certcollection.net

DCUFD

Designing Cisco Data Center Unfied Fabric

Volume 2

Version 5.0



Student Guide

Text Part Number: 97-3185-01

https://t.me/learningnets

certcollection.net



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)

DISCLAIMER WARRANTY: THIS CONTENT IS BEING PROVIDED “AS IS.” CISCO MAKES AND YOU RECEIVE NO WARRANTIES IN CONNECTION WITH THE CONTENT PROVIDED HEREUNDER, EXPRESS, IMPLIED, STATUTORY OR IN ANY OTHER PROVISION OF THIS CONTENT OR COMMUNICATION BETWEEN CISCO AND YOU. CISCO SPECIFICALLY DISCLAIMS ALL IMPLIED WARRANTIES, INCLUDING WARRANTIES OF MERCHANTABILITY, NON-INFRINGEMENT AND FITNESS FOR A PARTICULAR PURPOSE, OR ARISING FROM A COURSE OF DEALING, USAGE OR TRADE PRACTICE. This learning product may contain early release content, and while Cisco believes it to be accurate, it falls subject to the disclaimer above.

© 2012 Cisco and/or its affiliates. All rights reserved.

https://t.me/learningnets

certcollection.net

Table of Contents

Volume 2

Data Center Storage ***4-1***

Overview 4-1
Module Objectives 4-1

Introducing SAN **4-3**

Overview 4-3
Objectives 4-3
Data Storage and Fibre Channel 4-4
Fibre Channel Concepts 4-10
Fibre Channel Flow Control 4-24
Summary 4-29

Designing SAN **4-31**

Overview 4-31
Objectives 4-31
Storage Topologies 4-32
Storage Design Best Practices 4-36
Multitenant SANs 4-43
Summary 4-45
References 4-45

Designing Unified Fabric **4-47**

Overview 4-47
Objectives 4-47
Unified Fabric 4-48
FCoE Initialization Protocol 4-61
Unified Fabric Designs 4-72
Unified Fabric Designs with FEXs 4-76
Summary 4-79
References 4-79

Designing SAN Services **4-81**

Overview 4-81
Objectives 4-81
SAN-Based Services 4-82
SAN-Based Services Design Considerations 4-88
SAN-Based Data Replication 4-91
Long-Distance Fibre Channel Interconnects 4-97
Fibre Channel Long-Distance Acceleration Solutions 4-104
Summary 4-107
Module Summary 4-109
Module Self-Check 4-111
Module Self-Check Answer Key 4-113

Data Center Security ***5-1***

Overview 5-1
Module Objectives 5-1

Designing Data Center Application Security **5-3**

Overview 5-3
Objectives 5-3
Need for Data Center Security 5-4
Firewall Characteristics 5-13
Positioning Firewalls Within Data Center Networks 5-23
Secure Communication on Multiple Layers 5-32
Summary 5-39

https://t.me/learningnets

certcollection.net

Designing Link Security Technologies and Device Hardening 5-41
Overview 5-41
Objectives 5-41
Link Security 5-42
Device Hardening 5-46
Secure Management 5-60
Summary 5-67
Designing Storage Security 5-69
Overview 5-69
Objectives 5-69
Design Secure SAN 5-70
Data Security Solutions 5-81
Secure IP-Based Storage Design 5-86
Summary 5-87
Module Summary 5-89
Module Self-Check 5-91
Module Self-Check Answer Key 5-94
Data Center Application Services 6-1
Overview 6-1
Module Objectives 6-1
Designing Data Center Application Architecture 6-3
Overview 6-3
Objectives 6-3
Application Architecture and Design 6-4
Application Tiering 6-8
Wide-Area Application Optimization 6-12
Summary 6-19
Designing Application Services 6-21
Overview 6-21
Objectives 6-21
Server Load-Balancing Technologies 6-22
Application Delivery Services 6-35
Cisco ACE Virtualization 6-42
Secure Load-Balancing Design 6-48
Summary 6-51
Designing Global Load Balancing 6-53
Overview 6-53
Objectives 6-53
Need for GSLB 6-54
GSLB Solution Design 6-58
Site Selection Protocols 6-60
Site Selection Process 6-63
Summary 6-67
Module Summary 6-69
Module Self-Check 6-71
Module Self-Check Answer Key 6-73

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

Data Center Management 7-1

Overview 7-1
Module Objectives 7-1

Designing Data Center Management Solutions 7-3

Overview 7-3
Objectives 7-3
Need for Network Management 7-4
Cisco Data Center Management Tools 7-5
Network Management Scalability Limitations 7-21
Manage Multitenant Environments 7-22
Summary 7-23
Module Summary 7-25
Module Self-Check 7-27
Module Self-Check Answer Key 7-28

https://t.me/learningnets

certcollection.net

https://t.me/learningnets

certcollection.net

Module 4

Data Center Storage

Overview

In this module, you will learn about the foundation for unified fabric—the operation and design of SAN fabrics that are based on the Fibre Channel Protocol (FCP). This module also explains the transport for Fibre Channel over Ethernet (FCoE), along with design guidelines for converged networks.

Module Objectives

Upon completing this module, you will be able to present and design data center storage plans, solutions, and limitations of various storage technologies. This ability includes being able to meet these objectives:

- Explain the basics of Fibre Channel storage
- Design reliable, highly available, and flexible SANs
- Design unified fabric
- Design SAN-based Fibre Channel services

https://t.me/learningnets

certcollection.net

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

Lesson 1

Introducing SAN

Overview

This lesson explains basic data storage and Fibre Channel terms, explains basic concepts and technologies that are used on Fibre Channel networks, and introduces no-drop flow control mechanisms.

Objectives

Upon completing this lesson, you will be able to explain the basics of Fibre Channel storage. This ability includes being able to meet these objectives:

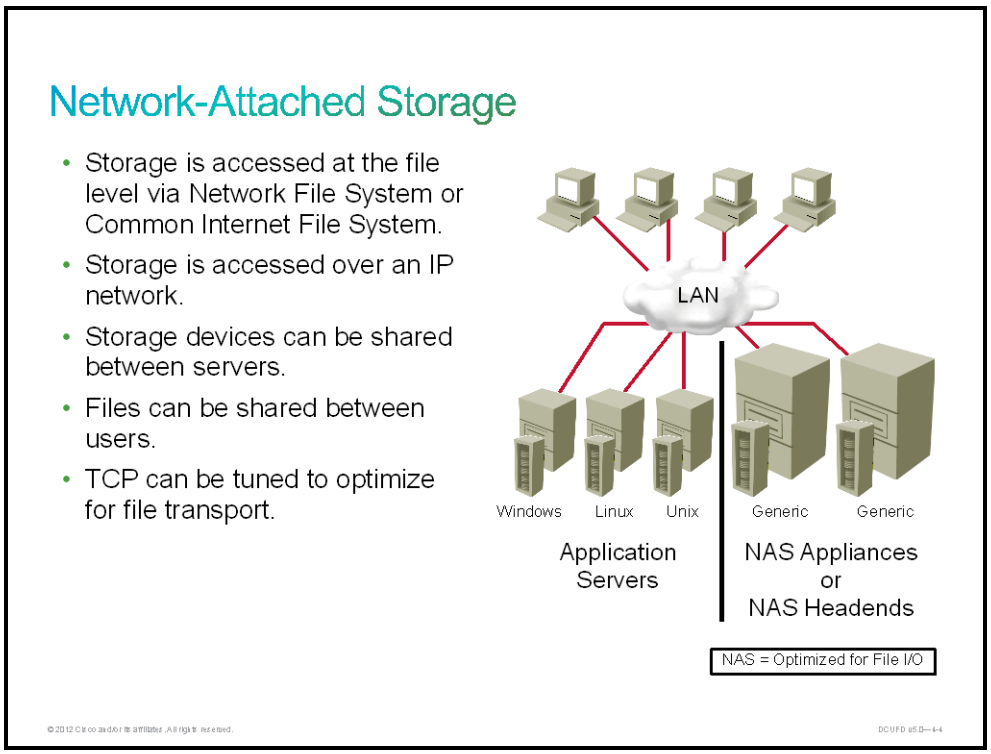
- Explain data storage and Fibre Channel basic terms
- Explain Fibre Channel basic concepts
- Explain Fibre Channel flow control mechanisms

https://t.me/learningnets

certcollection.net

Data Storage and Fibre Channel

This topic explains the basic terms for data storage and Fibre Channel.



Network-attached storage (NAS) is file-level computer data storage that is connected to a computer network that provides data access to heterogeneous clients. NAS devices allow users to attach scalable file-based storage directly to existing LANs based on IP and Ethernet, which provides easy installation and maintenance.

NAS systems are networked appliances that contain one or more hard drives, often arranged into logical, redundant storage containers or Redundant Array of Independent Disks (RAID). NAS removes the responsibility of file serving from other servers on the network. These servers typically provide access to files using network file-sharing protocols such as Network File System (NFS). An NAS unit is a computer that is connected to a network that only provides file-based data storage services to other devices on the network.

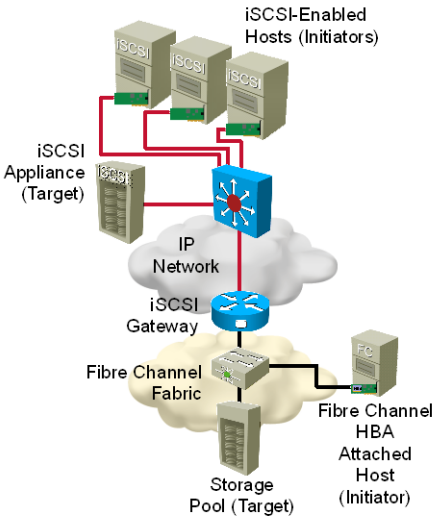
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

iSCSI Storage Network

- Storage accessed at block level via iSCSI
- Good performance via standard Ethernet NIC
- Enhanced performance with TCP/IP Offload Engine (TOE)
- Lower TCO relative to direct Fibre Channel HBA and fabric
- Standards-based



© 2012 Cisco and/or its affiliates. All rights reserved.

