Cisco MDS 9000 Family Blade Switch Solutions Guide

Introduction

This document provides design and configuration guidance for administrators implementing large-scale blade server deployments using the Cisco® MDS Blade Switch Series for HP c-Class BladeSystem and IBM BladeCenter blade chassis. Cisco MDS 9000 Family blade switches provide enterprise-class feature sets derived mainly from MDS 9000 Family director products. This design document highlights some of the unique features of the Cisco MDS 9000 Family blade switches and how they can be used to build enterprise-class blade switch environments.

Audience

This document is intended for SAN and server administrators who are engaged deploying blade servers.

Cisco Fibre Channel Blade Switch Overview

Cisco Fibre Channel blade switches are built using the same hardware and software platform as Cisco MDS 9000 Family Fibre Channel switches and hence are consistent in features and performance. Architecturally, blade switches share the same design as the Cisco MDS 9000 Family fabric switches.

Cisco offers Fibre Channel blade switch solutions for HP c-Class BladeSystem and IBM BladeCenter.

Cisco Fibre Channel Blade Switch for HP c-Class BladeSystem

The Cisco MDS HP Blade Switch (Figure 1) is a Fibre Channel blade switch for HP c-Class BladeSystem. It comes in two models: the base 12-port model and a 24-port model. A 12-port license upgrade is available to upgrade the 12-port model to 24 ports. In the 12-port model, 8 ports are for server connectivity, and 4 ports are for SAN connectivity. In the 24-port model, 16 ports are for server connectivity, and 8 ports are for SAN connectivity.

Figure 1. MDS 9124e FC Blade Switch for HP c-Class Bladesystem

Cisco Fibre Channel Blade Switch for IBM BladeCenter

The Cisco MDS IBM Blade Switch (Figure 2) is a Fibre Channel blade switch for IBM BladeCenter. It comes in two models: a 10-port switch and a 20-port switch. A 10-port upgrade license is available to upgrade from 10 ports to 20 ports. In the10-port model, 7 ports are for server connectivity, and 3 ports are for SAN connectivity. In the 20-port model, 14 ports are for server connectivity, and 6 ports are for SAN connectivity.

Figure 2. MDS FC Blade Switch for IBM Bladecenter

Features

Some of the unique features of the Cisco MDS 9000 Family switches are:

- **Virtual SANs (VSANs):** VSANs partition the physical SANs into logical SANs, which can be managed like separate SANs or can consolidate different physical SANs into fewer Fibre Channel switches. Each VSAN can be managed by different administrators using VSAN-aware role-based access control (RBAC). The VSANs can be configured with different traffic priorities to enable service differentiation.
- **Inter-VSAN Routing (IVR):** IVR allows sharing of scarce resources among the VSANs while maintaining the isolation between them. IVR increases the utilization of both switches and SAN resources by enabling flexibility of any-to-any connectivity.

HP c-Class BladeSystem

The HP c-Class BladeSystem blade chassis is offered in two models: HP c3000 and c7000. The enclosure holds up to 16 servers and 8 interconnect bays for I/O connectivity. Refer to the HP c-Class enclosure specifications for more information about setting up connectivity: http://h18000.www1.hp.com/products/blades/components/enclosures/c-class/.

IBM BladeCenter

IBM BladeCenter is built on the IBM System x architecture offered in many models. The enclosure holds up to 14 blades and 4 IO modules. Refer to IBM Bladecenter specification at http://www-03.ibm.com/systems/bladecenter/.

Enterprise-Class SAN Connectivity for Blade Server Deployments Challenges

As the scale of Fibre Channel SAN blade switch deployments increases, it introduces challenges in providing agility for both virtual and physical resources. For example, the capability for server administrators to move, add, and change servers without involving the SAN administrator provides flexibility and helps simplify management of resources. This bring several challenges in offering enterprise-class SAN connectivity to blade deployments.

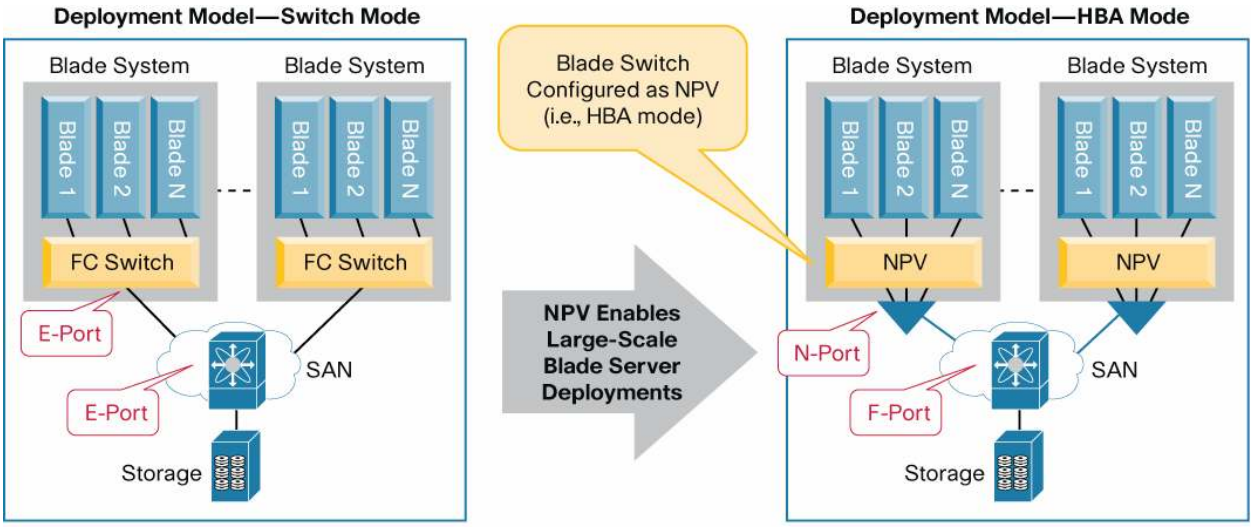
Management of a Large Number of Switches

Blade chassis consolidate many servers into one management unit, thereby saving space, energy, and management costs. Today most chassis are designed to consolidate only up to 16 or 20 servers in one chassis, which means that if a deployment needs 200 blade servers, for example, 100 blade chassis are required. Assuming that all the chassis require SAN access, 100 Fibre Channel switches are required (and 200 if redundancy is required). Hence, the blade deployment adds 100 small Fibre Channel switches per fabric, rather than 2 or 3 switches (assuming 24- or 48-port switches are used) if the servers were deployed without blades. Since every Fibre Channel switch needs a Fibre Channel domain ID, 100 domain IDs are also needed. Although, theoretically a VSAN (or fabric) could contain 239 domains, many Original Storage Manufacturers (OSMs) support a much smaller number (generally in the range of 40 to 50 switches).

Cisco N-Port Virtualizer

The Cisco N-Port Virtualizer (NPV) eliminates need for domain ID for a Fibre Channel switch. The blade switch works as a host bus adapter (HBA), aggregating many servers as shown in Figure 3. NPV gets Fibre Channel IDs (FCIDs) for attached devices from the SAN core switch to the SAN connectivity links using standards-based N-Port ID Virtualization (NPIV). In addition, the switches do not run many of the Fibre Channel protocols such as zoning and name server, making the management of switches much simpler. Refer to Cisco N-Port Virtualizer for Large-Scale Fibre Channel Blade Switch Deployments at http://www.cisco.com/en/US/prod/collateral/ps4159/ps6409/ps5989/ps7333/product_data_sheet0900aecd8068fa86.html.

