

Cisco Unified Computing System Ethernet Switching Modes

What You Will Learn

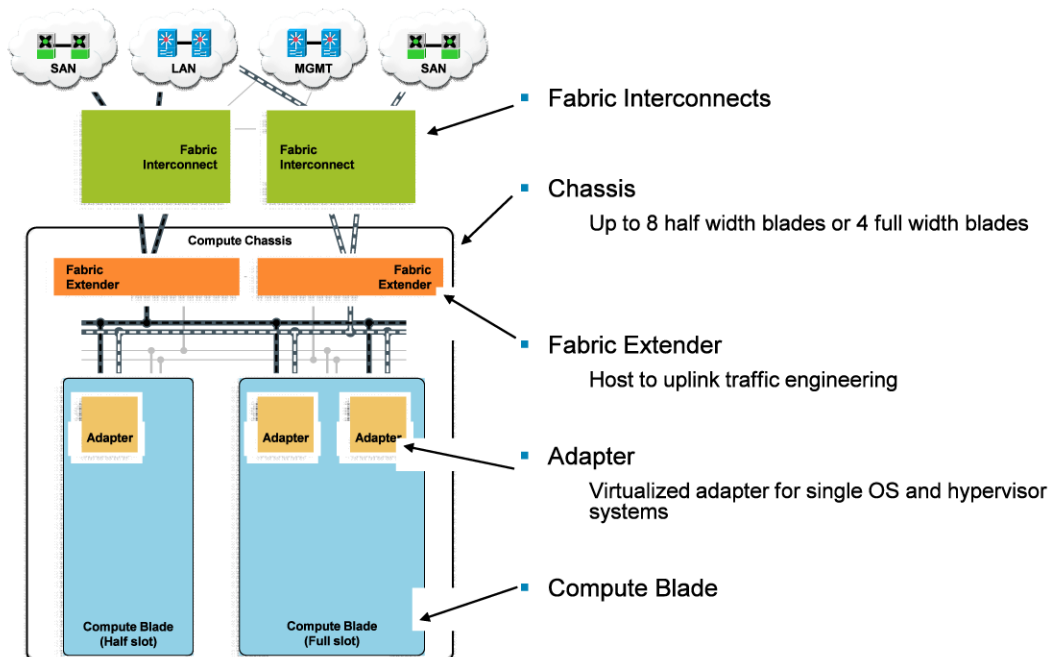
In Cisco Unified Computing System™ (Cisco UCS™) environments, two Ethernet switching modes determine the way that the fabric interconnects behave as switching devices between the servers and the network. In end-host mode, the fabric interconnects appear to the upstream devices as end hosts with multiple links. In end-host mode, the switch does not run Spanning Tree Protocol and avoids loops by following a set of rules for traffic forwarding. In switch mode, the switch runs Spanning Tree Protocol to avoid loops, and broadcast and multicast packets are handled in the traditional way. This document describes these two switching modes and discusses how and when to implement each mode.

Cisco UCS Architectural Overview

A fully redundant Cisco Unified Computing System consists of two independent fabric planes: Fabric A and Fabric B. Each plane consists of a central fabric interconnect (Cisco UCS 6100 or 6200 Series Fabric Interconnects) connected to an I/O module (fabric extender) in each blade chassis. The two fabric interconnects are completely independent from the perspective of the data plane; Cisco UCS can function with a single fabric interconnect if the other fabric is offline or not provisioned (Figure 1).

Figure 1. Cisco UCS Components (Logical)

Blade Server



All network endpoints, such as host bus adapters (HBAs), and management entities such as Cisco Integrated Management Controllers (IMCs; formerly referred to as baseboard management controllers, or BMCs), are dual-connected to both fabric planes and thus can work in an active-active configuration.

Virtual port channels (vPCs) are not supported on the fabric interconnects, although the upstream LAN switches to which they connect can be vPC or virtual switching system (VSS) peers.