DCUFD v5.0—4-5

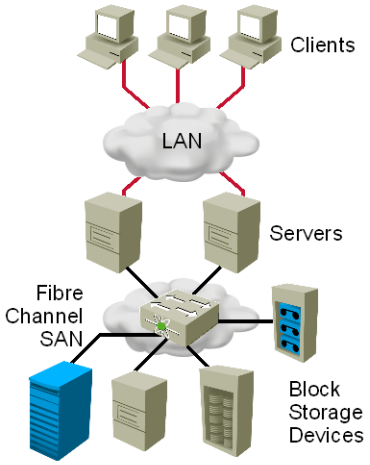
Internet Small Computer Systems Interface (iSCSI) represents an IP-based storage networking standard for linking data storage facilities. By carrying SCSI commands over IP networks, iSCSI makes data transfers possible over intranets and manages storage over long distances. iSCSI can transmit data over LANs and can enable location-independent data storage and retrieval. The protocol allows initiators to send SCSI commands to SCSI storage devices (targets) on remote servers. It allows organizations to consolidate storage into data center storage arrays while providing hosts (such as database and web servers) with the impression of locally attached disks. Unlike traditional Fibre Channel, which requires special-purpose cabling, iSCSI can be run over long distances using existing network infrastructure.

https://t.me/learningnets

certcollection.net

SAN Access

- Storage accessed at block level via SCSI
- High-performance interconnect providing high I/O throughput
- Lower TCO relative to direct attached storage; storage can be shared
- Limited vendor interoperability
- Complex management



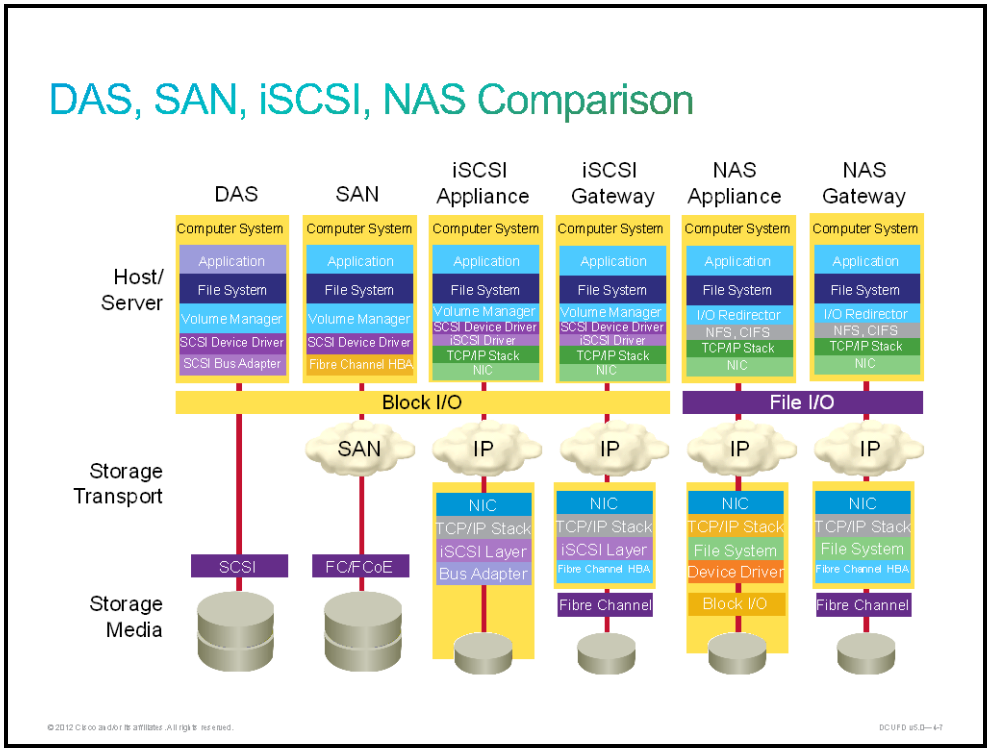
Separation of Storage from the Server

© 2012 Cisco and/or its affiliates. All rights reserved.

A SAN is a dedicated storage network that provides access to consolidated, block-level storage. SANs are primarily used to make storage devices accessible to servers so that the devices appear as being locally attached to the operating system. A SAN typically has its own network of storage devices that are generally not accessible through the regular network by regular devices. A SAN alone does not provide the “file” abstraction. It provides only block-level operations.

https://t.me/learningnets

certcollection.net



The major difference between direct-attached storage (DAS) and NAS is that DAS is simply an extension of an existing server and is not necessarily networked. NAS is designed as an easy and self-contained solution for sharing files over the network.

When both are available over the network, NAS might provide better performance than DAS because the NAS device can be tuned precisely for file serving, which is less likely to happen on a server that is responsible for other processing. Both NAS and DAS can have different amounts of cache memory, which greatly affects performance. When you are comparing the use of NAS with the use of local (non-networked) DAS, the performance of NAS depends mainly on the speed of the network and congestion on the network.

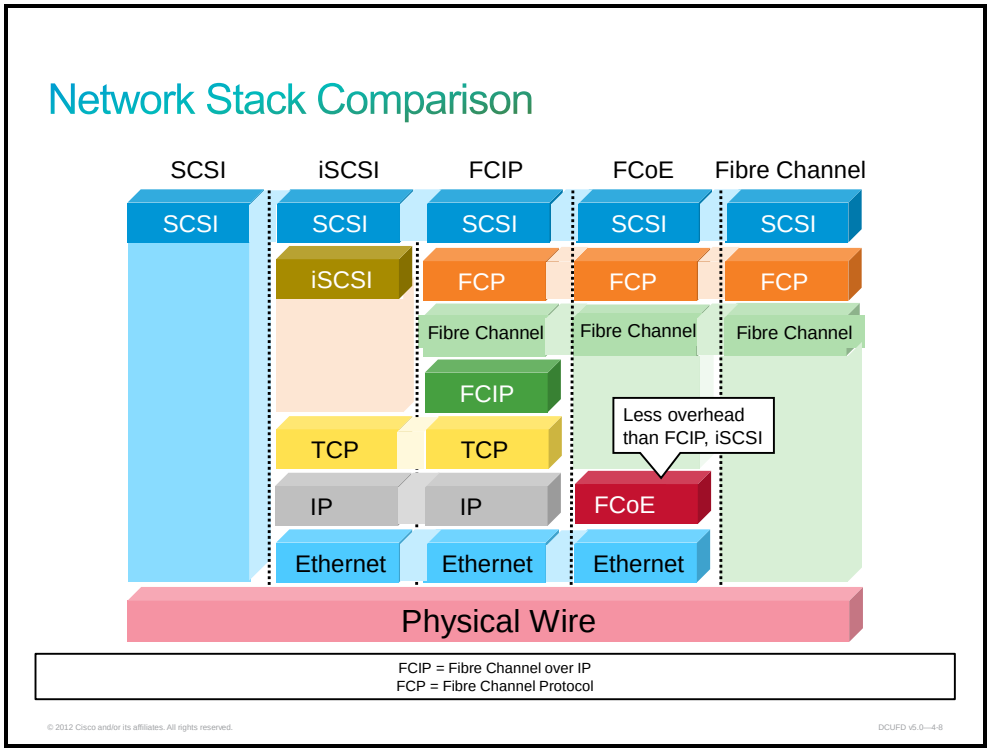
Despite their differences, SAN and NAS are not mutually exclusive. They can be combined as a SAN-NAS hybrid, which offers both file-level protocols (it serves up a file) and block-level protocols (it provides a disk drive) from the same system.

Many data centers use Ethernet for TCP/IP networks and Fibre Channel for SANs. With Fibre Channel over Ethernet (FCoE), Fibre Channel becomes another network protocol that runs on Ethernet alongside traditional IP traffic. FCoE operates directly above Ethernet in the network protocol stack, in contrast to iSCSI, which runs in addition to TCP and IP.

Because Classical Ethernet has no flow control—unlike Fibre Channel—FCoE requires enhancements to the Ethernet standard to support a flow control mechanism that prevents frame loss.

https://t.me/learningnets

certcollection.net



This figure shows the different elements of competing network stacks. While all stacks depend on SCSI mechanisms, different transport modules are used.

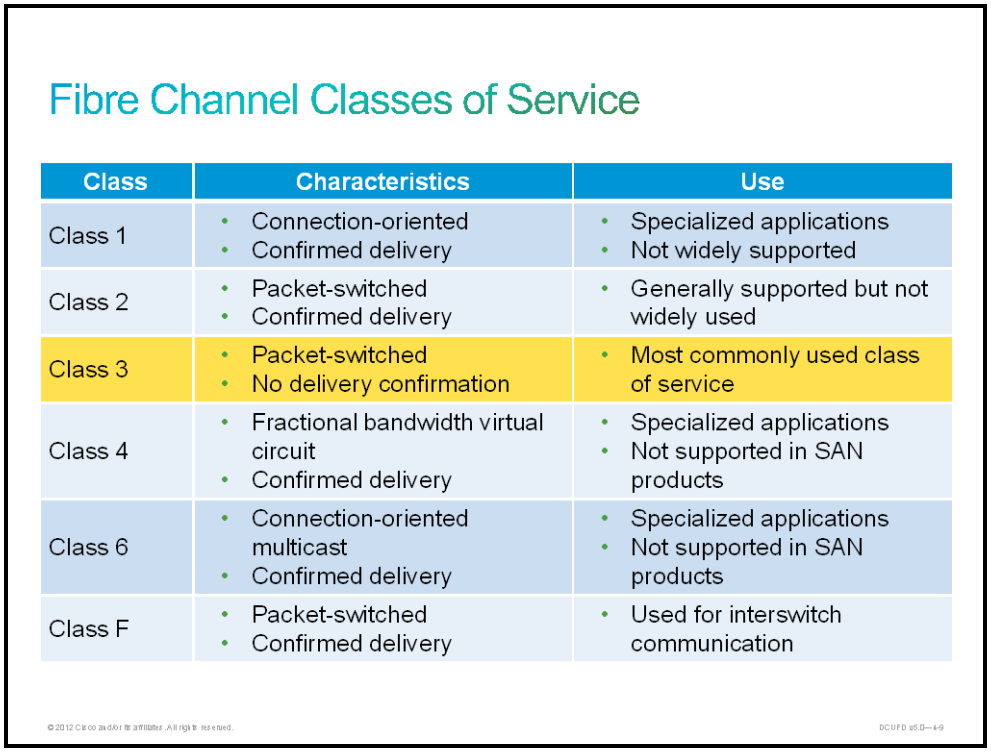
The SCSI is a mechanism that provides a low-overhead, High-Performance Parallel Interface (HIPPI) that is efficient in managing storage traffic within a chassis. Fibre Channel overcomes the distance and switching limitations that are inherent in SCSI. Fibre Channel carries SCSI as its higher-level protocol. SCSI does not respond well to lost frames, which can result in significant delays when recovering from a loss. Because Fibre Channel carries SCSI, it inherits the requirement for an underlying lossless network.

The ability of FCoE to work seamlessly with existing infrastructure makes it an evolutionary technology, one that data centers can deploy at the pace and to the extent that best serves their needs. FCoE allows IP and Fibre Channel network traffic to be carried over existing FCoE-aware drivers, network interface cards (NICs), and switches, which allows the use of a single cabling infrastructure within server racks. This technology simplifies network topology while reducing cabling cost and complexity, eliminating half of the I/O adapter cards in a rack, reducing power, and cooling overhead—all while improving bandwidth by leveraging 10-Gb/s Ethernet. As FCoE-enabled storage systems become available, data centers can implement a fully converged fabric, reaching from servers to storage using FCoE-aware switches.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net



This figure shows the uses of the Fibre Channel classes of service:

- Few commercially available Fibre Channel SAN products currently support Class 1.
- Many Fibre Channel products support Class 2, but it is not widely used.
- Class 3 is, by far, the most commonly used class of service on fabrics and it is often the only class that is supported on arbitrated loops. All Fibre Channel SAN products support Class 3.
- No commercially available Fibre Channel SAN products currently support Class 4 or Class 6.
- Class F is always used for interswitch communication.