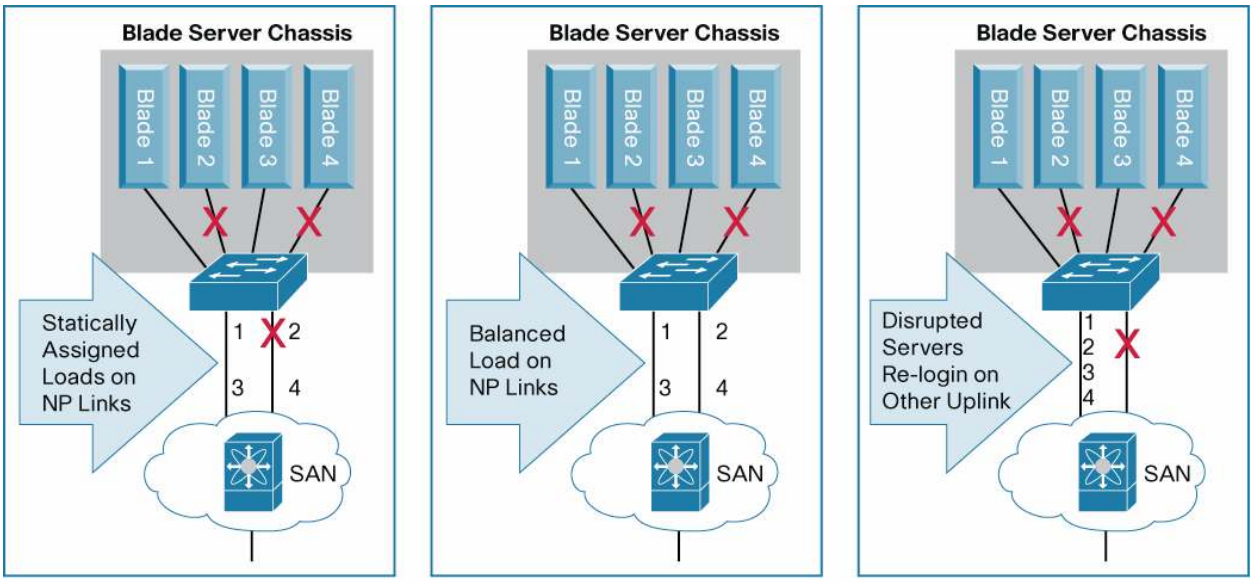
Figure 3. FC Blade Switch Working in NPV Mode



	Blade Switch Attribute	
Fibre Channel Switch Model (E-Port)	Deployment Model	HBA Model (N-Port)
One per Fibre Channel Blade Switch	Number of Domain IDs Used	None (Used Domain ID of Core Switch

The server uses the SAN connectivity links (uplinks) to log in to the SAN. The assignment of the servers to uplinks has a significant effect on overall uplink performance and utilization. With static assignment, any uplink failure will not only bring the servers but will also keep them down unless the uplink comes back up as shows in Figure 4.

Figure 4. Comparison of Uplink Assignment Models



The NPV automatic load-balancing feature distributes the server links among the available uplinks for optimal use of the uplink bandwidth as shown in Figure 4. Also, servers on the failed link are automatically transferred to an available uplink. A knob is also provided to redistribute the server loads when the failed links come back up.

The NPV traffic engineering feature allows customization of the uplink traffic by pinning servers to a particular uplink or set of uplinks.

In addition to addressing the manageability of large switches, NPV also solves the problem of interoperability. NPV uses standards-based NPIV and hence can interoperate with any Fibre Channel switch that supports NPIV.

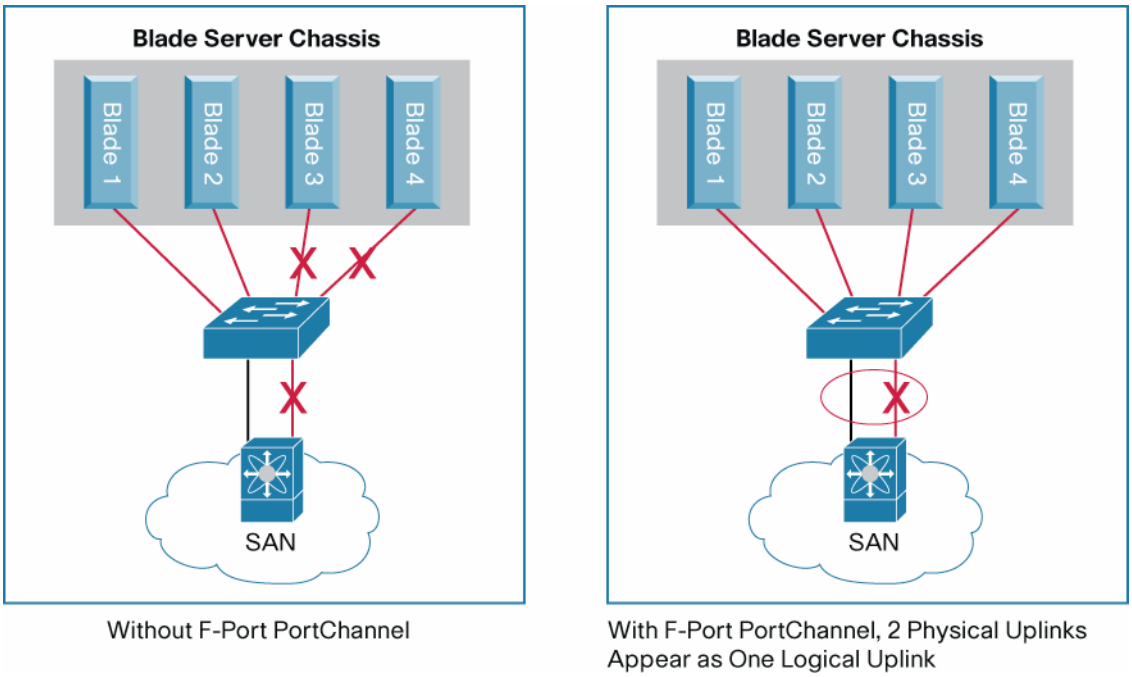
Need for Highly Available Network Resources in Large-Scale Blade Switch Deployments

Availability of network resources such as switches and network links are critical for any server deployment. Availability becomes even more critical in virtual server farms with a dynamic virtual machine scheduler, such as VMware Dynamic Resource Scheduler (DRS), in which smaller resource failures can cause churn in virtual machine scheduling, which can affect performance and functioning.

F-Port PortChannel

As mentioned earlier, uplink failures affect servers; although, with automatic load balancing features, the servers login again, the servers do experience traffic disruption. F-port PortChannel addresses this problem as illustrated in the Figure 5.

Figure 5. Enhancing Uplink Availability with F-PortChannels



As with Inter-Switch Link (ISL) PortChannels, F-port PortChannels are a logical link made up of one or more physical links. Multiple physical uplinks appear as one uplink, and the server login is on the PortChannel rather than the physical link. If a physical link goes down, the logical link does not go down as long as at least one physical link is up as illustrated in Figure 5.

In addition to providing high availability, PortChannels deliver higher aggregate bandwidth utilization by using all the links optimally.

A unique capability of Cisco's PortChannel implementation is that the physical links can be uniformly spread across multiple line cards, which makes it tolerate line-card failures, thereby enhancing uplink availability.

In Service Software Upgrade

Cisco MDS 9000 Family In Service Software Upgrade (ISSU) allows upgrading of the switch software with no negative effect on traffic. Hence, switch upgrading has no negative effect on business, and no planned downtime is needed to upgrade the blade switch infrastructure.

