Large SAN Design Best Practices Using the Cisco MDS 9710 Multilayer Director

What You Will Learn

As SANs continue to grow in size, many factors need to be considered to help scale and manage them. This document focuses on large SAN deployments in a data center and provides best practices and design considerations for the design of a large physical fabric. It does not address networks implementing Inter-VSAN Routing (IVR), Fibre Channel, or Fibre Channel over IP (FCIP) SAN extension, or intelligent fabric applications (for example, Cisco Data Mobility Manager [DMM] or I/O Acceleration [IOA]).

Design Parameters

In SAN environments, many design criteria need to be addressed, such as the number of servers that can access a shared storage frame, network topology, and fabric scalability. This document focuses on the following design parameters:

- 1000 or more end devices (servers, storage, and tape devices)
- Majority of end devices with connection speeds of 8 and 16 Gbps
- Identical dual physical fabrics (Fabric A and Fabric B)

Cisco MDS 9700 Series Multilayer Directors

The Cisco® MDS 9710 Multilayer Director is the newest-generation director-class multilayer switch. It supports up to 384 line-rate 16-Gbps Fibre Channel or 10-Gbps Fibre Channel over Ethernet (FCoE) ports. The Cisco MDS 9710 comes with dual supervisor modules and six fabric modules and provides up to 24 terabits per second (Tbps) of chassis throughput.

The Cisco MDS 9700 48-Port 16-Gbps Fibre Channel Switching Module delivers line rate nonblocking 16-Gbps Fibre Channel performance for scalability in virtualized data centers. Line-rate 16-Gbps performance provides high bandwidth throughput to enable consolidation of workloads from thousands of virtual machines while reducing the number of SAN components needed, providing scalability for future SAN growth. This line-card module is hot swappable and provides all the same features as in previous Cisco MDS 9000 Family products, including predictable performance, high availability, advanced traffic management, integrated virtual SANs (VSANs), high-performance Inter-Switch Links (ISLs), fault detection, isolation of errored packets, and sophisticated diagnostics. This module offers new hardware-based slow drain, real-time power consumption reporting, and improved diagnostics.

Table 1 lists the part numbers for ordering the Cisco MDS 9700 Series Multilayer Director components.

Table 1. Part Numbers for Cisco MDS 9700 Series Components

Part Number	Product Description
Cisco MDS 9700 Series Component	
DS-C9710	MDS 9710 Chassis, No Power Supplies, Fans Included
DS-X97-SF1-K9	MDS 9700 Series Supervisor-1

| Part Number	Product Description
DS-X9710-FAB1	MDS 9710 Crossbar Switching Fabric-1 Module
DS-X9448-768K9	48-Port 16-Gbps Fibre Channel Switching Module
Optional Licensed Software	
M97ENTK9	Enterprise package license for 1 MDS9700 switch
DCNM-SAN-M97-K9	DCNM for SAN License for MDS 9700

SAN Topology Considerations

It is common practice in SAN environments to build two separate, redundant physical fabrics (Fabric A and Fabric B) in the event that a single physical fabric fails. The topology diagrams in this document show a single fabric; however, customers would deploy two identical fabrics for redundancy. Most designs for large networks will use one of two types of topology for the physical fabric:

- Two-tier topology: Core-edge design
- Three-tier topology: Edge-core-edge design

In the two-tier design, servers connect to the edge switches, and storage devices connect to one or more core switches (Figure 1). This topology allows the core switch to provide storage services to one or more edge switches, thus servicing more servers in the fabric.