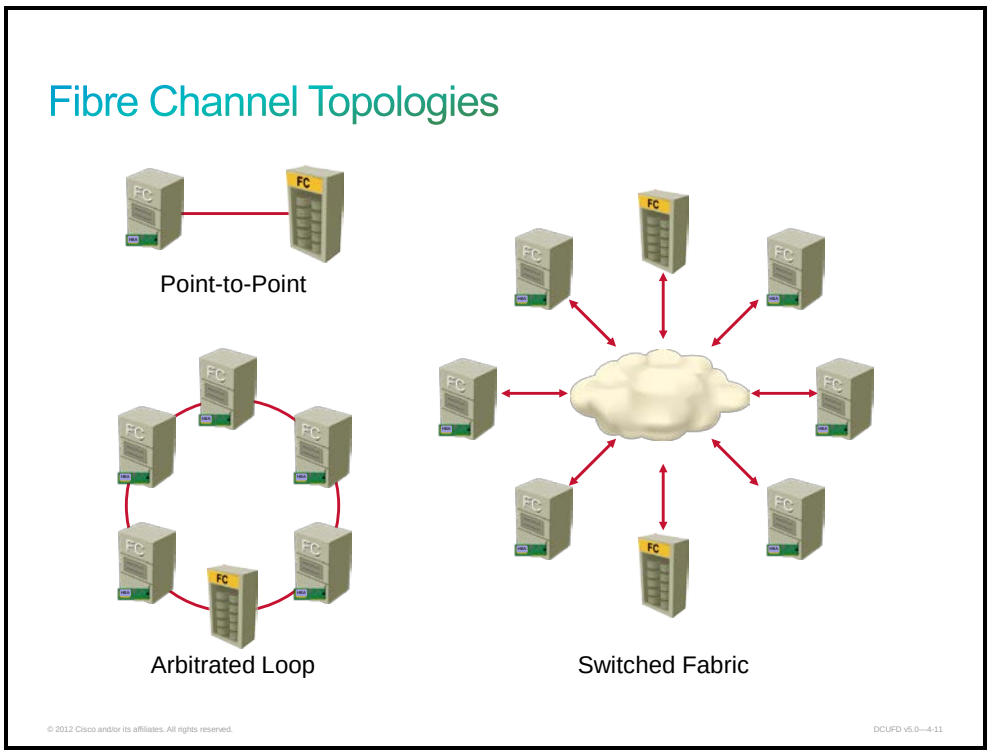
Note that Class 5 is not yet defined. Class 5 was intended to enable isochronous transactions by multiple ports, but it has not been completed. An isochronous connection is one in which bandwidth and data delivery rate are guaranteed. Class 5 would be appropriate for video delivery services.

https://t.me/learningnets

certcollection.net

Fibre Channel Concepts

This topic describes the basic concepts of Fibre Channel.



Several different Fibre Channel topologies exist:

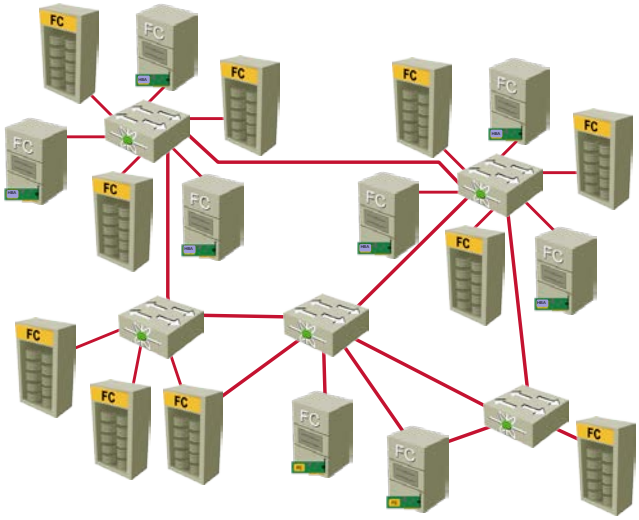
- **Point-to-point:** In this topology, two ports or devices are directly connected. This is the simplest topology, where each message has only one receiver.
- **Arbitrated loop:** In this topology, devices are connected in a loop that is similar to round-robin topology. Only two ports can communicate at the same time. This topology is rarely used for server-to-storage communication.
- **Switched fabric:** In this topology, devices are connected, one to another, via Fibre Channel switches. This topology has the best scalability properties. It is, however, the most expensive option. Currently, most Fibre Channel topologies are switched, because they are flexible and scalable.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

Switched Fabric Topology



Fibre Channel fabric is a switched network.

© 2012 Cisco and/or its affiliates. All rights reserved.

In a switched fabric topology, devices are connected, one to another, via Fibre Channel switches. Traffic paths between end nodes are determined by a routing protocol.

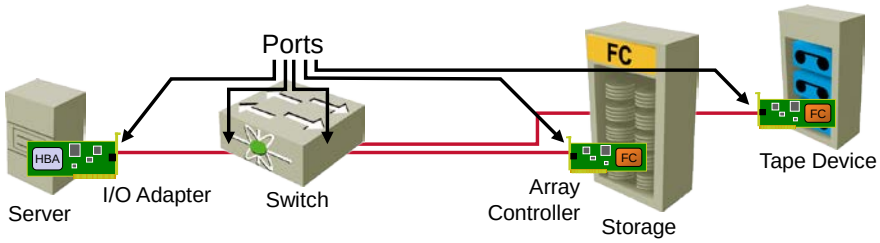
https://t.me/learningnets

certcollection.net

Fibre Channel Ports

Ports are intelligent interface points on the Fibre Channel network:

- Embedded in an HBA
- Embedded in an array or tape controller
- Embedded in a fabric switch
- Ports understand Fibre Channel



© 2012 Cisco and/or its affiliates. All rights reserved.

Fibre Channel ports are intelligent interface points (or structures) on the Fibre Channel network. They are embedded in host bus adapters (HBAs), array and tape controllers, and fabric switches.

Fibre Channel ports have sufficient logic to communicate with other devices. Fibre Channel ports on switches conform to rules to accept storage traffic from end hosts and to send or accept data from other Fibre Channel switches in the fabric.

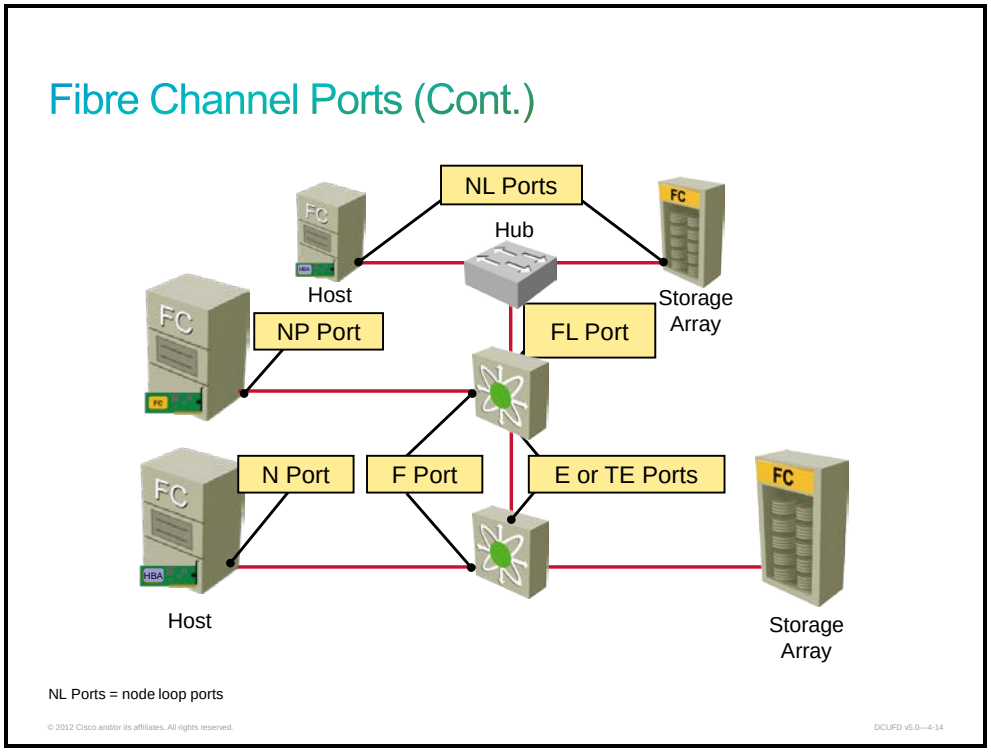
Logically, ports are of different types. The three most common ports are node ports, fabric ports, and extension ports. Every host or end device has a node port, which must connect to a fabric port on the other side. Fabric ports are hosted on switches.

Switches interconnect using extension ports. In addition to carrying storage traffic, extension ports are also used to maintain the Fibre Channel fabric.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net



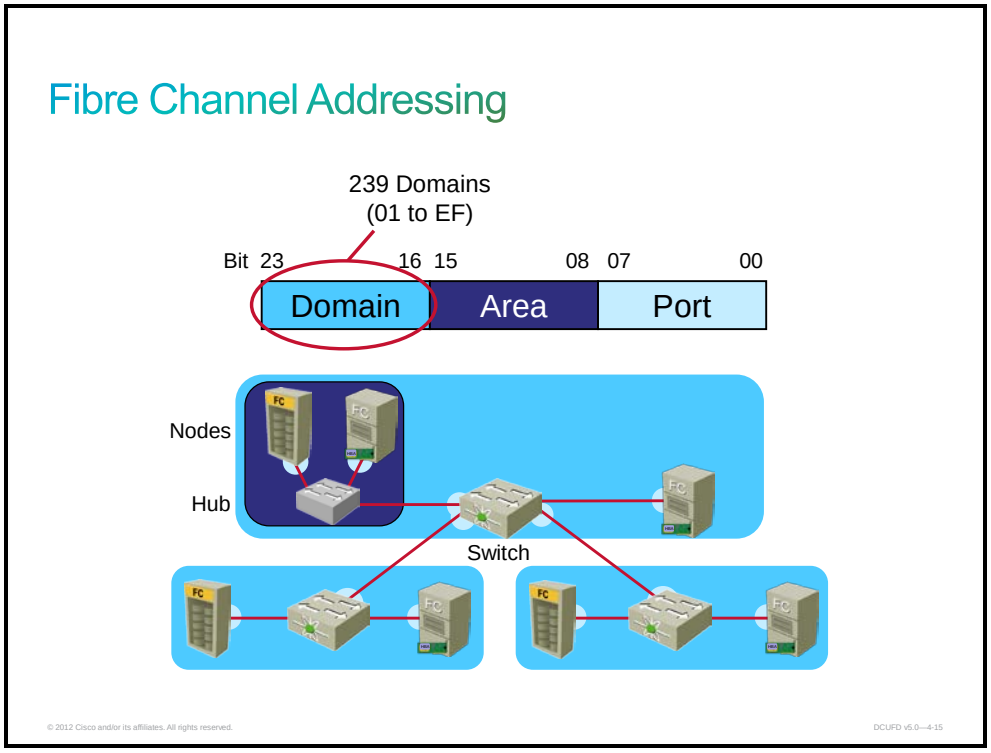
This table describes the different port types that are used when referring to Fibre Channel ports.

Different Port Types Used When Referring to Fibre Channel Ports

Port	Port Name	Description
N Port	Node port	This is a port on a node that connects to a fabric. N Ports can directly connect two devices in a point-to-point topology. Array controllers and I/O adapters have one or more N Ports.
NP Port	Proxy N port	This port behaves like an N Port except that, in addition to providing N Port behavior, it also functions as a proxy for multiple, physical N Ports.
F Port	Fabric port	This is a port on a switch that connects to an N Port.
E Port	Expansion port	This is a port on a switch that connects to an E Port on another switch.
TE Port	Trunking E port	This is an E Port that functions as a trunking expansion port. It may be connected to another TE Port to create an Extended Inter-Switch Link (EISL) between two switches. The TE Port provides not only standard E Port functions, but allows for routing of multiple virtual SANs (VSANs).
VE Port	Virtual E port	This port emulates an E Port over a non-Fibre Channel link.
VF Port	Virtual F port	This port emulates an F Port over a non-Fibre Channel link.
VN Port	Virtual N port	This port emulates an N Port over a non-Fibre Channel link.

https://t.me/learningnets

certcollection.net



The Fibre Channel point-to-point topology uses a one-bit addressing scheme. One port assigns itself an address of 000000 and then it assigns the other port an address of 000001.

The Fibre Channel arbitrated loop topology uses an eight-bit addressing scheme:

- The arbitrated loop physical address (ALPA) is an 8-bit address that provides 256 potential addresses. However, only a subset of 127 addresses is available due to 8b/10b encoding requirements.
- One address is reserved for a fabric loop port (FL Port), so there are 126 addresses that remain available for nodes.
- Addresses are cooperatively chosen during loop initialization.