Cisco UCS 6100 and 6200 Series Fabric Interconnect Switching Modes

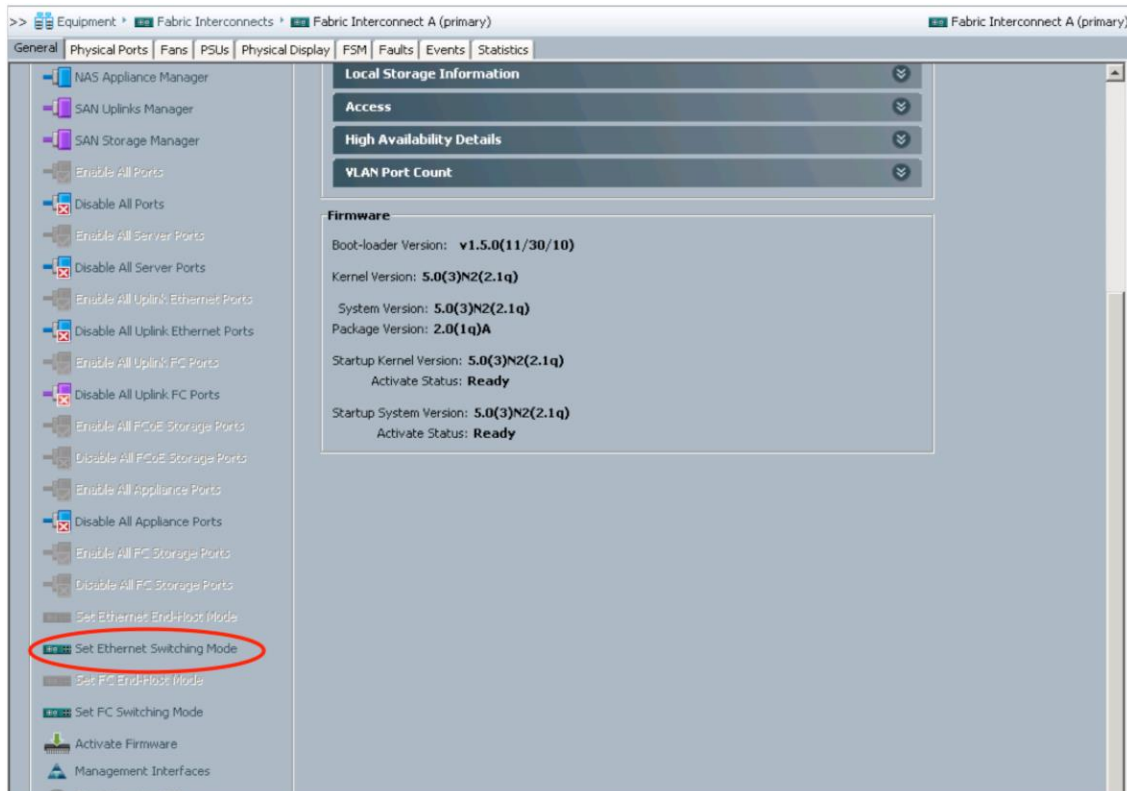
As shown in Figure 1, Cisco UCS supports connectivity to Ethernet LANs and Fibre Channel SANs to enable network and storage I/O from servers. The external interface operating modes to the LAN and the SAN balance the goals of administrative simplicity and utility in common deployment scenarios. The two operating modes supported on the fabric interconnect for Ethernet are:

- Ethernet end-host mode, sometimes referred to as Ethernet host virtualizer
- Traditional Ethernet switch mode

Note: This document does not discuss the operating mode available on the Fibre Channel ports.

For Ethernet, end-host mode is the default mode of operation. A change in the operating mode requires a fabric interconnect reboot to effect the change (Figure 2).

Figure 2. Setting the Switching Mode



Ethernet ports on the fabric interconnects are unconfigured by default. The ports can be configured to be:

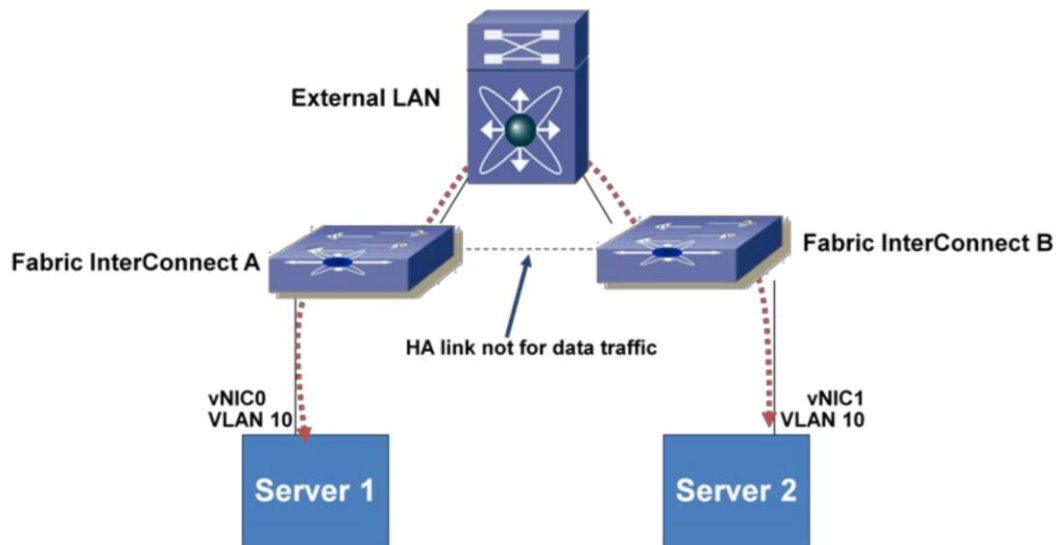
- Uplink ports
- Server ports
- Appliance ports
- Monitor ports