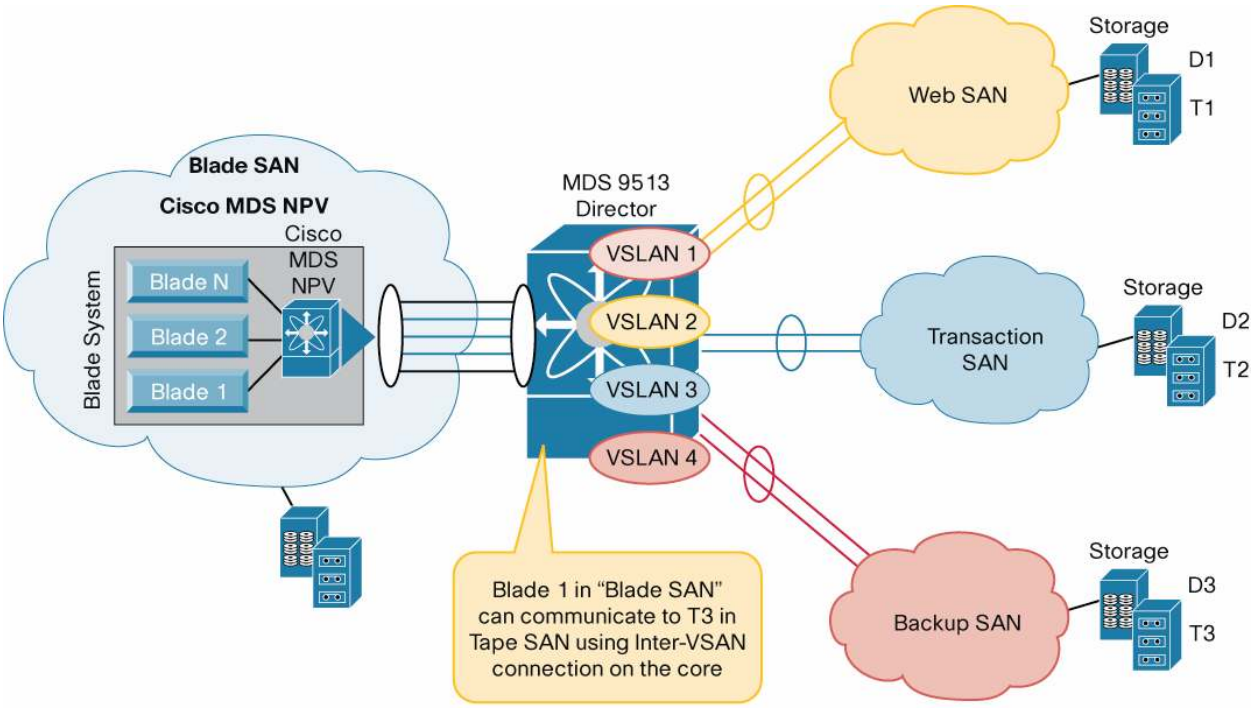
Isolation of SAN Resources While Sharing Scarce Resources

Blade chassis typically are managed by server administrators. Hence, the isolation of blade infrastructure from other SAN segments is important. The ability to segregate the larger SAN into smaller logical SANs to achieve isolation and management scalability is mandatory in blade switch deployments. Nevertheless, sharing of the existing valuable resources such as backup tape among the SAN islands is also very important.

VSANs and IVR

Cisco VSANs, the first and only switch virtualization solution, enable logical segregation of switches into logical SANs. In addition to traffic isolation, VSANs enhance switch utilization, enabling more scalable SANs. For example, the blade infrastructure can be deployed in its own logical VSAN to keep it separate from other SANs. The flexibility of VSAN-based quality of service (QoS) achieves optimized performance from the SAN. IVR enhances the flexibility of the Cisco Fibre Channel blade switches by enabling any-to-any connectivity between the shared resources in different VSANs (Figure 6).

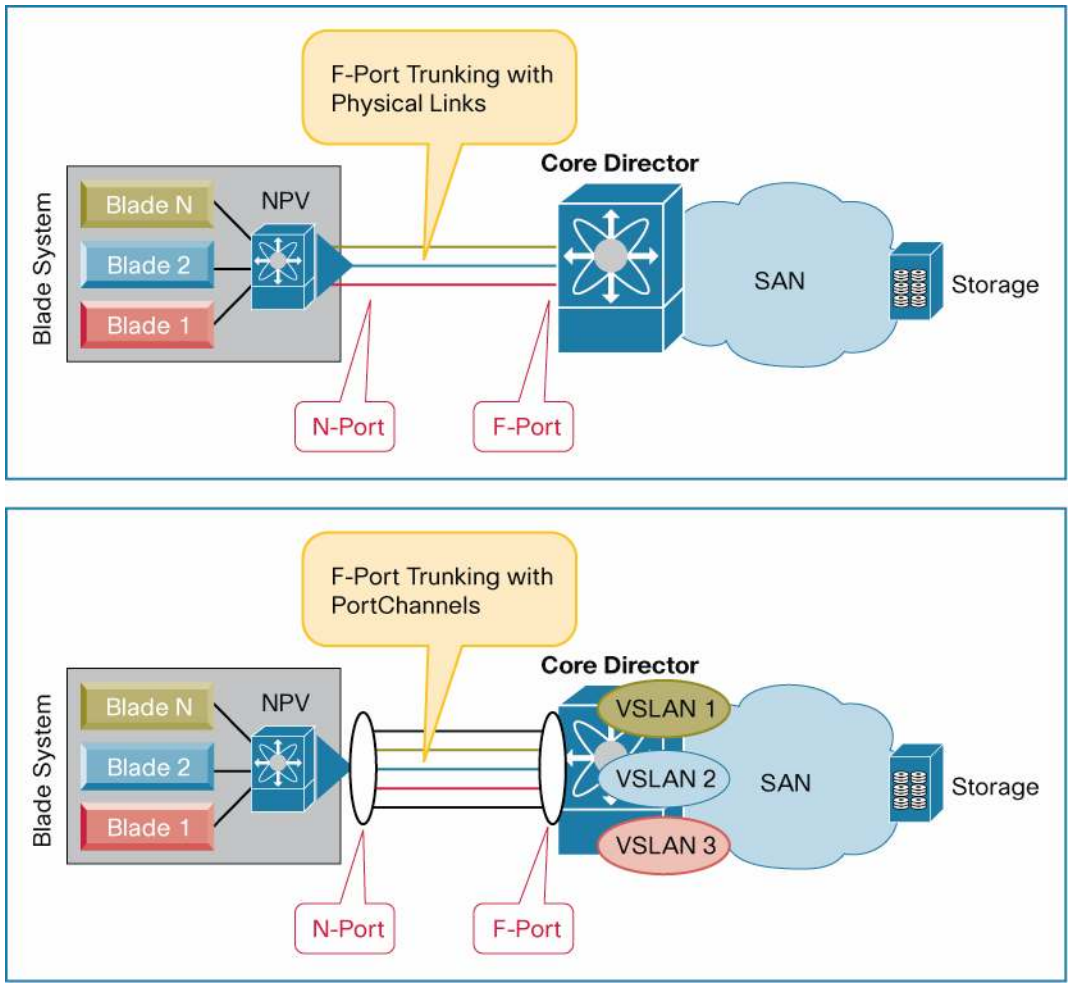
Figure 6. Any-To-Any Connectivity Using Inter-VSAN Routing