On switched Fibre Channel fabric, the 24-bit Fibre Channel address consists of three 8-bit elements:

- The domain ID defines a switch. Each switch receives a unique domain ID.
- The area ID identifies groups of ports within a domain. Areas can be used to group ports within a switch and are also used to uniquely identify fabric-attached arbitrated loops. Each fabric-attached loop receives a unique area ID.
- The port ID identifies each individual port within an area.

Although the domain ID is an 8-bit field, only 239 domains are available to the fabric:

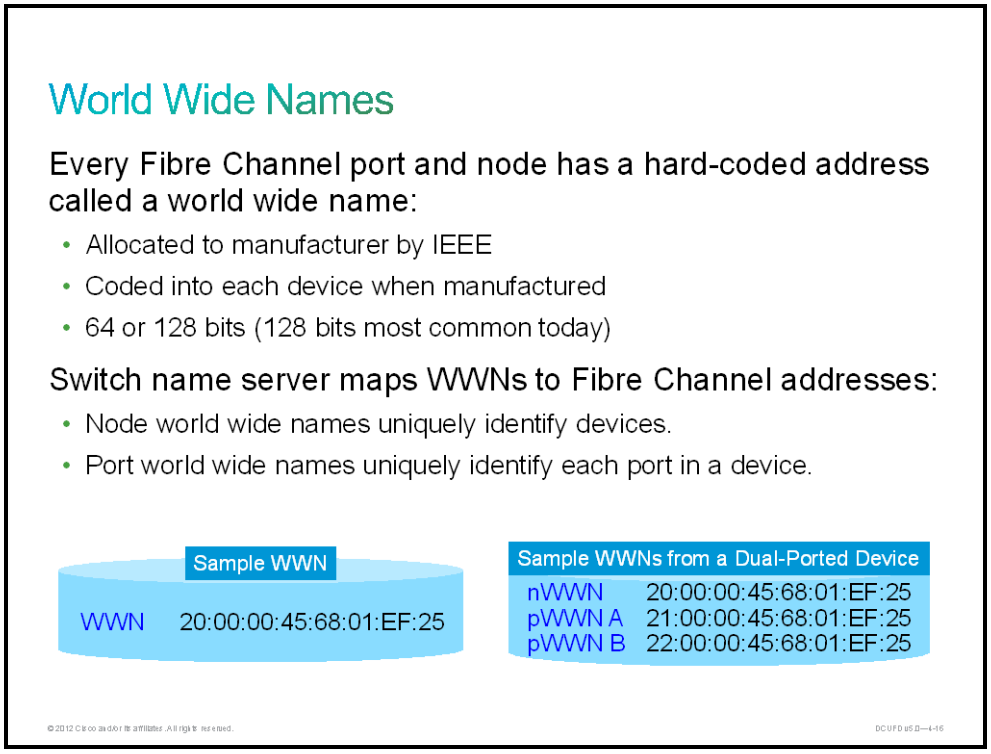
- Domains 01 to EF are available.
- Domains 00 and F0 to FF are reserved for use by switch services.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

Each switch must have a unique domain ID, so there can be no more than 239 switches in a fabric. The largest director-class switch that is available today has 256 ports, so the practical limit on the number of nodes that can be supported in a fabric is 61,184 ports (239 domains x 256 ports). With 16-port switches, the total port count is reduced to 3824 (239 domains x 16 ports), minus the number of ports that are used for Inter-Switch Links (ISLs). Note that these calculations do not take into account the ports that are consumed by ISLs (which reduces the number of ports) or the fact that an arbitrated loop with multiple loop ports (L Ports) can be attached to a single FL Port (which increases the potential number of ports).



World wide names (WWNs) are unique identifiers that are hardcoded into Fibre Channel devices. Every Fibre Channel port has at least one WWN. Vendors buy blocks of WWNs from the IEEE and allocate them to devices in the factory.

WWNs are important for enabling fabric services because they have these characteristics:

- Guaranteed to be globally unique
- Permanently associated with devices

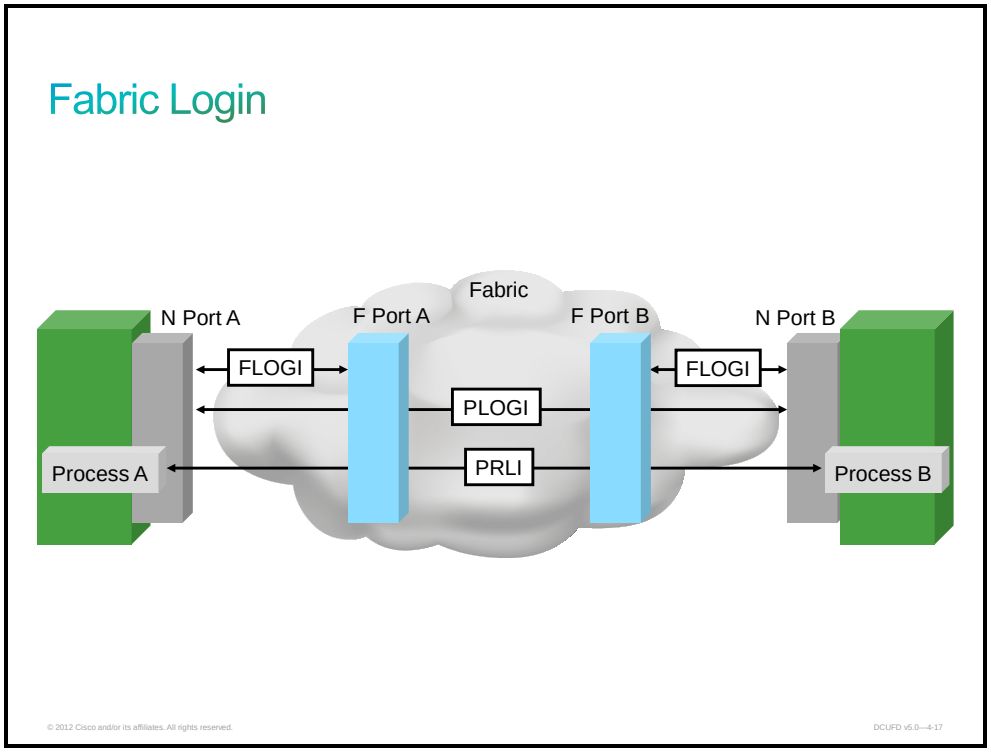
There are two types of WWNs:

- Node world wide names (nWWNs) uniquely identify devices. Every HBA, array controller, switch, gateway, and Fibre Channel disk drive has a single unique nWWN.
- Port world wide names (pWWNs) uniquely identify each port in a device. A dual-ported HBA has three WWNs: one nWWN, and a pWWN for each port.

nWWNs and pWWNs are both required because devices can have multiple ports. On single-ported devices, the nWWN and pWWN are usually the same. On multiported devices, however, the pWWN is used to uniquely identify each port. Ports must be uniquely identifiable because each port participates in a unique data path. nWWNs are required because the node itself must sometimes be uniquely identified. For example, path failover and multiplexing software can detect redundant paths to a device by observing that the same nWWN is associated with multiple pWWNs.

https://t.me/learningnets

certcollection.net



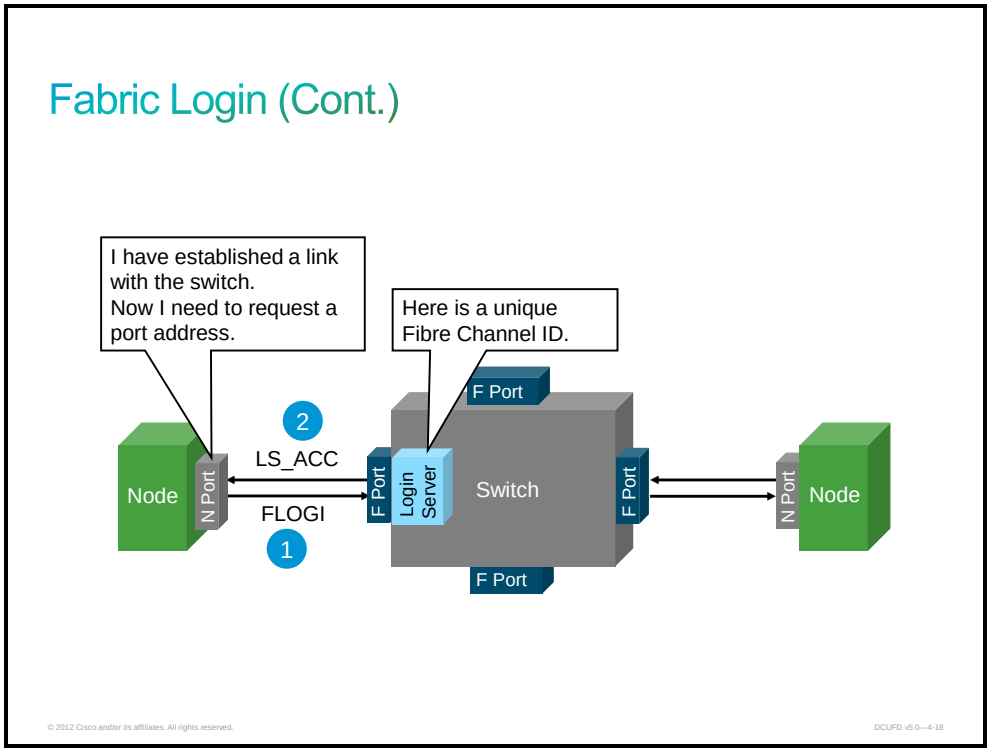
Before an N Port can begin exchanging data with other N Ports, three processes must occur:

- The N Port must log in to its attached F Port. This process is known as fabric login (FLOGI).
- The N Port must log in to its target N Port. This process is known as port login (PLOGI).
- The N Port must exchange information about upper-layer protocol (ULP) support with its target N Port to ensure that the initiator and target process can communicate. This process is known as process login (PRLI).

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net



The N Port follows this process:

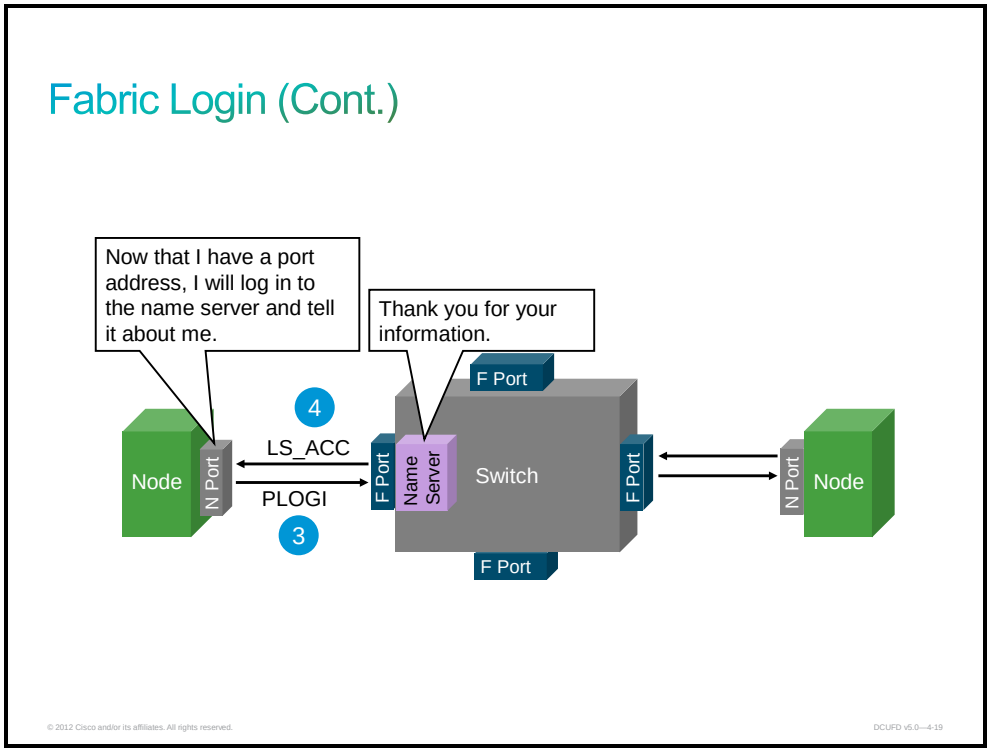
1. After the N Port has established a link to its F Port, the N Port obtains a port address by sending a FLOGI link services command to the switch login server (at well-known address 0xFFFFFE).