A port must be explicitly defined as a specific type, and this type defines the port behavior. For example, discovery of components such as fabric extenders or blades is performed only on server ports. Similarly, uplink ports are automatically configured as IEEE 802.1Q trunks for all VLANs defined on the fabric interconnect.

Note: In either Ethernet switching mode, a fabric interconnect does not require an upstream switch for Layer 2 traffic between two servers connected to it on the same fabric.

An external switch is required for switching Layer 2 traffic between servers if virtual network interface cards (vNICs) belonging to the same VLAN are mapped to different fabric interconnects (Figure 3).

Figure 3. External Switch Required for Switching Layer 2 Traffic Between vNICs on Different Fabric Interconnects



Fabric Failover for Ethernet: High-Availability vNIC

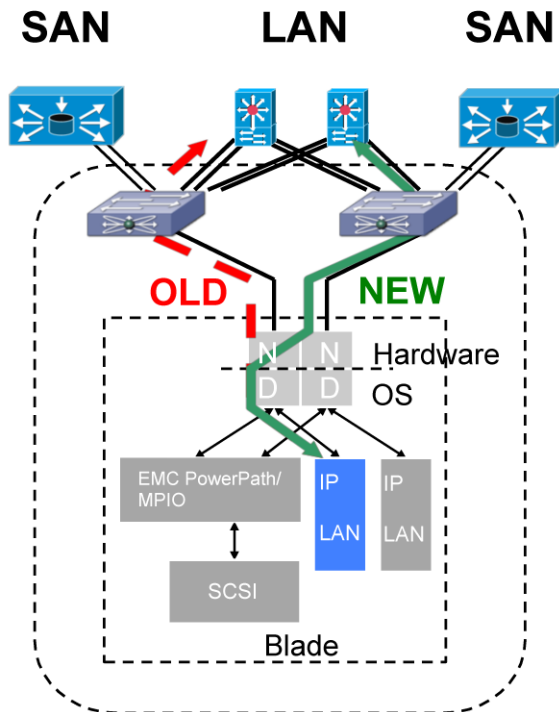
To understand the switching mode behavior, you need to understand the fabric-based failover feature for Ethernet in Cisco UCS.

Each adapter in Cisco UCS is a dual-port adapter that connects to both fabrics (A and B). The two fabrics in Cisco UCS provide failover protection in the event of planned or unplanned component downtime in one of the fabrics. Typically, host software—such as NIC teaming for Ethernet and EMC PowerPath or multipath I/O (MPIO) for Fibre Channel—provides failover across the two fabrics (Figure 4).

A vNIC in Cisco UCS is a host-presented PCI device that is centrally managed by Cisco UCS Manager. The fabric-based failover feature, which you enable by selecting the high-availability vNIC option in the service profile

definition, allows network interface virtualization (NIV)-capable adapters (Cisco M71KR-Q QLogic and M71KR-E Emulex Converged Network Adapters [CNAs], Cisco Virtual Interface Card (VIC) 1280, and Cisco M81KR VIC) and the fabric interconnects to provide active-standby failover for Ethernet vNICs without any NIC-teaming software on the host. Host software (EMC PowerPath or MPIO) is still required to handle failover for Fibre Channel virtual HBAs (vHBAs).

Figure 4. Cisco UCS Fabric-Based Failover



For unicast traffic failover, the fabric interconnect in the new path sends gratuitous Address Resolution Protocols (gARPs). This process refreshes the forwarding tables on the upstream switches.