Figure 1. Sample Core-Edge Design
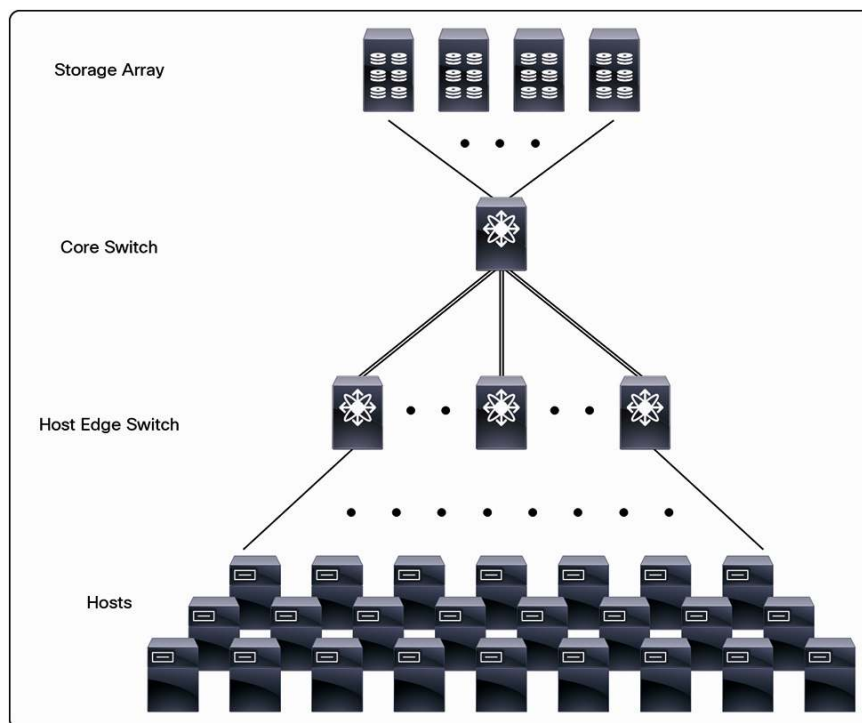
In environments in which projections for future growth of the network estimate that the number of storage devices may exceed the number of ports available at the core switch, a three-tier design may be preferred (Figure 2). This type of topology still uses a set of edge switches for server connectivity, but it adds another set of edge switches for storage devices. Both sets of edge switches connect to a core switch through ISLs.

Figure 2.   Sample Edge-Core-Edge Design
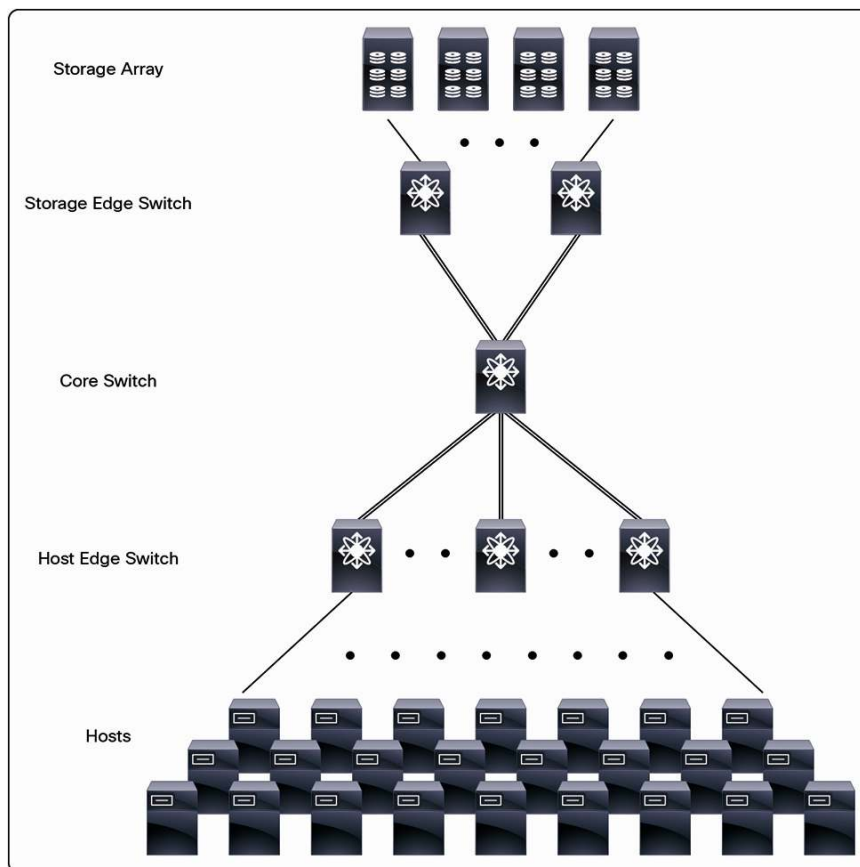
Network Considerations