F-Port Trunking

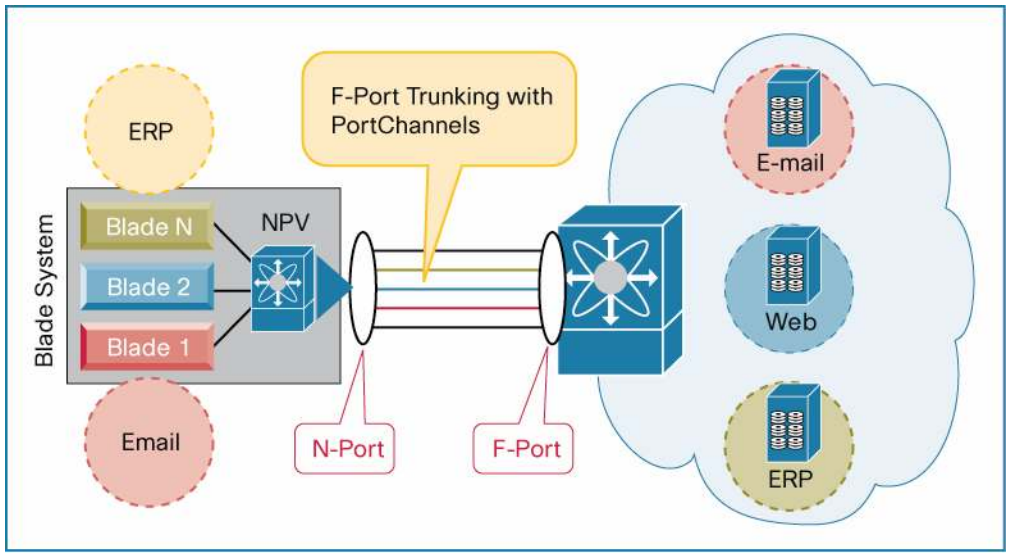
VSANs can used to consolidate physical SANs, using fewer switches, while maintaining traffic isolation. Multiple VSANs can be configured on the NPV switch. However, the uplinks can carry just one VSAN. This approach may not use the uplink bandwidth optimally, and with PortChannels, even this becomes very restrictive since the whole PortChannel can carry just one VSAN. With F-port trunking, the uplinks can carry multiple links, proving the flexibility of using PortChannels to maintain uplink availability (Figure 7).

Figure 7. F-Port Trunking with Physical and Port Channel Links



F-port trunking provides the capability to host different applications on blades in different VSANs while maintaining application and traffic protection. In combination with Cisco MDS 9000 Family VSAN-based QoS, F-port trunking enables the building of security and services differentiation for blade deployments (Figure 8).

Figure 8. Application Consolidation and Isolation Flexibility



SAN and Server Administrator Server Management Interaction Overhead

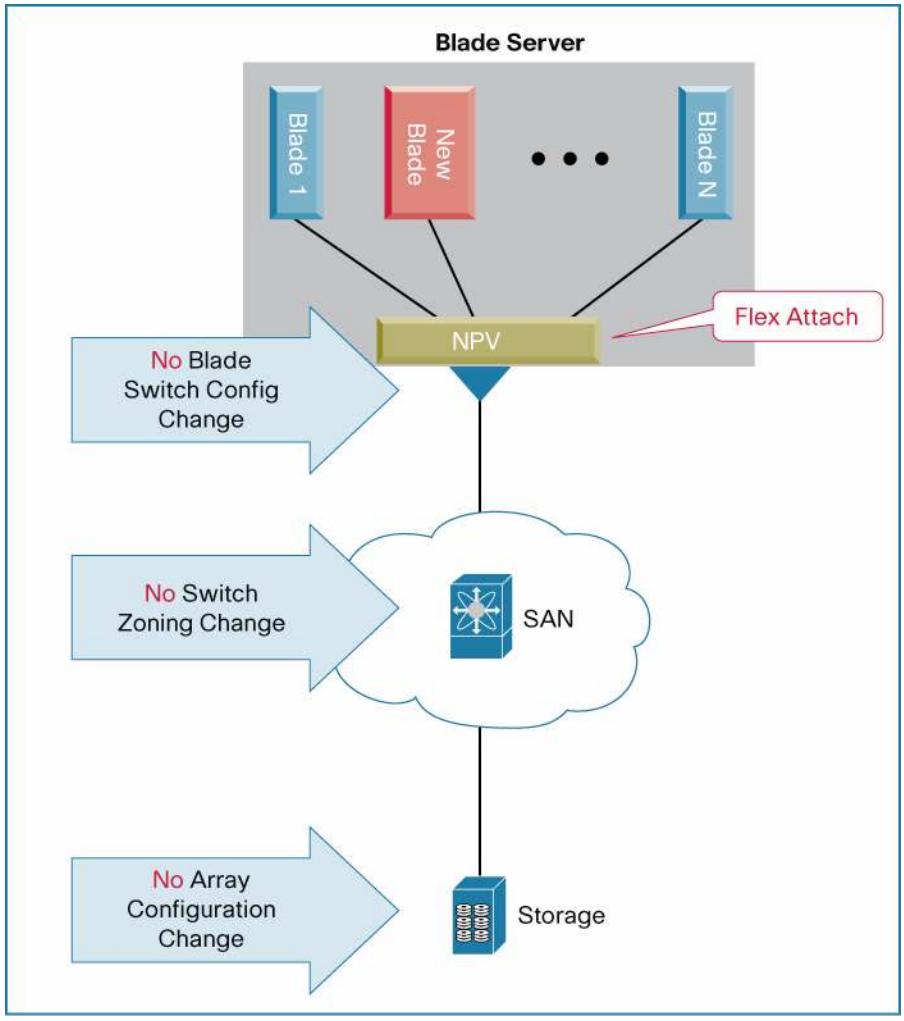
According to Share Survey[1], the two of top virtualization priorities for business are providing greater flexibility and agility in the IT environment and simplifying management and administration of resources. As the scale of Fibre Channel SAN blade switch deployment increases, it introduces challenge in providing agility for both virtual and physical resources. For example, flexibility enabling server administrators to move, add, and change servers without involving the SAN administrator helps simplify management.

Cisco FlexAttach

Cisco FlexAttach eliminates the need for SAN and server administrators to coordinate server changes, thereby increasing the availability of servers and the efficiency of operations. It assigns a virtual identity to attached servers in the form of a virtual port worldwide names (WWNs) to increase the agility of the server.

When the physical server is replaced, because of failure or for some other reason, no reconfiguration of SAN switches (for zoning) or storage arrays (logical unit number [LUN] masking) is required because the virtual identity remains with the port, and the new server gets the same virtual port WWN. Servers can even be moved across blade server chassis, with the virtual WWNs moved with the physical server (Figure 9).

Figure 9. FlexAttach Increases Operation Efficiency



[1] Conducted by Unisphere Research, a division of Information Today, Inc.

Thus, server administrators can now add, move, or replace servers without having to coordinate every change with the SAN administrator. FlexAttach provides fabricwide server mobility using Cisco Fabric Services, which distributes data across all the NPV switches.

FlexAttach can be used in following used cases:

- **SAN pre-provisioning:** Typically, SAN provisioning of end devices requires port WWNs, and so unless the servers are physically present, SAN provisioning is not possible. Servers on order are present for planned change control, and when the servers arrive, special change control needs to be planned. If virtual identities can be generated for required blade switch ports, the virtual port WWNs can be used to configure the SAN. When the servers arrive, they just need to plugged in.
- **Server replacement:** When servers are replaced, the port WWN changes, and hence SAN reconfiguration is needed. With virtual port WWNs, no such reconfiguration is needed. The replacement can be for the same port or to a spare port. In the case of spare ports, the virtual port WWN must be moved from the old port to the spare port.
- **Server move:** When a server is moved to another port or switch, the virtual port WWN can be moved along with it. No SAN reconfiguration is needed.