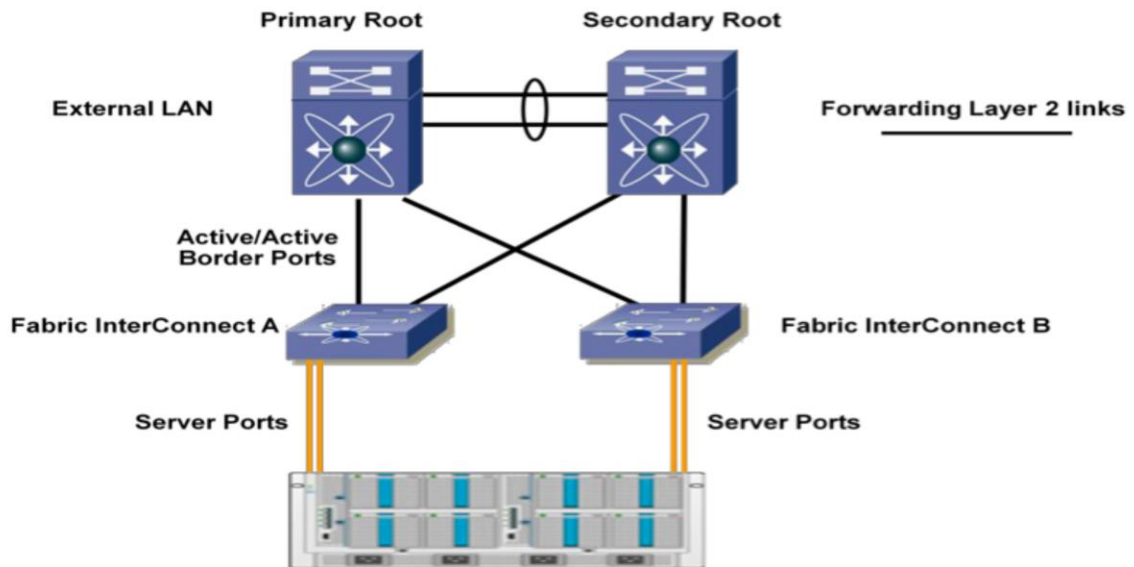
For multicast traffic, the new active fabric interconnect sends an Internet Group Management Protocol (IGMP) Global Leave message to the upstream multicast router. The upstream multicast router responds by sending an IGMP query that is flooded to all vNICs. The host OS responds to these IGMP queries by rejoining all relevant multicast groups. This process forces the hosts to refresh the multicast state in the network in a timely manner.

Cisco UCS fabric failover is an important feature because it reduces the complexity of defining NIC teaming software for failover on the host. It does this transparently in the fabric based on the network property that is defined in the service profile.

Ethernet End-Host Mode

In end-host mode, Cisco UCS presents an end host to an external Ethernet network. The external LAN sees the Cisco UCS fabric interconnect as an end host with multiple adapters (Figure 5).

Figure 5. Active-Active Links in End-Host Mode



End-host mode features include:

- Spanning Tree Protocol is not run on both the uplink ports and the server ports.
- MAC address learning occurs only on the server ports; MAC address movement is fully supported.
- Links are active-active regardless of the number of uplink switches.
- The system is highly scalable because the control plane is not occupied.

Server links (vNICs on the blades) are associated with a single uplink port, which may also be a PortChannel. This association process is called pinning, and the selected external interface is called a pinned uplink port. The pinning process can be statically configured when the vNIC is defined or dynamically configured by the system. In end-host mode, pinning is required for traffic flow to a server.

Static pinning is performed by defining a pin group and associating the pin group with a vNIC (Figures 6 and 7).

Figure 6. Creating a LAN Pin Group for Static Pinning

Create LAN Pin Group

Name:

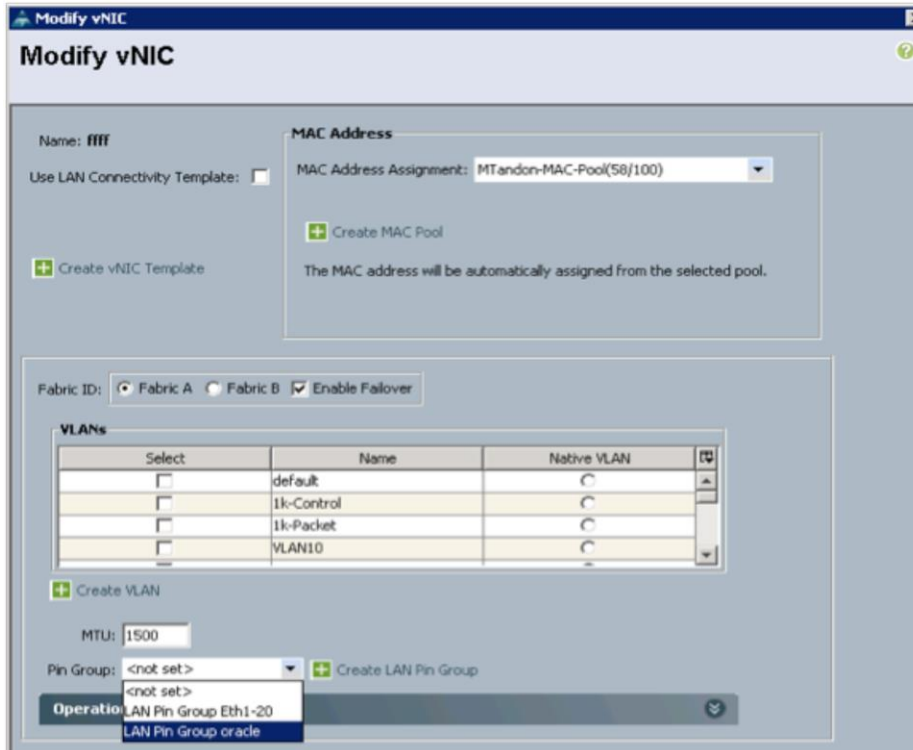
Description:

Targets

- Fabric A
Interface:
- Fabric B
Interface:

OK Cancel

Figure 7. Static Pinning: Associating a Pin Group with a vNIC



Static pinning should be used in scenarios in which a deterministic path is required. When the target (as shown on Figure 6) on Fabric Interconnect A goes down, the corresponding failover mechanism of the vNIC goes into effect, and traffic is redirected to the target port on Fabric Interconnect B.

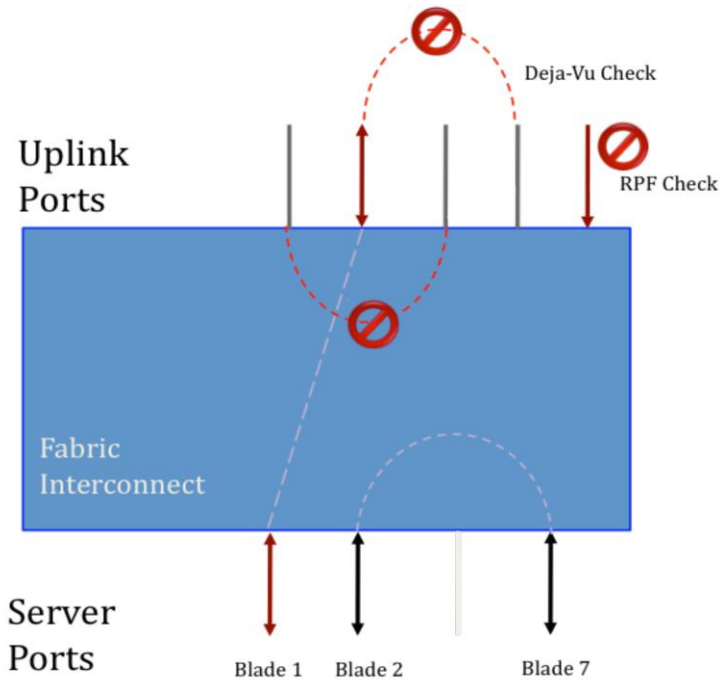
If the pinning is not static, then the vNIC is pinned to an operational uplink port on the same fabric interconnect, and the vNIC failover mechanisms are not invoked until all uplink ports on that fabric interconnect fail. In the absence of Spanning Tree Protocol, the fabric interconnect uses various mechanisms for loop prevention while preserving an active-active topology

Unicast Traffic Summary

Unicast traffic paths in Cisco UCS are shown in Figure 8. Characteristics of unicast traffic in Cisco UCS include:

- Each server link is pinned to exactly one uplink port (or PortChannel).
- Server-to-server Layer 2 traffic is locally switched.
- Server-to-network traffic goes out on its pinned uplink port.
- Network-to-server unicast traffic is forwarded to the server only if it arrives on a pinned uplink port. This feature is called the reverse path forwarding (RPF) check.
- Server traffic received on any uplink port, except its pinned uplink port, is dropped (called the deja-vu check)
- The server MAC address must be learned before traffic can be forwarded to it.