When designing a large Cisco MDS 9000 Family SAN, you should consider the following:

- Fan-in, fan-out, and oversubscription ratios
- VSANs
- PortChannels
- ISLs
- Fabric login
- Zone type
- Smart zoning

Fan-In, Fan-Out, and Oversubscription Ratios

To efficiently and optimally use resources and to reduce deployment and management costs, SANs are designed to share array port, ISL, and line-card bandwidth. The terms used to describe this sharing include fan-in ratio, fan-out ratio, and oversubscription ratio. The term used depends on the point of reference being described. In general, the fan-in ratio is the ratio of host-port bandwidth to storage-array-port bandwidth, and the fan-out ratio is the ratio of storage-array-port bandwidth to host-port bandwidth.

Oversubscription is a networking term that is generally defined as the overall ratio of bandwidth between host and storage array ports. See Figure 3 for details.

Figure 3.   Fan-In; Fan-Out, and Oversubscription Ratios
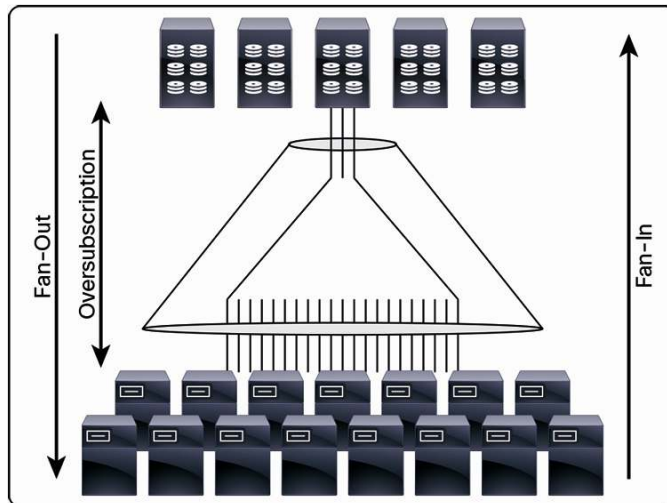

Virtual SANs

Cisco MDS 9000 Family switches support VSAN technology, which provides a simple and secure way to consolidate many SAN islands into a single physical fabric. Separate fabric services (for example, per-VSAN zoning, name services, domains, and role-based management) are provided for each VSAN, providing separation of both the control plane and the data plane.

VSANs have multiple use cases: for example, you can create a VSAN for each type of operating system (for instance, a VSAN for Microsoft Windows and for HP-UX) or create VSANs on the basis of business function (for instance, a development VSAN, a production VSAN, and a lab VSAN). VSAN 1 is created on the Cisco MDS 9000 Family switch by default and cannot be deleted. As a best practice, you should use VSAN 1 as a staging area for unprovisioned devices, and you should create other VSANs for production environments. With each VSAN having its own zones and zone sets, Cisco MDS 9000 Family switches enable secure, scalable, and robust networks.

Inter-Switch Links

An ISL is a connection between Fibre Channel switches. The number of ISLs required between Cisco MDS 9000 Family switches depends on the desired end-to-end oversubscription ratio. The storage port oversubscription ratio from a single storage port to multiple servers can be used to determine the number of ISLs needed for each edge-to-core connection. Figure 4 shows three examples of storage, server, and ISL combinations, all with the same oversubscription ratio of 8:1. The first example shows one 16-Gbps storage port with eight 16-Gbps server ports connected over one 16-Gbps ISL. The second example shows one 16-Gbps storage port with sixteen 8-Gbps server ports connected over one 16-Gbps ISL. The third example shows eight 16-Gbps storage ports with sixty-four 16-Gbps server ports connected over eight 16-Gbps ISLs. A 1:1 ratio of storage bandwidth to ISL bandwidth is recommended for SAN design. Additional ISL bandwidth can be added with additional ISLs to provide greater availability in the event of link failure.