Cisco Fabric Services

Switch discovery and data distribution on NPV switches is achieved using Cisco Fabric Services. Normally, Cisco Fabric Services uses the Fibre Channel connectivity for fabricwide distribution. Because the NPV switch does not provide ISL connectivity to the rest of the SAN, Cisco Fabric Services uses IP. Any switch in NPV mode with IP connectivity will be included in the discovery and distribution. Cisco Fabric Services over IP requires Layer 2 multicast forwarding enabled along the path leading to all the NPV switches in the fabric. Since any switch in NPV mode is included in the distribution group, it is important to segregate the switches, especially in different physical fabrics, into different groups using Cisco Fabric Services regions. By default, an NPV switch is in Cisco Fabric Services region 201.

Use Cisco Fabric Services Regions to Separate Cisco Fabric Services Distribution Groups

In a dual-fabric deployment, the scope of discovery and distribution should be limited to one fabric using Cisco Fabric Services regions. For more information about Cisco Fabric Services regions, refer to the discussion of Cisco Fabric Services regions in the latest Cisco MDS 9000 Family configuration guide at http://www.cisco.com/en/US/docs/switches/datacenter/mds9000/sw/4_1/configuration/guides/cli_4_1/cfs.html#wp1306478. As guideline, leave the region on fabric A at the default 201 and use region 202 for fabric B.

Use Cisco Fabric Services Static List if Multicast Forwarding Is Not Supported

Cisco Fabric Services over IP switch discovery needs multicast forwarding support in entire path connecting the blade switches. If multicast forwarding is disabled, for example, in the current IBM blade center chassis, the Cisco Fabric Services static list should be used. The Cisco Fabric Services static list is a mechanism for manually adding all the blade switches on one of the switches. Cisco Fabric Manager provides a wizard to display all the switches in the fabric and add the peers to the list (Figure 10).

Figure 10. Using CFS Static List If Multicast Forwarding is Disabled

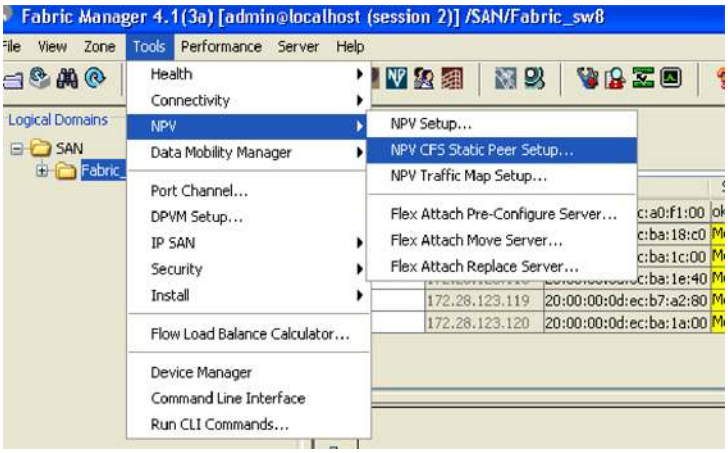
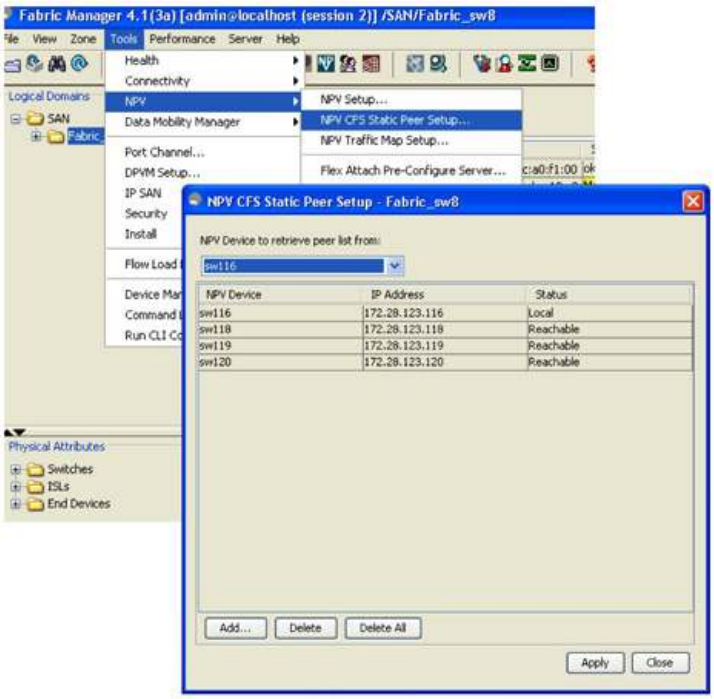


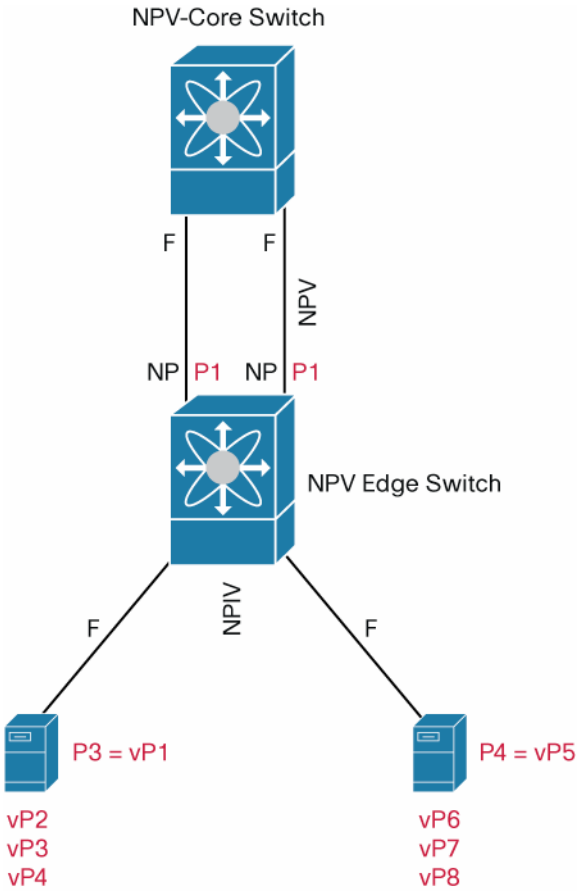
Figure 11. Using CFS Static List If Multicast Forwarding is Not Enabled



Virtual Machines with Port WWNs (Nested NPIV)

If the virtual machine has SAN identities, they can controlled and monitored in the SAN like any other end device. Virtual machines can be given SAN identities by assigning port WWNs and getting different Fibre Channel IDs (FCIDs) in the SAN using NPIV. Cisco MDS 9000 Family blade switches support NPIV and can support multiple logins on the same F-port. This capability is called nested NPIV since end devices use NPIV to log in to NPV, and NPV uses NPIV to connect to the SAN core as illustrated in Figure 12.

Figure 12. Support for Port WWNs for Virtual Machine Hosted on Blade Servers



It is important to follow the guidelines from the virtual machine vendors for assigning port WWNs to virtual machines. For example, VMware requires the use of Raw Device Mode (RDM) instead of Virtual Machine File System (VMFS) to get access to raw LUNs.