Figure 8. Unicast Traffic Paths

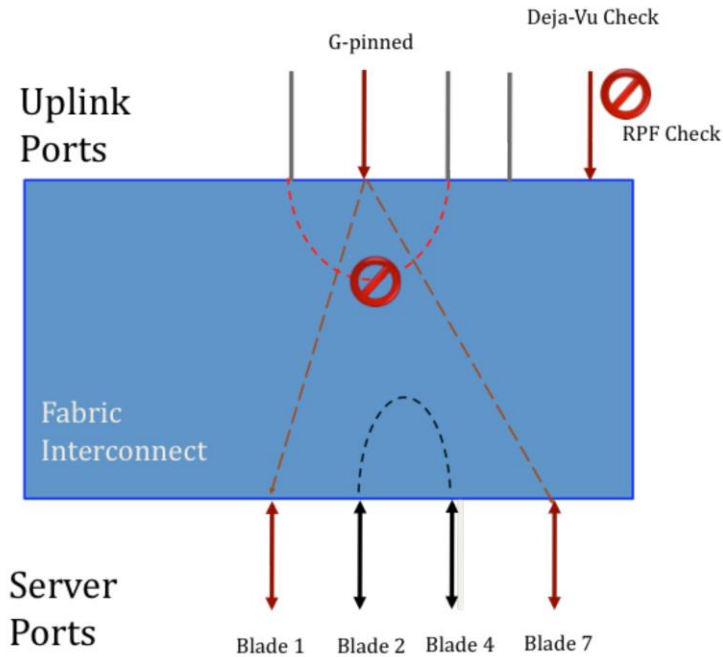


Multicast and Broadcast Forwarding Summary

Multicast traffic paths in Cisco UCS are shown in Figure 9. Characteristics of multicast traffic in Cisco UCS include:

- Broadcast traffic is pinned on exactly one uplink port in Cisco UCS Manager Release 1.4 and earlier and is dropped when received on the other uplink ports. In Cisco UCS Manager Release 2.0, the incoming broadcast traffic is pinned on a per-VLAN basis, depending on uplink port VLAN membership.
- IGMP multicast groups are pinned based on IGMP snooping. Each group is pinned to exactly one uplink port.
- Server-to-server multicast traffic is locally switched.
- RPF and deja-vu checks also apply to multicast traffic.

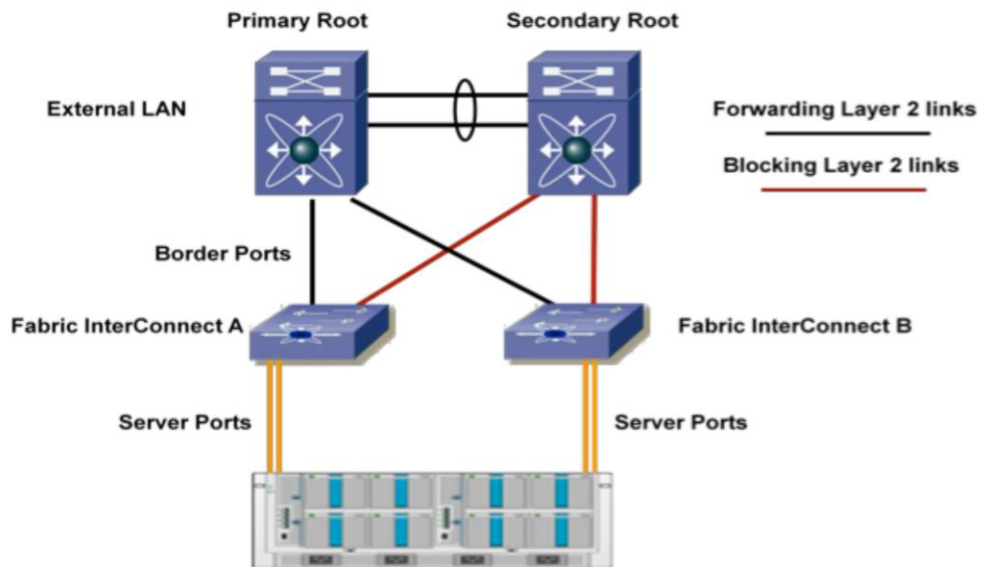
Figure 9. Multicast and Broadcast Traffic Summary



Ethernet Switching Mode

In Ethernet switching mode (Figure 10), the Cisco UCS 6100 Series Fabric Interconnects act like traditional Ethernet switches with support for Spanning Tree Protocol on the uplink ports.