Figure 4.   Number of ISLs Needed to Maintain Oversubscription Ratio
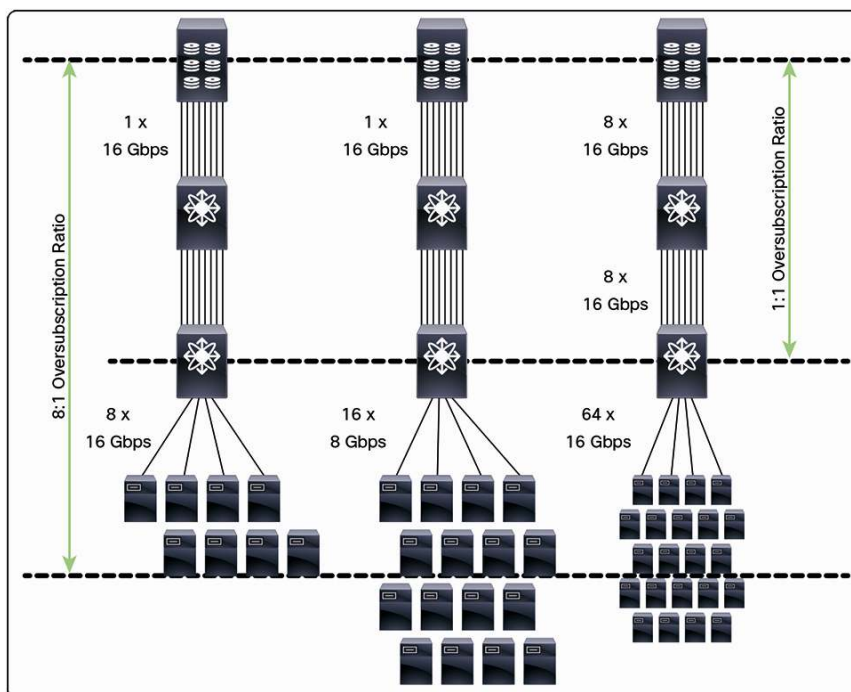
PortChannels

A PortChannel is an aggregation of multiple physical interfaces into one logical interface to provide higher aggregated bandwidth, load balancing, and link redundancy while providing fabric stability in the event of member failure. PortChannels can connect to interfaces across different switching modules, so a failure of a switching module does not bring down the PortChannel link.

A PortChannel has the following functions:

- It provides a single logical point-to-point connection between switches.
- It provides a single VSAN ISL (E-port) or trunking of multiple VSANs over an EISL (TE-port). EISL ports exist only between Cisco switches and carry traffic for multiple VSANs.
- It increases the aggregate bandwidth on an ISL by distributing traffic among all functional links in the channel. PortChannels can contain up to 16 physical links and can span multiple modules for added high availability. Multiple PortChannels can be used if more than 16 ISLs are required between switches.
- It performs load balancing across multiple links and maintains optimum bandwidth utilization. Load balancing is configured per VSAN (source ID and destination ID [SID and DID] or source ID and destination ID and exchange ID [SID and DID and OXID]).
- It provides high availability on an ISL. If one link fails, traffic is redistributed to the remaining links. If a link goes down in a PortChannel, the upper protocol is not aware of it. To the upper protocol, the link is still there, although the bandwidth is diminished. The routing tables are not affected by link failure.

Fabric Login

The number of actual physical ports in the fabric is greater than the number of end devices (server, storage, and tape ports) in the physical fabric. The Cisco MDS 9000 Family supports up to 10,000 fabric logins in a physical fabric, independent of the number of VSANs in the network. Typically when designing a SAN, the number of end devices determines the number of fabric logins. The increase in blade server deployments and the consolidation of servers as a result of server virtualization technologies affect the design of the network. With the use of features such as N-Port ID Virtualization (NPIV) and Cisco N-Port Virtualization (NPV), the number of fabric logins needed has increased even more (Figure 5). The proliferation of NPIV-capable end devices such as host bus adaptors (HBAs) and Cisco NPV-mode switches makes the number of fabric logins needed on a per-port, per-line-card, per-switch, and per-physical-fabric basis a critical consideration. The fabric login limits determine the design of the current SAN as well as its potential for future growth. The total number of hosts and NPV switches determine the number of fabric logins required on the core switch.