2. The login server sends an accept (ACC) reply that contains the N Port address in the destination ID field.

When an N Port is performing FLOGI and receives an ACC frame that indicates that the ACC came from another N Port, the N Port that is logging in assumes that it is in a point-to-point configuration. In this case, the N Port immediately initiates PLOGI with the other N Port after completing FLOGI.

https://t.me/learningnets

certcollection.net



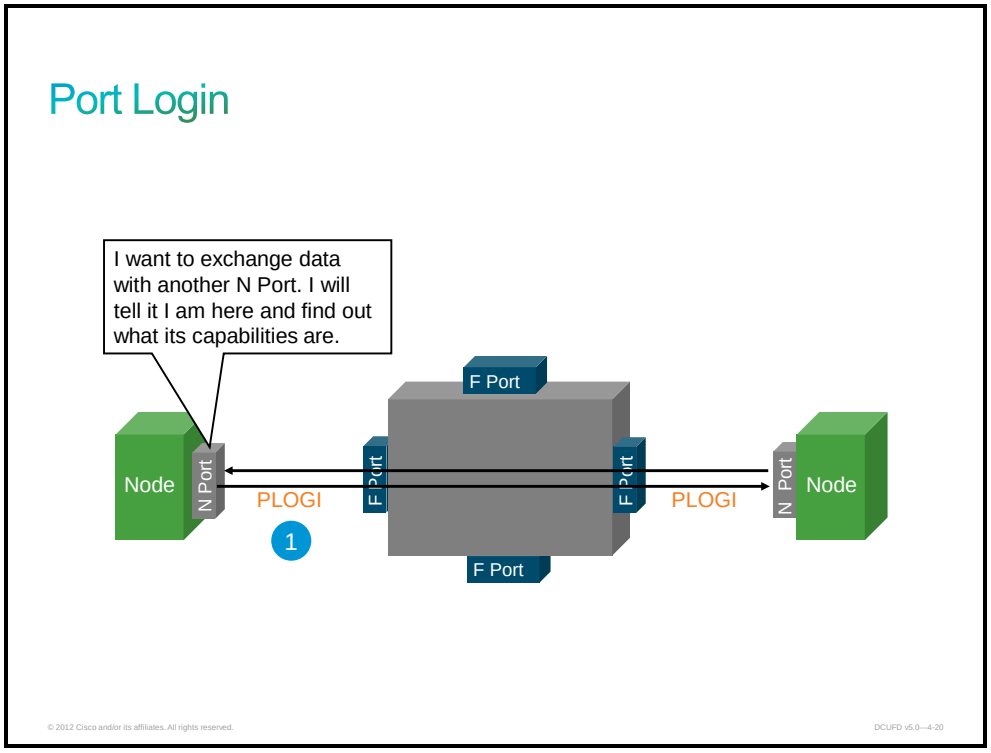
3. After receiving a port address, the N Port logs into the fabric name server at address 0xFFFFFC and transmits its service parameters, such as the number of buffer credits it supports, its maximum payload size, and supported classes of service.

4. The name server responds with an LS_ACC frame.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net



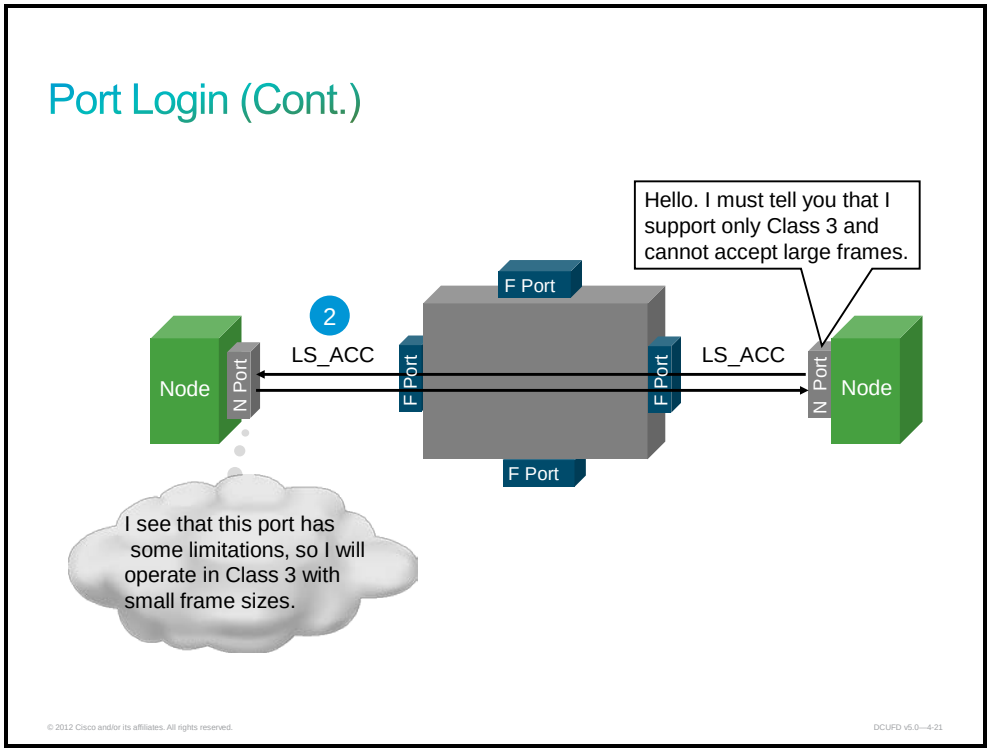
After completing the FLOGI process, the N Port can log into another N Port using the PLOGI protocol. PLOGI must be completed before the nodes can perform any ULP operations.

The PLOGI protocol follows this process:

5. The initiator N Port sends a PLOGI frame that contains the N Port operating parameters encapsulated in the payload.

https://t.me/learningnets

certcollection.net



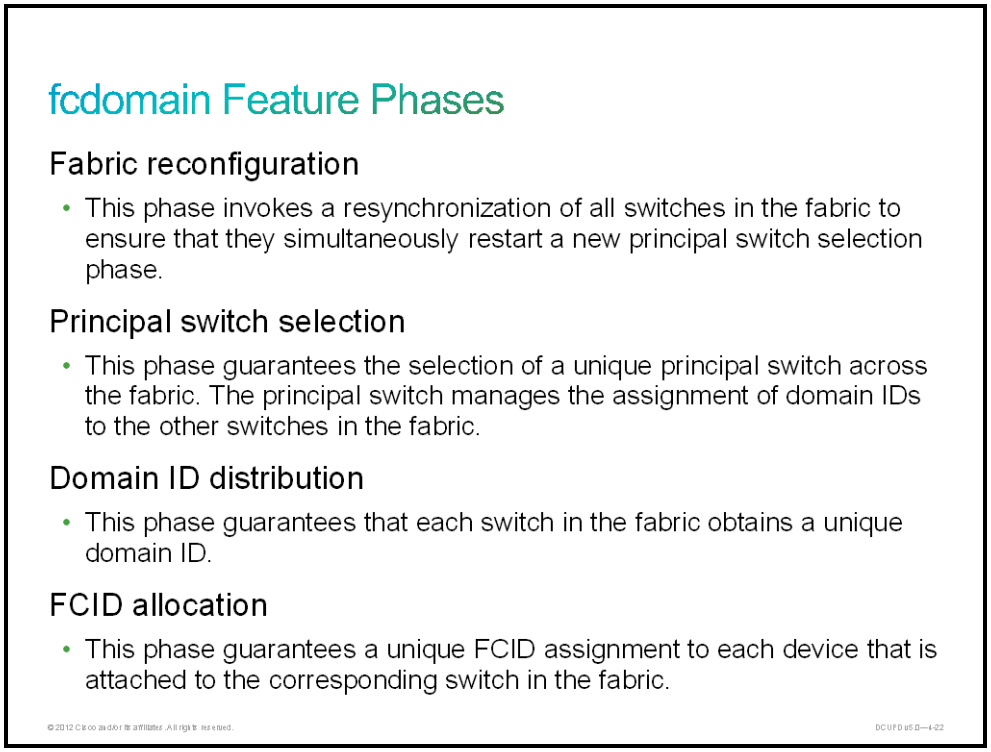
6. The target N Port responds to the initiator N Port by sending an ACC frame that specifies the target N Port operating parameters. The operating system driver that manages the initiator N Port stores this information in a parameter block.

An N Port can be logged into multiple N Ports simultaneously. N Ports typically perform port logout only when one of the nodes goes offline.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net



The Fibre Channel domain (fcdomain) feature performs principal switch selection, domain ID distribution, Fibre Channel ID (FCID) allocation, and fabric reconfiguration functions as described in the FC-SW-2 standards. The domains are configured on a per-VSAN basis, and if you do not configure a domain ID, the local switches use a random ID.

To successfully configure domain parameters and prevent fabric segmentation, it is necessary to understand the anticipated behavior of the fcdomain feature phases:

- **Fabric reconfiguration:** During fabric reconfiguration, the entire process of SAN initialization is restarted and traffic across the SAN is stopped.
- **Principal switch selection:** This phase guarantees the selection of a unique principal switch across the fabric.

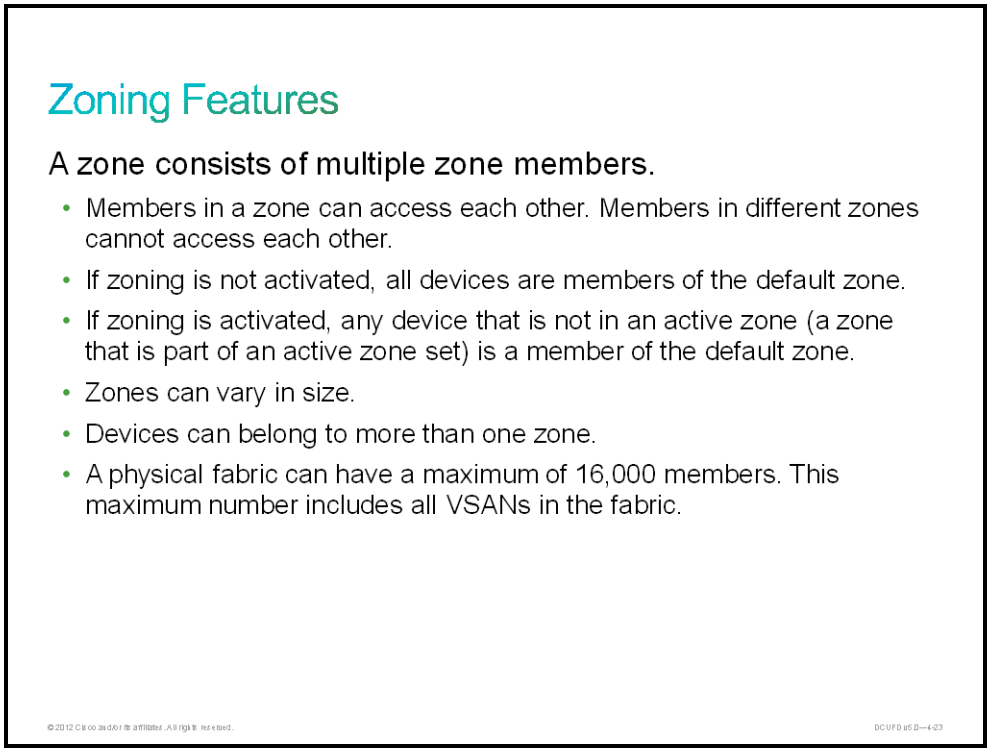
Note	The principal switch should be a highly available device, such as Cisco MDS 9500, and it should be located in the SAN core.