Figure 10. Forwarding Paths in Switch Mode Without VSS or vPC Upstream



Ethernet switching mode features include:

- Spanning Tree Protocol is run on the uplink ports per VLAN as defined by Cisco® Per-VLAN Spanning Tree Plus (PVST+)
- Configuration of Spanning Tree Protocol parameters (bridge priority, hello timers, etc.) is not supported.
- VLAN Trunk Protocol (VTP) is not supported.
- MAC address learning and aging occur on both the server and uplink ports as in a typical Layer 2 switch.
- Upstream links are blocked according to Spanning Tree Protocol rules.

Design Consideration for Running Switch or End Host Mode

In most cases, end-host mode is preferable because it offers scalability and simplicity for server administrators when connecting to an upstream network. However, there are other factors to consider when selecting the appropriate switching mode, including:

- Scalability
- Efficient use of bandwidth
- Fabric failover
- Active-active link utilization
- Disjoint Layer 2 domain or a loop-free topology
- Optimal network behavior for the existing network topology
- Application-specific requirements

Scalability

Without Spanning Tree Protocol running, end-host mode provides the most scalability because the control plane is not occupied. Additionally, because there is no MAC address learning on the uplink ports, the MAC address table can scale to support as many virtual machines as the number of entries available in the MAC address forwarding table.

For scalability, the recommended Ethernet switching mode is end-host mode.

Efficient Bandwidth Use

Using static pin groups in end-host mode, an administrator can explicitly define the upstream port for a particular vNIC. This approach provides control over bandwidth and deterministic behavior for a vNIC.

For efficient bandwidth use, the recommended Ethernet switching mode is end-host mode.

Fabric Failover

Cisco UCS fabric failover, depending on the status of the uplink ports, is available only in end-host mode. In switch mode, OS-based teaming (active-passive) software is required to provide failover. The pinning feature provides the association between the uplink port and the vNIC that is used for tracking and for providing failover.

For fabric failover, the recommended Ethernet switching mode is end-host mode.

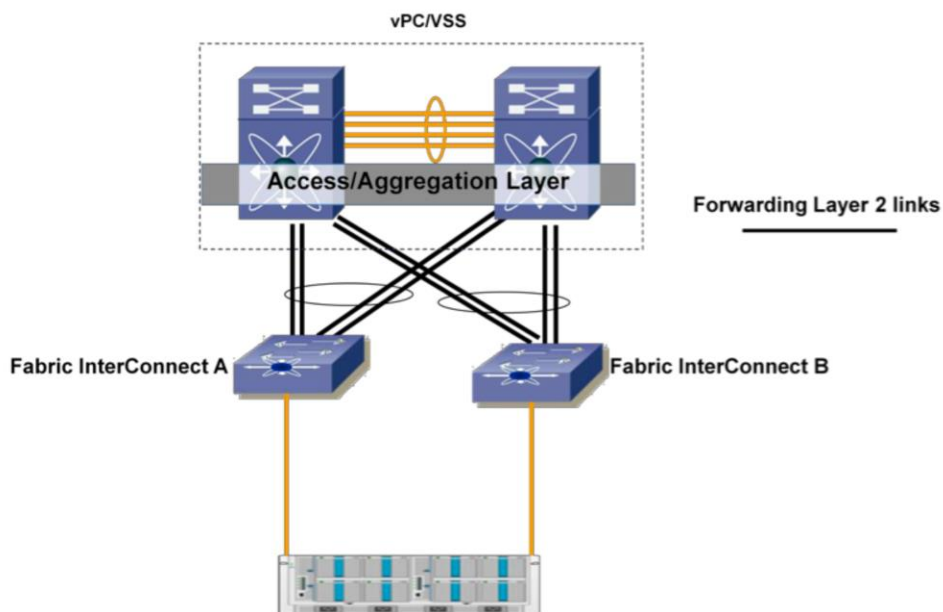
Active-Active Link Utilization

Without the use of technologies such as vPC or VSS upstream, switch mode blocks ports that depend on spanning root selection (Figure 10). By using pin groups in end-host mode, effective upstream bandwidth utilization is possible.

Note: Pinning is applicable only in end-host mode.

For active-active links without vPC or VSS upstream, the recommended switching mode is end-host mode. When vPC or VSS is available upstream (Figure 11), either end-host mode or switch mode can be used.