Figure 5.   Cisco NPV-Enabled Switches and Fabric Logins
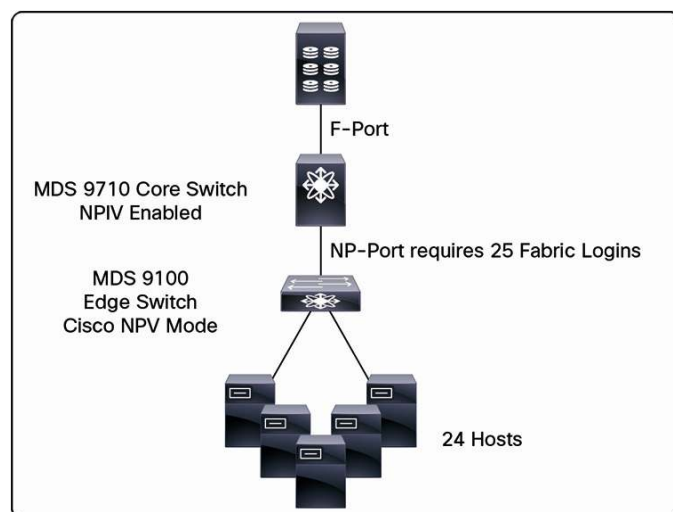
Note: Prior to NPIV and Cisco NPV, a single port supported a maximum of one fabric login. With NPIV and Cisco NPV-enabled switches, a single port can now support multiple fabric logins.

Zone Type

Each VSAN has only one active zone set, which contains one or more zones. Each zone consists of one or more members to allow communication between the members. The Cisco MDS 9000 Family SAN-OS and NX-OS Software provide multiple ways to identify zone members, but the commonly used ones are:

- World Wide Port Name (WWPN) of the device (most commonly used)
- Device alias, an easy-to-read name associated with a single device's WWPN

Depending on the requirements of the environment, the type of zone members used is a matter of preference. A recommended best practice is to create a device alias for end devices when you manage the network. The device alias provides an easy-to-read name for a particular end device. For example, a storage array with WWPN 50:06:04:82:bf:d0:54:52 can be given a device-alias name of Tier1-arrayX-ID542-Port2. In addition, with a device alias, when the actual device moves from one VSAN (VSAN 10) to a new VSAN (VSAN 20) in the same physical fabric, the device alias follows that device. Therefore, you do not need to reenter the device alias for each port of the moved array in the new VSAN.

Note: As a best practice for large SAN deployments, you should have more zones with two members rather than a single zone with three or more members. This practice is not a concern in smaller environments.

Smart Zoning

Smart zoning supports the zoning of multiple devices in a single zone by reducing the number of zoning entries that need to be programmed. With smart zoning, multiple member zones consisting of multiple initiators and multiple zones can be zoned together without increasing the size of the zone set. Smart zoning can be enabled at the zone level, zone-set level, zone-member level, or VSAN level.

Sample Use Case Deployments

Figures 6 and 7 shows two sample deployments of large-scale Cisco MDS 9000 Family fabrics.