Deployment and Management

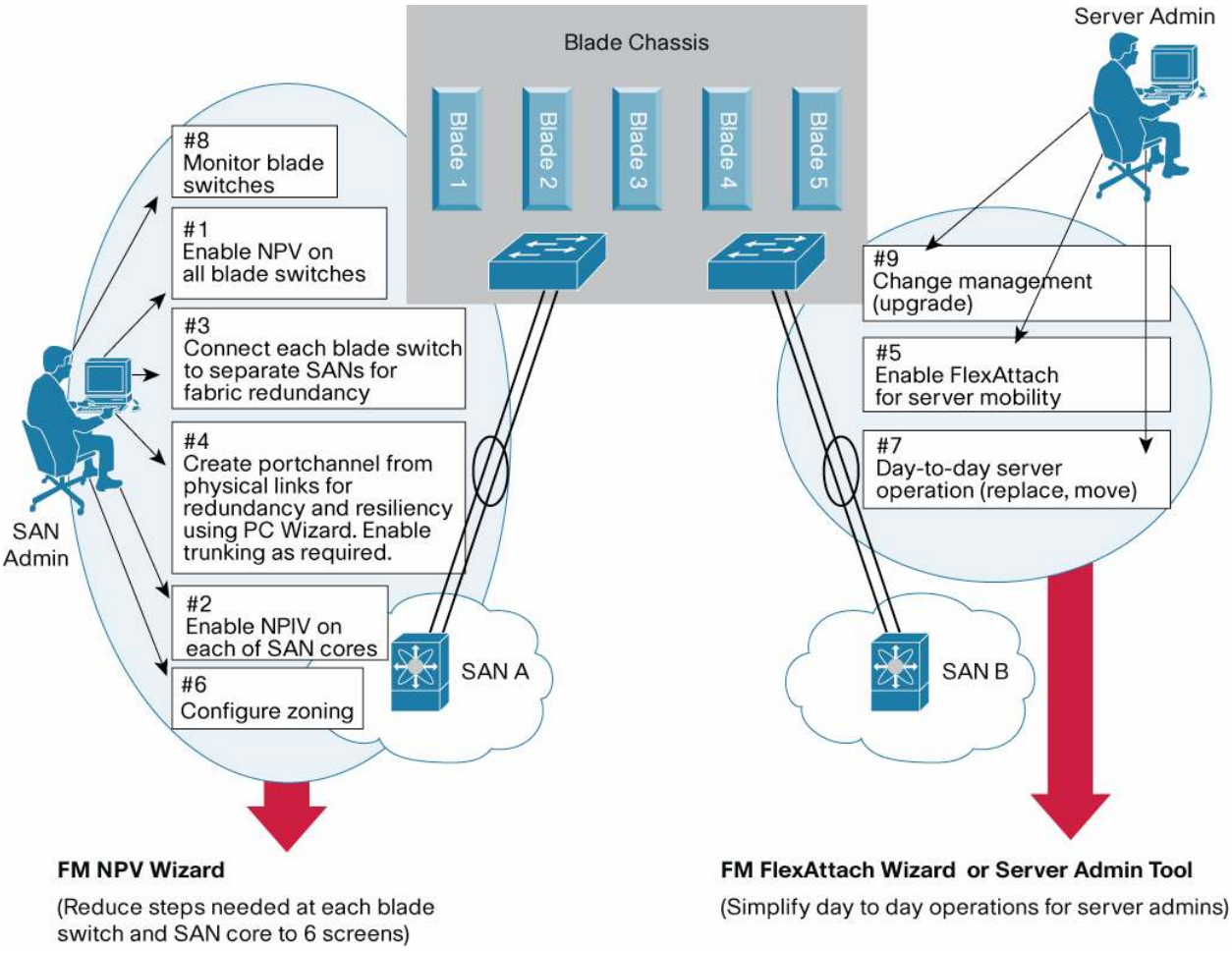
Cisco Fabric Manager provides comprehensive, enterprise-class feature deployment configuration and monitoring tools. Cisco Fabric Manager also has performance and status monitoring capabilities, with its fabric map, which shows all the switches and devices in the SAN, and its zoning wizard, which configures zoning in a few simple steps. Figure 13 provides screen captures for fabric view map, device manager and real-time monitoring.

Figure 13. Integrated and Comprehensive Management Tools

The blade switches that are part of the SAN are traditionally managed by the SAN administrator. However, because they are housed inside a blade chassis, they could also be managed by the server administrator. The complexity of the switch may be a barrier for the server administrator. However, the fact that NPV is not a Fibre Channel switch but more an HBA aggregator should address that concern. The question of who manages the blade switches then depends on organizational policies and processes. The Cisco blade switch management strategy is to provide a very comprehensive and flexible toolset that can adapt to the organizational policies and processes.

Deployment of a large number blade switches can be accomplished using the Cisco Fabric Manager NPV wizard. Cisco Fabric Manager also provides tools to deploy the PortChannels and trunking ports. Depending on the model of internal management, either the Cisco Fabric Manager FlexAttach wizard or Server Admin tool can be used to provision the FlexAttach feature. Figure 14 shows the steps needed to deploy enterprise-class SAN connectivity for blade environments using unique features provided by Cisco MDS 9000 Family blade switches.

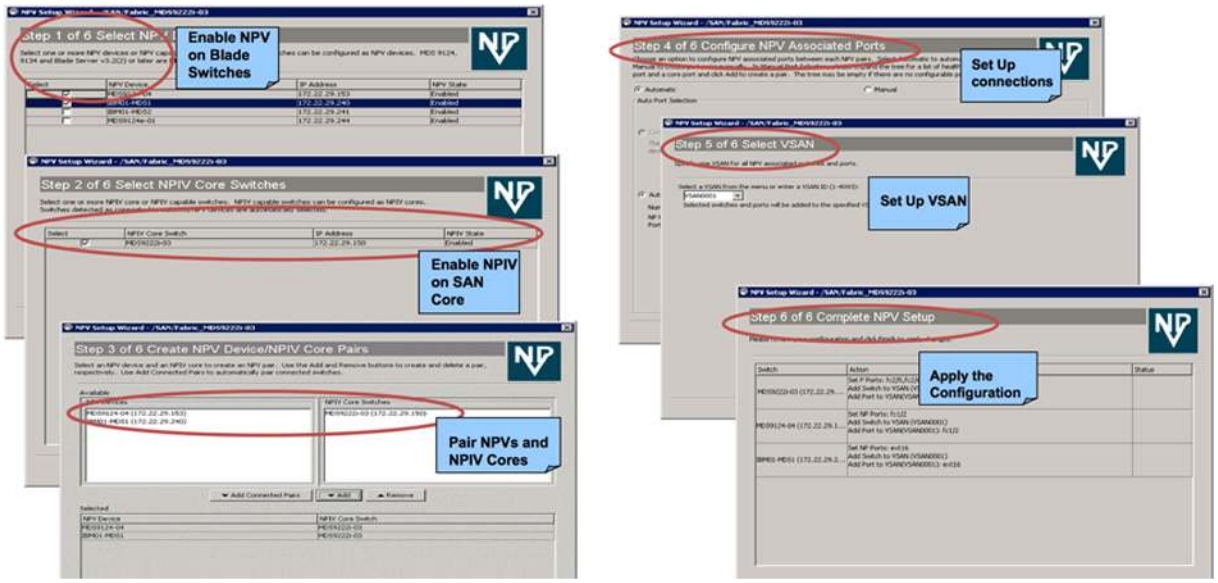
Figure 14. Illustration of Steps Needed for SAN Connectivity Deployment for Blade Servers



Cisco Fabric Manager provides the following tools to help with this deployment:

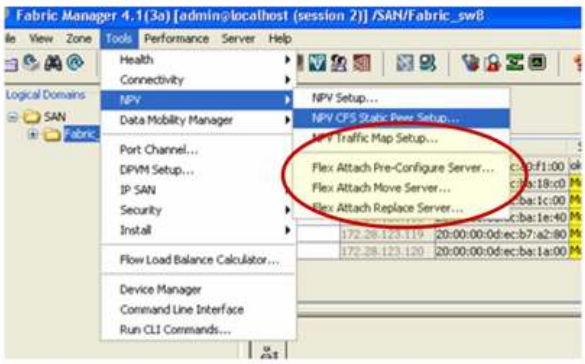
- **NPV wizard:** This wizard provides screens to list all the NPIV-capable switches, enable NPIV, list all the blade switches, enable NPV, and provide connectivity between NPV and NPIV core switches. Figure 15 shows screen shots for the steps.