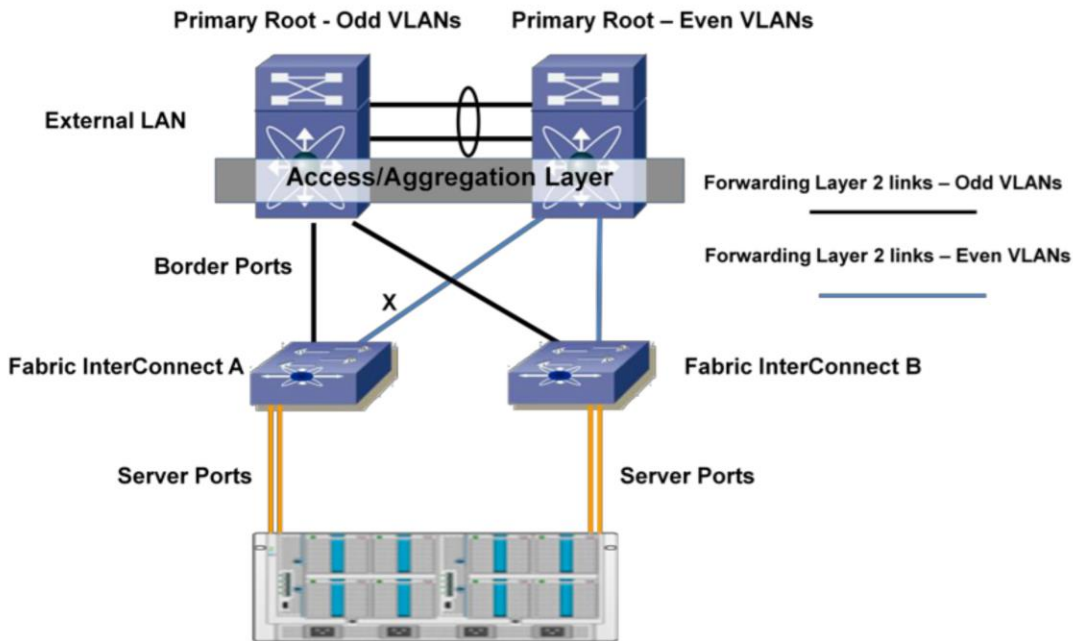
Figure 11. Forwarding Paths with VSS or vPC Upstream



Optimal Network Behavior for the Network Topology

In some commonly deployed LAN topologies, switch mode provides the best network behavior. A typical example is a switch directly connected to a pair of Hot Standby Router Protocol (HSRP) routers that are the Spanning Tree Protocol roots on different VLANs. Switch mode provides the optimal path because of the use of Spanning Tree Protocol. For example, a vNIC belonging to an odd-numbered VLAN can be dynamically pinned to link X on Fabric Interconnect A (Figure 12). As a result of this process, traffic traverses an extra hop (the Inter-Switch Link [ISL]) to the HSRP primary.

Figure 12. VLANs Load-Balanced Across a Pair of Switches



When a switch is directly connected to a pair of HRSP routers, the recommended Ethernet switching mode is switch mode, because it provides the optimal path. End-host mode can be used if static pinning is employed.

Application-Specific Requirements

Certain applications, such as Microsoft network load balancing in unicast mode, require unknown unicast flooding to operate. In end-host mode, unknown unicast broadcasts are not flooded, switch mode is required.

In applications that require load balancing in unicast mode, the recommended Ethernet switching mode is switch mode.

Conclusion

The external interface switching modes to a LAN that are supported in Cisco UCS provide the flexibility that server and network architects need to meet the most complex data center connectivity challenges. End-host mode provides a simple, scalable, and nondisruptive approach in most cases when integrating Cisco UCS into a network.

For More Information

- Cisco vPC white paper: http://www.cisco.com/en/US/prod/collateral/switches/ps9441/ps9402/white_paper_c11-516396.html
- Cisco vPC and VSS video presentation: http://www.youtube.com/watch?v=8edu9z_m8fl



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.

Cisco and the Cisco logo are trademarks or registered trademarks of Cisco and/or its affiliates in the U.S. and other countries. To view a list of Cisco trademarks, go to this URL: www.cisco.com/go/trademarks. Third party trademarks mentioned are the property of their respective owners. The use of the word partner does not imply a partnership relationship between Cisco and any other company. (1110R)

Printed in USA

C11-701962-00 03/12