Sample Deployment 1

Figure 6.   Use Case 1 Topology with Mix of 8-Gbps and 16-Gbps Hosts Connected to 16-Gbps Storage Ports
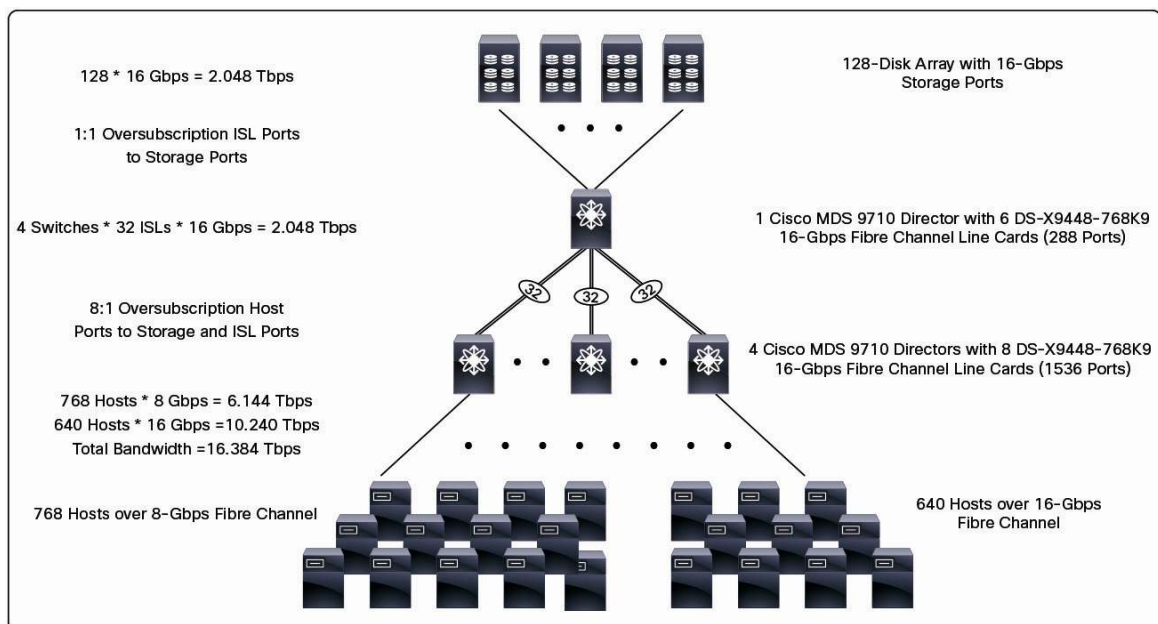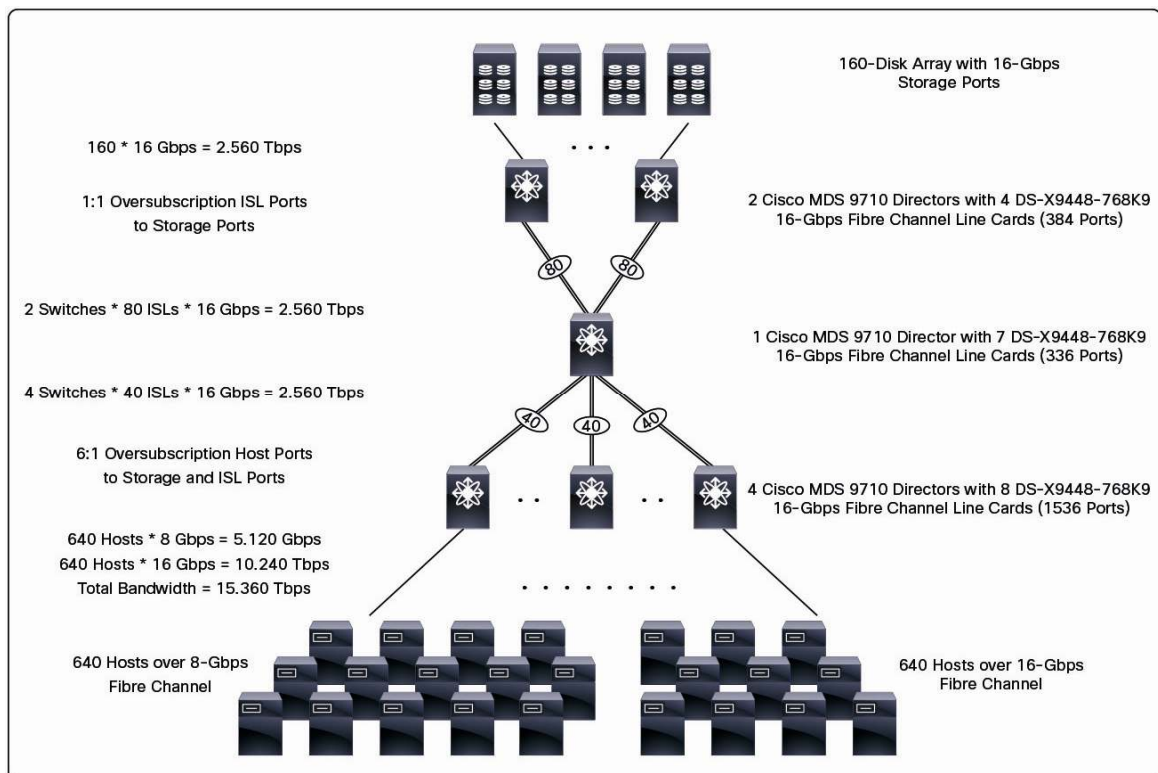