Figure 15. Screen Shots for NPV Wizard



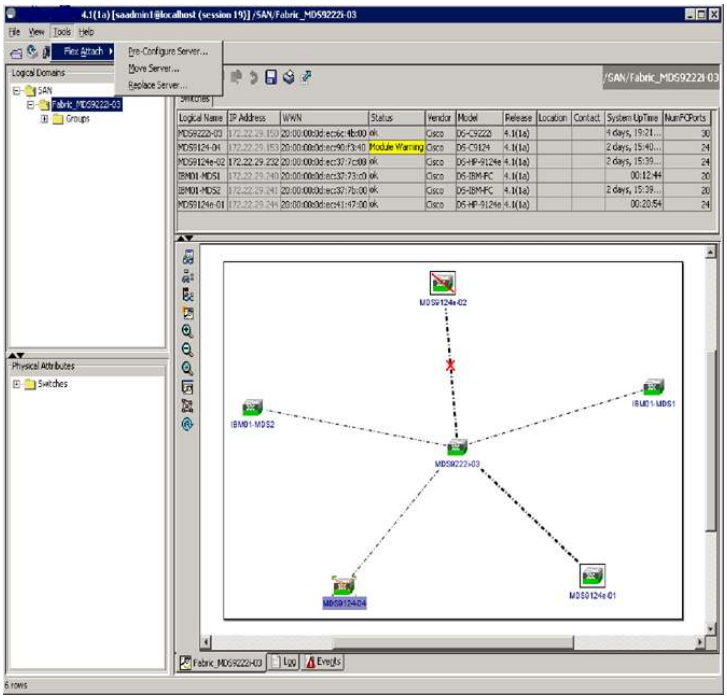
- **PortChannel wizard:** This wizard provides screens to list the uplink connections between the NPV and NPIV switches and help create PortChannels. Refer to the Cisco Fabric Manager user guide for more information about the PortChannel wizard.
- **FlexAttach wizards:** This set of wizards provides screens for addressing each of the FlexAttach use cases listed earlier in the “Cisco FlexAttach” section. In this case, blade switch management is assumed to be performed by the SAN administrator (Figure 16).

Figure 16. FlexAttach Wizards



- **Server Admin tool:** In organizations in which server administrators manage the blade switches, Cisco Fabric Manager may not be the right tool. NPV initial deployment is performed by SAN administrators, and day-to-day operations are handled by server administrators. For server administrators, the Server Admin tool is the more relevant tool. In addition to providing wizards for FlexAttach use cases, it provides a map of all the NPV switches along with the SAN cores (Figure 17).

Figure 17. Server Admin Tool



Blade Server SAN Connectivity Design

Cisco MDS 9000 Family blade solutions provide a high level of availability and flexibility for building enterprise-class SAN connectivity for blade deployments. This section introduces a model design and discusses several considerations. To achieve enterprise-class SAN connectivity, the following baseline requirements are needed:

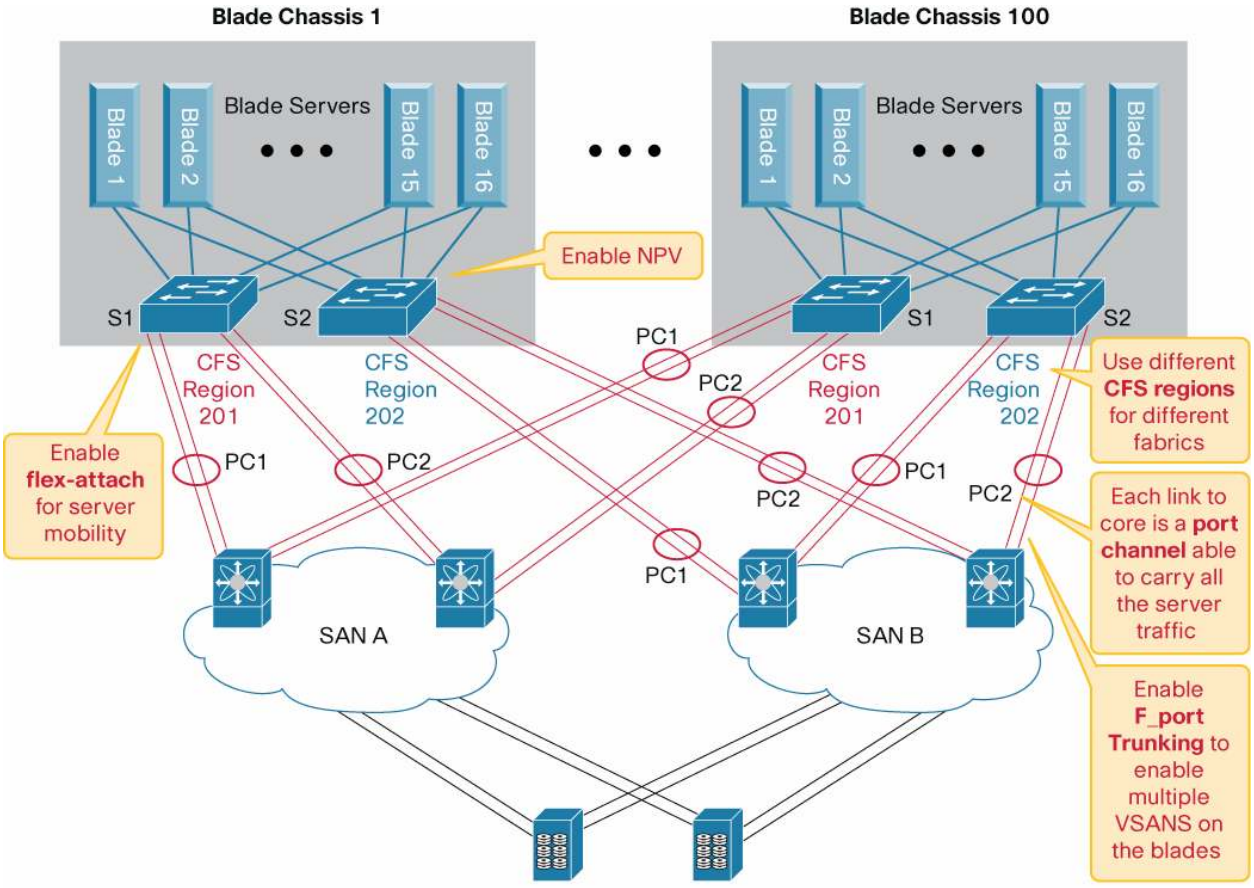
- Set blade switches to NPV mode.
- Make sure the blade server is dual homed.
- Connect different blade switches to physically separate SAN fabrics.
- Use Cisco Fabric Services regions to assign the blade switches to different Cisco Fabric Services distribution groups to isolate the switches into physically separate fabrics. Make sure that switches in the same fabric have the same region numbers.
- Use PortChannels to achieve the highest level of availability. On the SAN core, choose PortChannel members from different line cards. Refer to the “Scalability Considerations” section later in this document for information about how to choose the members on the NPV switches. The best approach is to use all six uplink ports, although by default only three are available.
- Use FlexAttach on the NPV switches to enhance the flexibility and efficiency in the data center.
- Use either a Cisco MDS 9509 or 9513 Multilayer Director or Cisco MDS 9222i Multiservice Modular Switch as the SAN core.
- If you need to isolate the blade environment from other SAN environments, consolidate different SANs using VSANs to achieve the isolation. If a resource is required, use IVR to zone only required device pairs. Refer to the “VSANs and IVR” section earlier in this document.