When adding a new switch (Cisco MDS or Cisco Nexus 7000, 5000, or 5500) to an existing Cisco MDS 9500-based fabric, ensure that the principal switch priority is lower than the priority of the current principal switch. Otherwise, fabric reconfiguration will occur and disrupt traffic across the SAN.

- **Domain ID distribution:** This phase guarantees that each switch in the fabric obtains a unique domain ID.
- **FCID allocation:** This phase guarantees a unique FCID assignment to each device that is attached to the corresponding switch in the fabric.
- **Fabric reconfiguration:** This phase guarantees a resynchronization of all switches in the fabric to ensure that they simultaneously restart a new principal switch selection phase.

https://t.me/learningnets

certcollection.net



Zones can consist of multiple zone members:

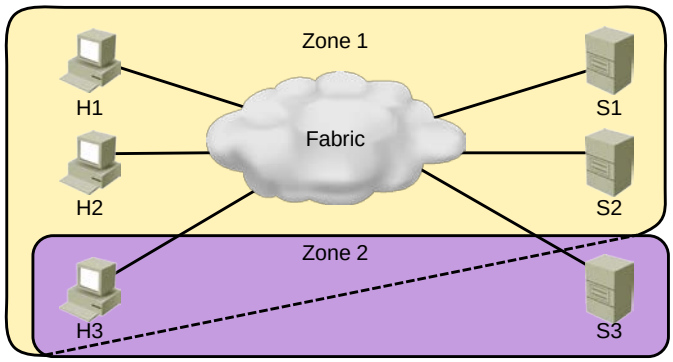
- Members in a zone can access each other. Members in different zones cannot access each other.
- If zoning is not activated, all devices are members of the default zone.
- If zoning is activated, any device that is not in an active zone (that is, a zone that is part of an active zone set) is a member of the default zone.
- Zones can vary in size.
- Devices can belong to more than one zone.
- A physical fabric can have a maximum of 16,000 members. This maximum number includes all VSANs in the fabric.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

Fabric with Two Zones



© 2012 Cisco and/or its affiliates. All rights reserved.

This figure shows a zone set with two zones—Zone 1 and Zone 2—in a fabric. Zone 1 provides access from all three hosts (H1, H2, and H3) to the data that resides on storage systems S1 and S2. Zone 2 restricts the data on S3 to access only by H3. H3 resides in both zones.

https://t.me/learningnets

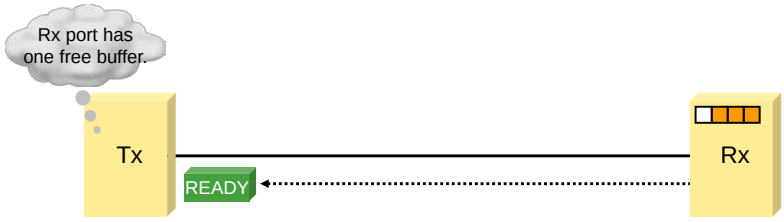
certcollection.net

Fibre Channel Flow Control

This topic describes Fibre Channel flow control mechanisms.

Credit-Based Flow Control

- Fibre Channel uses a credit-based strategy:
  - The transmitter does not send a frame until the receiver tells the transmitter that the receiver can accept another frame.
  - The receiver is always in control.
- There are benefits to credit-based flow control:
  - Prevents loss of frames due to buffer overruns
  - Maximizes performance under high loads



Flow Control in Fibre Channel

© 2012 Cisco and/or its affiliates. All rights reserved.

To improve performance under high-traffic loads, Fibre Channel uses a credit-based flow control strategy in which the receiver must issue a credit for each frame that is sent by the transmitter before that frame can be sent.

A credit-based strategy ensures that the receive (Rx) port is always in control. The Rx port must issue a credit for each frame that is sent by the transmitter. This strategy prevents frames from being lost when the Rx port runs out of free buffers. Preventing lost frames maximizes performance under high-traffic load conditions because the transmit (Tx) port does not have to resend frames.

The figure shows a credit-based flow control process:

- The Tx port counts the number of free buffers at the Rx port.
- Before Tx can send a frame, Rx must notify Tx that Rx has a free buffer and is ready to accept a frame. When Tx receives the notification, it increments its count of the number of free buffers at Rx.
- Tx sends frames only when it knows that Rx can accept them.

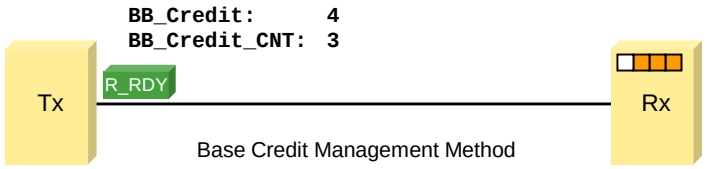
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

Base Credit Management Method

1. At login, Rx tells Tx how many buffers Rx has (BB_Credit).
2. Tx sets BB_Credit_CNT = 0 at login.
3. Tx increments BB_Credit_CNT when it sends a frame.
4. Rx sends an R_RDY message when it processes the frame.
5. Tx decrements BB_Credit_CNT when it receives the R_RDY message. Tx sends only when the BB_Credit_CNT value is less than the BB_Credit value.



© 2012 Cisco and/or its affiliates. All rights reserved.

The base credit management method works as follows:

- When the Tx port sends a port login request, the Rx responds with an ACC frame that includes information about the size and number of frame buffers it has (buffer-to-buffer credit [BB_Credit]). The Tx port stores the BB_Credit value in a table.
- The Tx port also stores another value called “BB_Credit_CNT,” which represents the number of used buffer credits. BB_Credit_CNT is set to zero after the ports complete the login process.
- Each time the Tx port sends a frame, it increments BB_Credit_CNT.
- Upon receiving the frame, the Rx processes the frame and moves it to ULP buffer space. The Rx port then sends a receiver ready (R_RDY) acknowledgment signal back to the Tx port, informing it that a buffer is available.
- When the Tx port receives the R_RDY signal, it then decrements its BB_Credit_CNT.

To prevent overrunning the Rx port buffers, the Tx port can never allow BB_Credit_CNT (the count of frames that have not yet been acknowledged) to exceed BB_Credit (the total number of buffers in the Rx port). In other words, if it cannot confirm that the Rx port has a free buffer, it does not send any more frames.

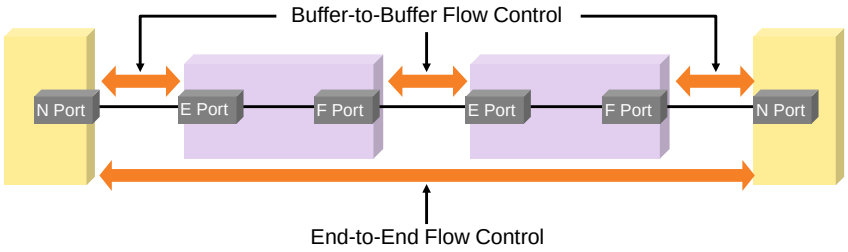
https://t.me/learningnets

certcollection.net

Types of Flow Control

Fibre Channel defines two types of flow control:

- Buffer-to-buffer (port to port)
- End-to-end (source to destination)



© 2012 Cisco and/or its affiliates. All rights reserved.

Fibre Channel defines two types of flow control:

- Buffer-to-buffer flow control takes place between two ports that are connected by a Fibre Channel link, such as an N Port and an F Port, or two E Ports, or two L Ports.
- End-to-end flow control takes place between the source node and the destination node.

Note that buffer-to-buffer flow control is performed between E Ports in the fabric, but it is not performed between the incoming and outgoing ports in a given switch. In other words, Fibre Channel buffer-to-buffer flow control is not used between two F Ports or between an F Port and an E Port within a switch. Fibre Channel does not define how switches route frames across the switch.

Buffer-to-buffer flow control is used in the following situations:

- Class 1 connection request frames use buffer-to-buffer flow control, but Class 1 data traffic uses only end-to-end flow control.
- Class 2 and Class 3 frames always use buffer-to-buffer flow control.
- Class F service uses buffer-to-buffer flow control.
- In an arbitrated loop, every communication session is a virtual dedicated point-to-point circuit between a source port and destination port. Therefore, there is little difference between buffer-to-buffer flow control and end-to-end flow control. Buffer-to-buffer flow control alone is generally sufficient for arbitrated loop topologies.

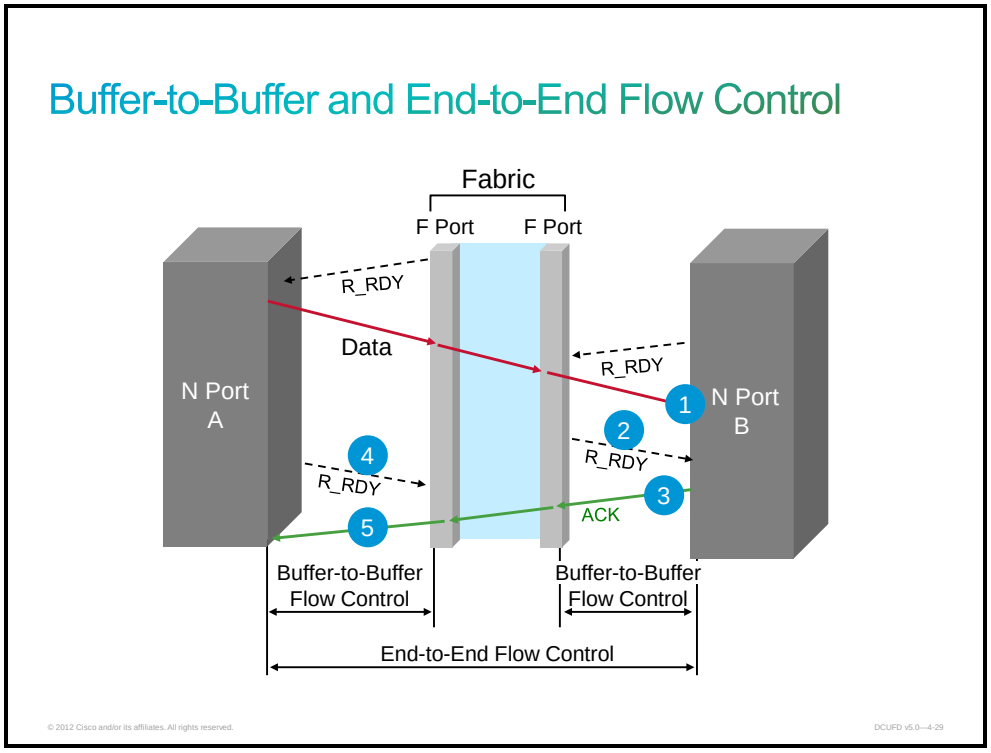
End-to-end flow control is used in the following situations:

- Classes 1, 2, 4, and 6 use end-to-end flow control.
- Class 2 service uses both buffer-to-buffer and end-to-end flow control.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net



This figure shows buffer-to-buffer flow control in Class 3:

1. Before N Port A can transmit a frame, it must receive the R_RDY signal from its attached F Port. The R_RDY signal tells N Port A that its F Port has a free buffer.
2. When it receives the R_RDY signal, N Port A transmits a frame.
3. The frame is passed through the fabric. Buffer-to-buffer flow control is performed between every pair of E Ports, although this is not shown here.
4. At the other side of the fabric, the destination F Port must wait for an R_RDY signal from N Port B.
5. When N Port B sends an R_RDY, the F Port transmits the data frame.