The deployment shown in Figure 6 allows scaling to more than 1500 devices in a single fabric. The actual production environment has 128 storage ports running at 16 Gbps and roughly 1400 host ports. The environment requires a minimum of 8:1 oversubscription in the network, which requires each host edge switch to have 2 Tbps of ISL bandwidth. The number of storage ports will not increase quite as rapidly, and the core switch has room to grow to add more host edge switches.

The network in this environment was managed as follows:

- A total of four VSANs were created:
  - VSAN 1 was used for staging new SAN devices
  - VSAN 100 was used for the development SAN
  - VSAN 200 was used for the lab SAN
  - VSAN 300 was used for the production SAN
- TACACS+ was used for authorization and authentication of Cisco MDS 9000 Family switches.
- Role-based access control (RBAC) was used to create separate administrative roles for different VSANs.
- Device aliases were used for logical device identification.
- Two member zones were used with the device alias.

Sample Deployment 2

Figure 7.   Use Case 2 Topology with Mix of 8-Gbps and 16-Gbps Hosts Connected to 16-Gbps Storage Ports

The deployment shown in Figure 7 scales to nearly 1500 devices in a single fabric. The actual production environment has 128 storage ports running at 16 Gbps and about 1300 host ports. The environment requires a minimum of 6:1 oversubscription in the network, which requires each host edge switch to have 2.560 Tbps of ISL bandwidth. Again, the number of storage ports will not increase quite as rapidly, and the core switch has room to grow to add more host edge switches.

The data center in this environment was managed as follows:

- A total of five VSANs were created:
  - VSAN 1 was used for staging new devices.
  - Four VSANs were based on business operations.
- TACACS+ was used for authorization and auditing of Cisco MDS 9000 Family switches.
- Separate administrative roles were created for the VSANs.
- A device alias was created for the environment.
- The dynamic port VSAN membership (DPVM) feature was enabled. This feature dynamically assigns VSAN membership to ports by assigning VSANs based on the device WWN. DPVM eliminates the need to reconfigure the port VSAN membership to maintain the fabric topology when a host or storage device connection is moved between two Cisco SAN switches or two ports within a switch. It retains the configured VSAN regardless of where a device is connected or moved.
- A mixture of two- and three-member zones was used.

Conclusion

With data centers continually growing, SAN administrators must design networks that both meet their current needs and can scale for demanding growth. Cisco MDS 9710 Multilayer Directors provide the embedded features that SAN administrators need for these tasks. SAN administrators deploying large Cisco SAN fabrics can use the design parameters and best practices discussed in this document to design optimized and scalable SANs.

For More Information

- Cisco MDS 9700 Series Multilayer Directors
- Cisco MDS 9710 Multilayer Director data sheet
- Cisco MDS 9700 Series Supervisor-1 Module data sheet
- Cisco MDS 9700 48-Port 16-Gbps Fibre Channel Switching Module data sheet
- Compare models: Learn about the similarities and differences among the models in this product series
- Data sheets and product literature
- At-a-Glance documents
- Data sheets
- Presentations
- White papers

Printed in USA