bringdown** | The VDC stays in the failure mode. The switch needs to be reloaded for recovery. |
| **Reset** | If the switch is occupied by two supervisors in active-standby mode, a supervisor switchover to the standby module is performed. If only a single supervisor is present, that module will be reloaded, thereby affecting all VDCs. |
| **Restart** | The failed VDC is deleted and re-created. |

Enhancements made to the Cisco Nexus 7000 Series VDC model allows the ingress link to belong to two different VDCs based on the frame Ethernet type. As shown in Figure 2, the Ethernet driver at the ingress interface identifies the FCoE control plane traffic and directs it to the storage VDC for processing. All other traffic is forwarded to the LAN VDC.

Use of a storage VDC would segregate LAN and SAN control plane traffic ingress to the Cisco Nexus 7000 Series Switch and allow better management of SAN traffic through a traditional Ethernet switch while enabling the administrator to take advantage of the high-availability features provided by the VDC to help ensure the integrity of their storage traffic. However, deployment of a storage VDC is not required. As shown in Figure 2, storage traffic can also be handled in a LAN VDC by enabling the FCoE feature. In this scenario, all the resources available to the VDC are shared by both the Ethernet and storage-specific processes.

Overall, VDCs on the Cisco Nexus 7000 Series create an effective separation of control traffic and management domains suitable for infrastructure consolidation.

**Figure 2.** Treatment of Storage Traffic in VDCs




## Link-Level High availability

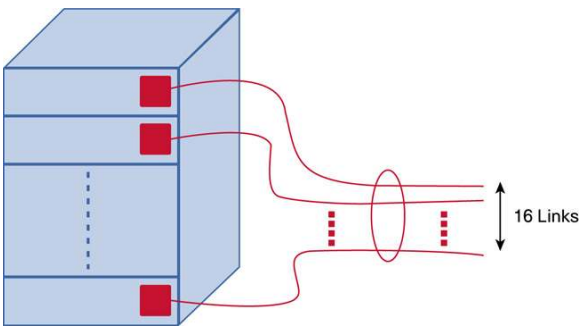
Another level of high availability can be achieved at the link level.

On the Cisco Nexus 7000 Series and Cisco MDS 9500 Series, 16 FCoE links can be bundled together to form a PortChannel to increase the aggregate bandwidth and create link resiliency. The links can reside anywhere on one chassis and connect to any port on a neighboring chassis. The PortChannel forms a logical link between the two switches.

This logical connection will stay active as long as at least one port in the group is operational and active. Line-card failures on the neighboring switch may bring down individual ports; however, the PortChannel will stay active, forwarding and load balancing the traffic. This feature reduces the state change activity at the upper layers and provides a more stable operating environment because neighboring switch failures are isolated.

As illustrated in Figure 3, to increase the resiliency of the PortChannel, administrators should spread the FCoE links in a PortChannel across the whole chassis. This approach isolates a PortChannel failure from the failure of one particular line card.

**Figure 3.** Spreading the Load of a PortChannel Across the Whole Chassis



These software and hardware resiliency features help eliminate the possibility of network downtime and enhance the profile of the Cisco Nexus 7000 Series and Cisco MDS 9500 Series products as true director switches for converged LAN and SAN networks.

### In-Service Software Upgrade

Cisco NX-OS also offers In-Service Software Upgrade (ISSU), which enables administrators to perform nondisruptive software upgrades. This feature allows upgrades to supervisor and switching modules with little to no negative effect on the data forwarding plane, particularly on the FCoE traffic.

### High Availability in Multi-Hop FCoE Network

Network availability depends on the physical availability of the networking equipment and the design of the network. The design of the data center network will have direct affect the mean time to repair (MTTR), which is one of the determinants of network availability. MTTR is generally defined as the average time needed to restore full function to the network. Since storage data loss has catastrophic consequences, data center designers and administrators must have the tools necessary to increase device uptime and to build redundancy into the network.

For storage traffic, appropriate network design can greatly reduce the MTTR. For example, network serialization tends to increase MTTR, whereas building parallelism into the network decreases it because points of failure are eliminated, therefore making the storage network highly available.

To decrease the MTTR, storage network designers should consider several factors:

- Individual component failure
- System failure
- Single points of failure

Figure 4 shows a typical multi-hop FCoE network that addresses these design considerations.

#### Individual Component Failure

Failure to switch traffic by any individual switch in the network would affect the overall availability of the network. In the typical multi-hop FCoE network shown in Figure 4, the Cisco Nexus 7000 Series and Cisco MDS 9500 Series director switches in the path of the storage traffic are highly available individual switches. Dual supervisors, VDCs, and redundant power supplies allow the switch to forward traffic nonstop in the event of a failure. These features reduce the MTTR due to individual component failure.