End-to-end flow control is designed to overcome the limitations of buffer-to-buffer flow control. The figure also shows end-to-end flow control in Class 2:

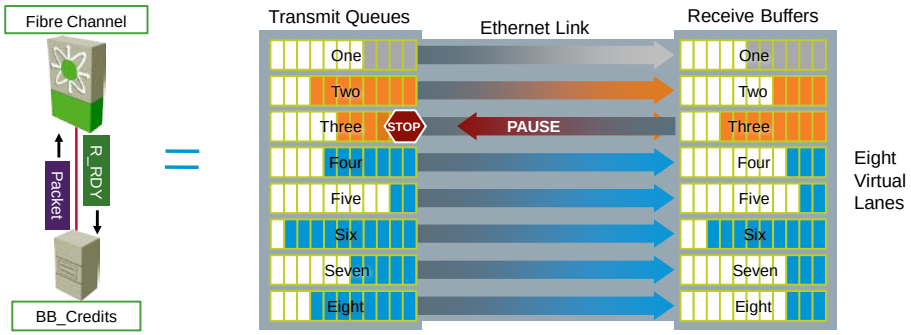
1. Standard buffer-to-buffer flow control is performed for each data frame.
2. After destination N Port B receives a frame, it waits for an R_RDY signal from the F Port.
3. When N Port B receives an R_RDY signal, it sends an acknowledgment (ACK) frame back to N Port A.
4. At the other side of the fabric, the initiator F Port must wait for an R_RDY signal from N Port A.
5. When N Port A sends an R_RDY signal, the F Port transmits the ACK frame.

https://t.me/learningnets

certcollection.net

Priority-Based Flow Control

- Enables lossless Ethernet using the PAUSE feature, based on a CoS as defined in IEEE 802.1p
- When the link is congested, the CoS that is assigned to FCoE is paused so that traffic is not dropped.
- Other traffic that is assigned to other classes of service continues to transmit and relies on upper-layer protocols for retransmission.



© 2012 Cisco and/or its affiliates. All rights reserved.

Priority-based flow control (PFC) is based on class of service (CoS) bits in the IEEE 802.1p standard. PFC enables selective pausing of the traffic that is waiting in the buffer to be sent across the Data Center Bridging (DCB) ISL. A physical link can be split into up to eight selective virtual lanes, using the CoS bits.

In this example, the traffic in the third virtual lane is FCoE traffic that is being paused. PFC is enforced on this virtual lane in order to achieve lossless behavior for other virtual lanes.

A NIC or a converged network adapter (CNA) can support per-priority based flow control. Note that when the interface buffer limit is exceeded, the pause signal is sent for each virtual lane when a queue threshold is exceeded. In this way, congestion management protects traffic on the other virtual lanes.

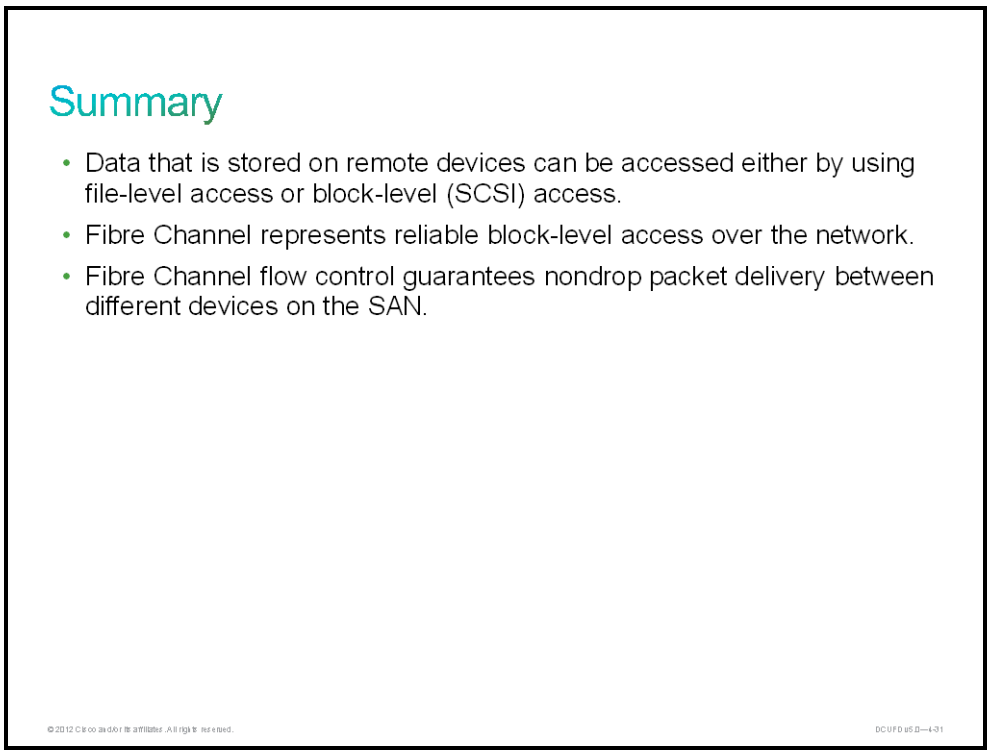
The third lane is the default no-drop virtual lane for FCoE traffic. However, manual adjustment of traffic into different lanes might be required, such as when using unified communications on Cisco Unified Computing System (UCS). Correct CoS mappings must be done so that there are separate virtual lanes for VoIP and FCoE.

https://t.me/learningnets

certcollection.net

Summary

This topic summarizes the primary points that were discussed in this lesson.



https://t.me/learningnets

certcollection.net

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

Lesson 2

Designing SAN

Overview

In this lesson, you will learn how SAN fabrics are designed. Depending on the size of the fabric, different approaches can be taken to accommodate the required number of ports, bandwidth oversubscription, and level of redundancy.

When designing SAN fabrics, redundancy is achieved by having two separate, but identical, fabrics to which the servers connect. Depending on the size of the fabrics, a type of topology is chosen—the core-edge, collapsed core, or edge-core-edge.

Objectives

Upon completing this lesson, you will be able to design reliable, highly available, and flexible SANs. This ability includes being able to meet these objectives:

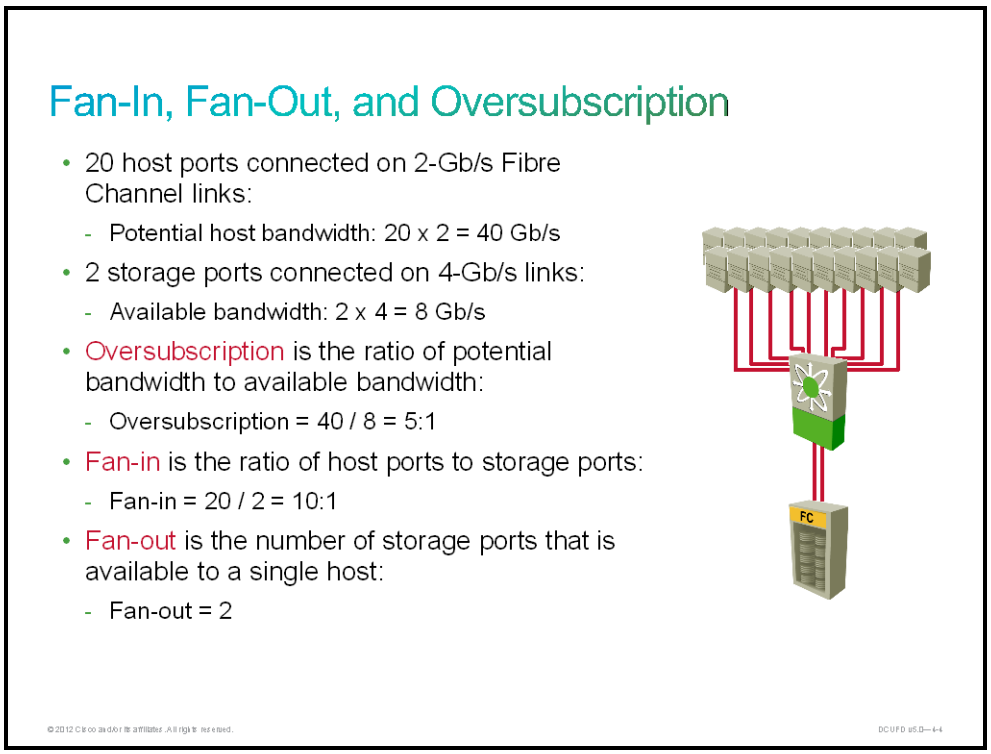
- Explain different storage designs and topologies
- Design SANs using Cisco best practices and Cisco Validated Designs
- Design scalable SANs with provisions for multitenancy

https://t.me/learningnets

certcollection.net

Storage Topologies

This topic describes different storage designs and topologies.



This figure explains the difference between oversubscription, fan-in, and fan-out. There are many definitions but these are from a server perspective:

Oversubscription: Calculates the ratio of potential server bandwidth to available storage bandwidth by multiplying the number of ports at each layer by the link speed. Oversubscription is a necessary requirement in SAN design and is possible because most hosts do not utilize the full bandwidth that is available on each host port. However, it is important to recognize that some applications, like backup and video streaming, require sustained bandwidth and so may have higher link utilization.

Fan-in: Coarse measurement of the number of server ports that share a smaller number of storage ports. It takes no account of bandwidth or the speed of each port, so it is only a rough guide.

Fan-out: Coarse measurement of the number of storage ports that are available to a single host. It is an indication of the number of paths that a host can take to reach the storage logical unit number (LUN) and can be used to check for high availability.

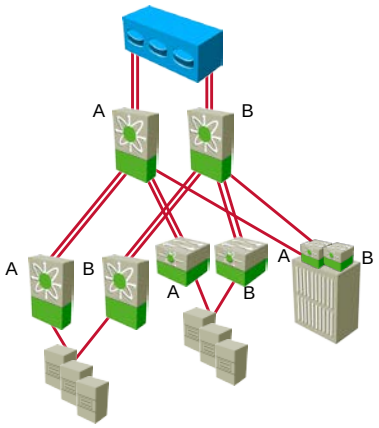
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

Core-Edge

- Traditional SAN design for growing SANs
- High-density directors in core and fabric switches; directors or blade switches on edge
- Predictable performance
- Scalable growth up to core and ISL capacity
- Servers at the edge:
  - End and middle of row with directors
  - Top of rack with fabric switches
  - Blade chassis with blade switches



© 2012 Cisco and/or its affiliates. All rights reserved.

Core-edge is a traditional SAN design that is comparable to a traditional LAN design.

While high-density directors that are located in the core provide high-speed connectivity with storage, fabric and blade switches in the edge provide high port density to provide connectivity to end devices.

Core-Edge Fabric Evaluation



Advantages of core-edge design:

- Highest scalability
- Scalable performance (core switches)
- Scale nondisruptively (edge switches)
- Deterministic latency
- Easy to analyze and tune performance
- Cost-effective for large SANs



Disadvantages of core-edge design:

- Many devices to manage
- Many interconnections to manage
- Large number of ISLs (lower port efficiency)
- Higher oversubscription



© 2012 Cisco and/or its affiliates. All rights reserved.