Following are optional best-practices requirements:

- If application consolidation and traffic isolation are required simultaneously, use multiple VSANs on a blade switch to achieve isolation. Use F-port trunking on the uplinks to carry the multiple VSANs.
- If virtual machines need to controlled and monitored by the SAN, enable NPIV on the blade switches.

Figure 18 shows a model design that meets the requirements described here.

Figure 18. Suggested SAN Connectivity Design for Blade Servers



Scalability Considerations

NPV aggregates many device logins on a single port on the SAN core switch, and nested NPIV aggregates more logins in the form of virtual machines. Hence, you should understand some of the restrictions on the SAN core and the blade switches while deploying the blade switches. Refer to the document on large SAN design best practices at http://www.cisco.com/en/US/prod/collateral/ps4159/ps6409/ps5990/white_paper_C11-515630.html to learn more about designing the SAN core. To achieve the best performance, the design should try to distribute the NPV blade switches across different line cards and switches. This document also lists the maximum number of logins per port and per switch on both the Cisco MDS IBM and HP Blade Switches.

PortChannels on NPV

To understand the implications of PortChannels, consider the port group mapping for both the Cisco MDS IBM and HP Blade Switches. Port groups are functional blocks that share the same resources within the switch architecture. The best performance is achieved if the PortChannel members are in a port group. Use Figures 19 and 20 as guidelines for deploying ports on both Cisco MDS IBM and HP Blade Switches.

Figure 19. Suggested Port Channel Member Grouping for MDS Blade Switches for HP c-Class Enclosures

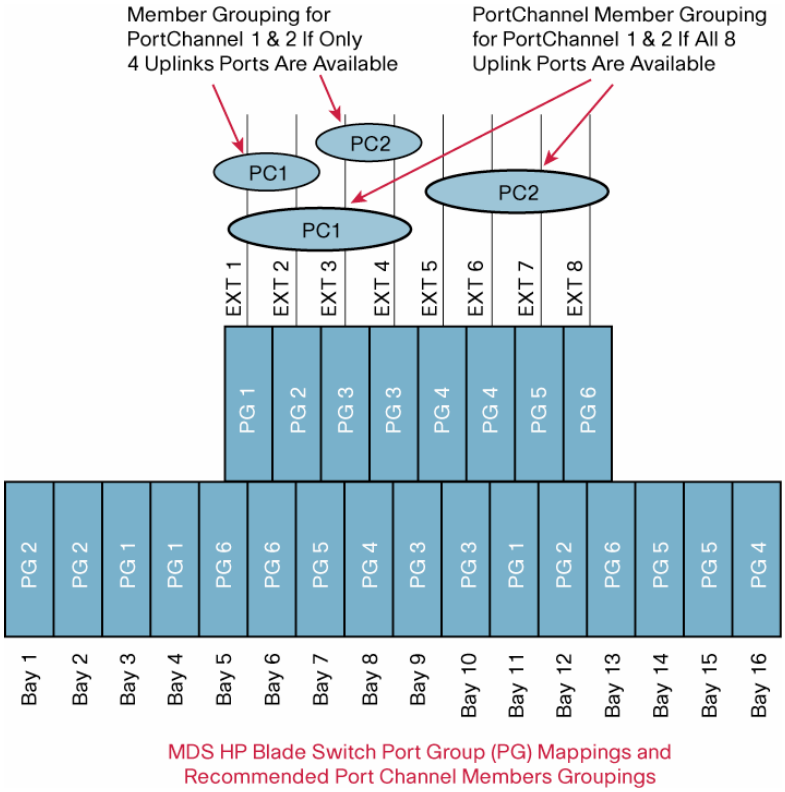
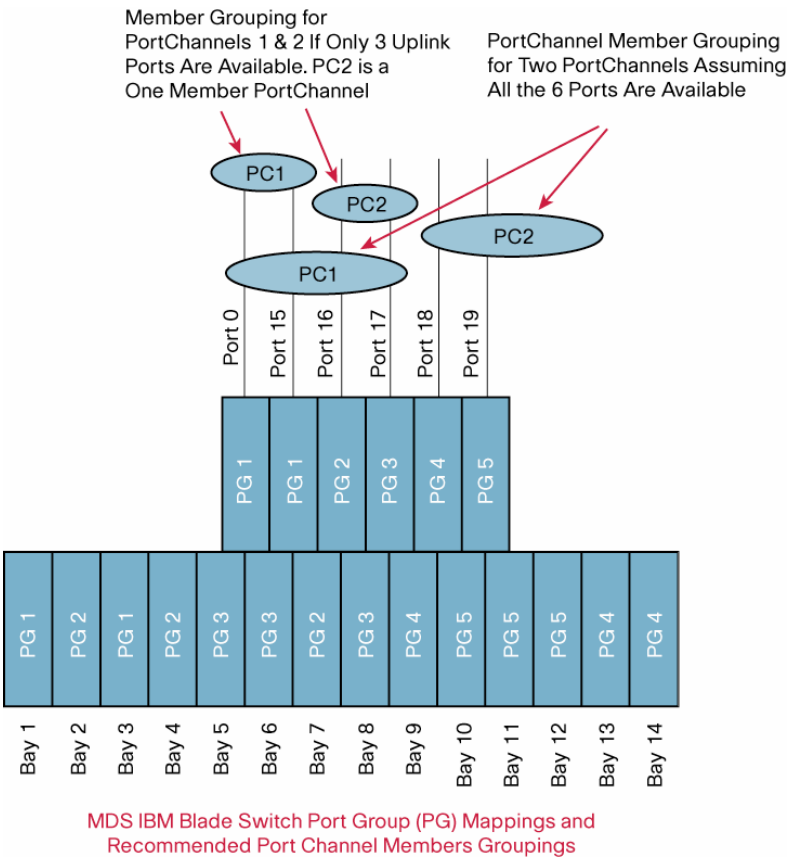


Figure 20. Suggested Port Channel Member Grouping for MDS Blade Switches for IBM Bladecenter Enclosures



Conclusion

The Cisco MDS Blade Switch Series for IBM and HP blade enclosures provide enterprise-class features. In addition, the Cisco Fabric Manager Management toolset helps administrators deploy and manage the features and the switches.

For More Information

MDS Fibre Channel Blade Switch for IBM BladeCenter and HP c-Class BladeSystem

http://www.cisco.com/en/US/prod/collateral/ps4159/ps6409/ps5987/product_data_sheet0900aecd805f187a.html

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 C07-539900-00 05/09