#### System Failure

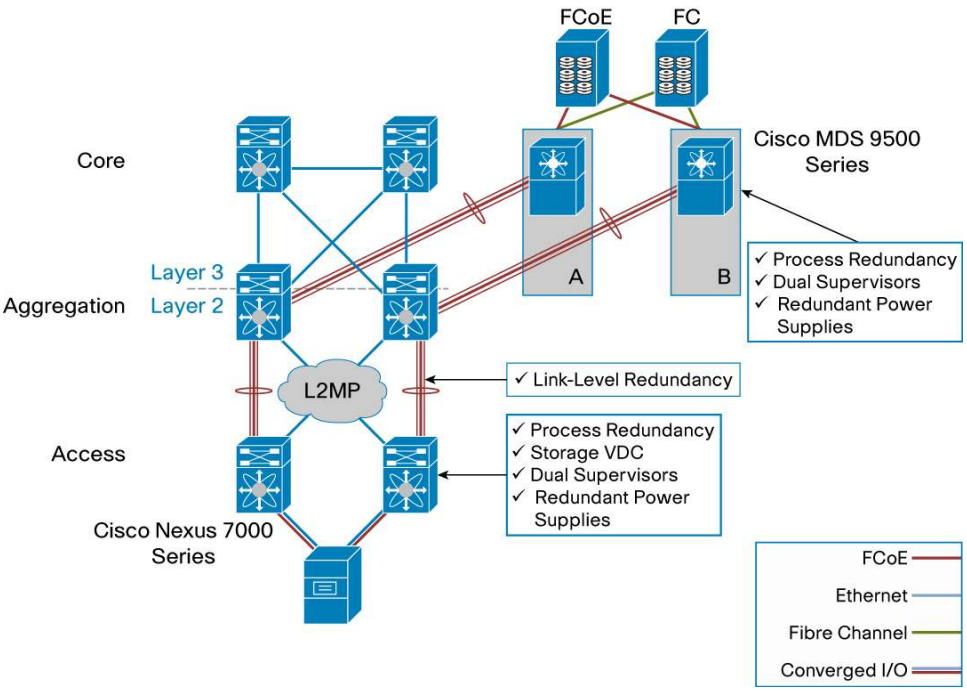
System failure refers to failures due to external factors such as accidental removal of cables. In a converged network in which multiple sets of both network and storage administrators manage the switches, the opportunities for failures due to accidental actions are increased. Use of PortChannels for storage traffic alleviates storage traffic loss due to accidental cable removal. Redundant power supplies that reside on different power circuits in the data center prevent switch downtime due to an accidental power outage.

#### Single Points of Failure

It is important to build parallelism into the network so that traffic can still have a path to its intended target in the event of a failure. The storage network shown in Figure 4 does not have any single points of failure. From the server to any of the hosts, there are two distinct paths to traverse, and within each segment, through the use of PortChannels, the FCoE traffic can choose from multiple paths. Building this type of parallelism into the network allows the network operators and storage administrators to reduce downtime, lowering MTTR.

These Cisco NX-OS features, available on the Cisco Nexus 7000 Series Switches and Cisco MDS 9500 Series Multilayer Directors, position these devices as ideal candidates for deployment in a converged network in which high availability is of utmost importance.

**Figure 4.** Highly Available Converged Network




## Conclusion

As data centers continue to expand, data center architects face increasing pressure to design networks that not only meet current needs, but also have the potential to grow to satisfy future requirements. They are challenged to achieve more with less equipment. Converging the disparate data center networks into one ubiquitous infrastructure reduces capital and operating costs and at the same time allows the data center to grow to meet business requirements. Consolidation of these networks can also increase the opportunities for downtime for critical processes because points of failure are reduced to fewer devices. As discussed in this document, the Cisco Nexus 7000 Series and Cisco MDS 9500 Series director switches provide data center administrators with ample technologies to increase network availability. These devices help ensure that the processes in the switch stay active despite failures and that data paths through the network stay operational despite individual link failures.

The software and hardware features discussed in this document highlight the resiliency of the Cisco Nexus 7000 Series Switches and Cisco MDS 9500 Series Multilayer Directors that makes them excellent choices for data center consolidation.

## For More Information

http://www.cisco.com/go/unfiedfabric

**Americas Headquarters**
Cisco Systems, Inc.
San Jose, CA

**Asia Pacific Headquarters**
Cisco Systems (USA) Pte. Ltd.
Singapore

**Europe Headquarters**
Cisco Systems International BV Amsterdam,
The Netherlands

Cisco has more than 200 offices worldwide. Addresses, phone numbers, and fax numbers are listed on the Cisco Website at **www.cisco.com/go/offices.**



Printed in USA C11-644817-00 03/11