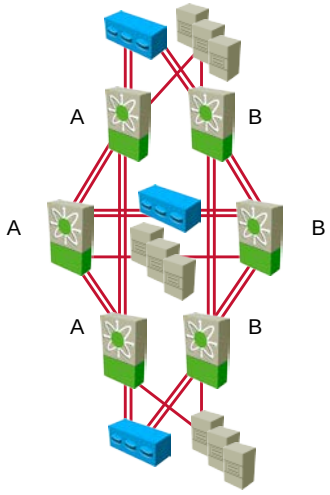
The core-edge design has only one notable disadvantage for a large SAN and that is that it requires many switches and interconnections. While the symmetrical nature of the core-edge design simplifies performance analysis and tuning, there are still many switches to manage.

https://t.me/learningnets

certcollection.net

Collapsed Core-Edge

- SAN design to take full advantage of high-density directors
- Most traffic localized, reducing number of ISLs
- Oversubscription primarily in chassis and line cards
- Potential to scale further than traditional core-edge design



© 2012 Cisco and/or its affiliates. All rights reserved.

Collapsed core-edge utilizes high port density of high-level directors. Separated edge devices are eliminated and most of traffic is local to a single director. This topology has the potential to grow separate edge segments if needed.

Collapsed-Core Fabric Evaluation

Advantages of collapsed fabrics:

- No ISLs
  - All purchased ports available for nodes
  - Increased reliability and simplified management
- Scales easily (hot-swap blade architecture)
- Single management interface
- Highest performance
- Cost-effective

Disadvantages of collapsed fabrics:

- Scalability limitations for very large fabrics
- Potential disaster-tolerance issue

© 2012 Cisco and/or its affiliates. All rights reserved.

The collapsed core topology includes the features of the core-edge topology but delivers required port densities in a more efficient manner. Configuration and management is simple and there are no ports that are used for interswitch links (ISLs).

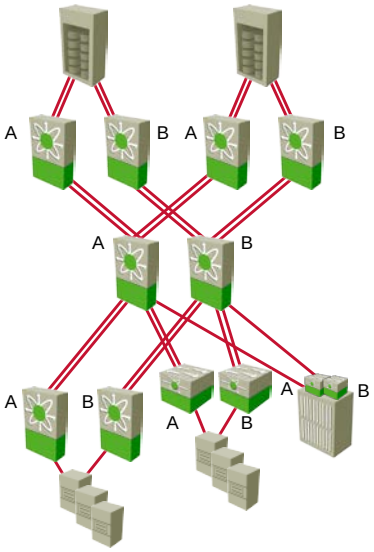
https://t.me/learningnets

certcollection.net

Edge-Core-Edge Design

Edge-core-edge is a less common SAN design option:

- Storage on directors separate, attached to core
- Core directors provide routing and services
- Servers at the edge:
  - End and middle of row with directors
  - Top of rack with fabric switches
  - Blade chassis with blade switches



© 2012 Cisco and/or its affiliates. All rights reserved.

Edge-core-edge is a SAN design that is mostly used when there are several storage blocks. This design is common to cases of consolidating two SANs to a new common SAN.

Edge-Core-Edge Fabric Evaluation

Advantages of edge-core-edge fabrics:

- Scales easily on storage side
- Can manage large number of devices
- High flexibility and scalability

Disadvantages of edge-core-edge fabrics:

- Complex fabric topology
- Higher cost
- Many devices to manage
- Many interconnections to manage

© 2012 Cisco and/or its affiliates. All rights reserved.

The edge-core-edge topology enables independent scaling of both storage and computing edge.

Disadvantages of edge-core-edge fabrics are complexity and cost. There is a higher number of ISLs and a higher number of devices.

https://t.me/learningnets

certcollection.net

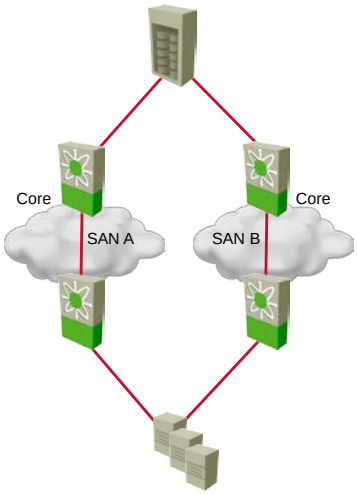
Storage Design Best Practices

This topic describes how to design SANs using Cisco best practices and Cisco Validated Designs.

Design Requirements

SAN Design—Two-Tier Topology

- Edge-core or edge-core-edge topology
- Servers connect to an edge switch in both fabric A and fabric B for redundancy:
  - Multipathing is handled at the server level.
- Storage devices connect to one or more core switches.
- Core switches provide advanced storage services to the edge switches, therefore servicing more servers in the fabric.
- ISLs are designed based on the fan-in ratio of servers to storage and end-to-end oversubscription.
- High availability is achieved in two physically separate, but identical, redundant SAN fabrics.



© 2012 Cisco and/or its affiliates. All rights reserved.

It is common practice in SAN environments to build two separate, redundant physical fabrics (fabric A and fabric B) in case a single physical fabric fails.

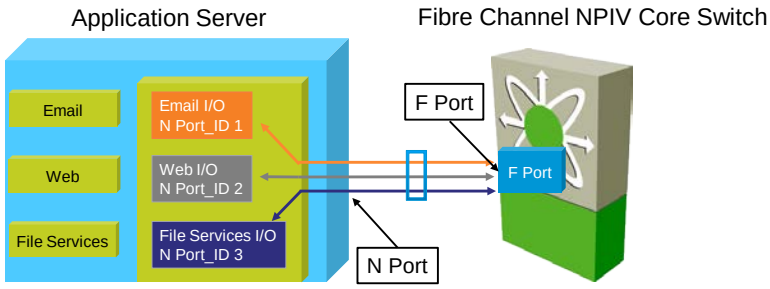
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

N-Port ID Virtualization

- NPIV provides a means to assign multiple FCIDs to a single N Port:
  - A limitation exists in Fibre Channel where only a single FCID can be handed out per F Port. Therefore, an F Port can only accept a single FLOGI.
- Allows multiple applications to share the same Fibre Channel adapter port.
- Usage applies to applications such as VMware, Microsoft Virtual Server, and Citrix.



© 2012 Cisco and/or its affiliates. All rights reserved.

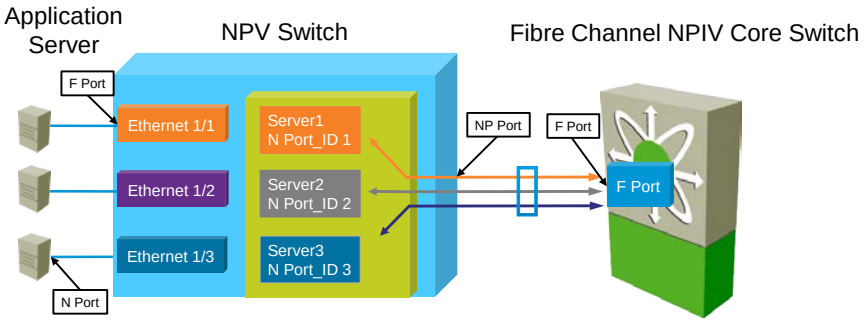
N-Port ID Virtualization (NPIV) enables virtualization of numerous host bus adapters (HBAs) that become virtualized, on a single physical HBA. While traditional Fibre Channel allows a single N Port to be connected to a single F Port, NPIV enables numerous virtual N Ports to connect to a single F Port. Each of the virtual N Ports performs its own fabric login (FLOGI) and receives its own Fibre Channel ID (FCID).

https://t.me/learningnets

certcollection.net

N-Port Virtualizer

- NPV uses NPIV functionality to allow a switch to act like a server, performing multiple logins through a single physical link.
- Physical servers that are connected to the NPV switch log in to the upstream NPIV core switch.
- No local switching is done on a Fibre Channel switch in NPV mode.
- A Fibre Channel edge switch in NPV mode does not take up a domain ID:
  - Helps to alleviate domain ID exhaustion in large fabrics



© 2012 Cisco and/or its affiliates. All rights reserved.

When a Fibre Channel switch operates as an NPV switch, its role is to proxy all traffic from its F Port-facing servers to a Node Proxy (NP) port-facing Fibre Channel fabric. The NPV switch does not provide any fabric services and does not need a domain ID. All Fibre Channel operations are done on the Fibre Channel NPIV core switch. Additionally, no local switching is performed on the NPV switch.

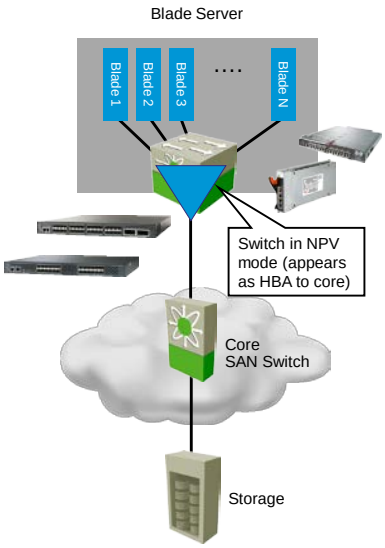
© 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

Enabling Large-Scale Blade Server Deployments—NPV

- NPV converts Fibre Channel switches to HBAs from a connectivity perspective.
- Simplifies deployment and management of large-scale SANs:
  - Reduces number of domain IDs
  - Minimizes interoperability issues with core SAN switches
  - Minimizes coordination between server and SAN administrators

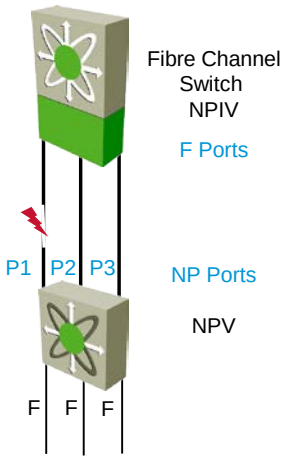


© 2012 Cisco and/or its affiliates. All rights reserved.

Fibre Channel switches, while operating in NPV mode, appear as a single HBA to the rest of the fabric. The domain ID is not used on the switches and the SAN design is simplified.

NPV Design: What Happens When an NP Port Goes Down

- All F Ports are mapped (or pinned) to available NP Ports in a round-robin fashion.
- If an NP Port fails but there are other NP Ports in the same VSAN that are available, the following occurs:
  - N Ports must log in to the fabric again
  - Multiple login attempts are pinned to and distributed among the remainder of the available NP Ports.
- If no NP Port is available, F Ports remain in the "down" state and wait for an NP Port to come up.
- When the failed NP Port comes back up, the logins are *not* redistributed (so as to avoid disruption).



© 2012 Cisco and/or its affiliates. All rights reserved.

Failure of the NP port is managed in two ways. If there is an available NP port in the same virtual SAN (VSAN) where the failed port exists, the NPV switch will login again to the fabric on the remaining port. If there is no available NP port, the NPV switch will shut down the F Port in order to propagate the failure to the hosts. In that case, the host must continue processing SAN traffic over another SAN.

https://t.me/learningnets

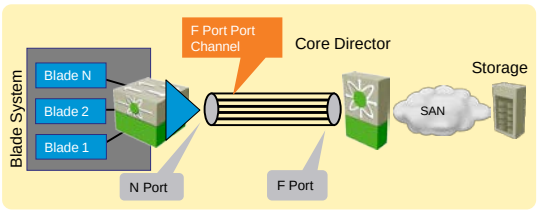
certcollection.net

NPV High Availability: F Port Trunking and Channeling

F Port Port Channels:

- Bundle multiple ports in to one logical link
- Any port, any module
- High availability
- Blade servers are transparent if a cable, port, or line cards fails
- Traffic management
- Higher aggregate bandwidth
- Hardware-based load balancing

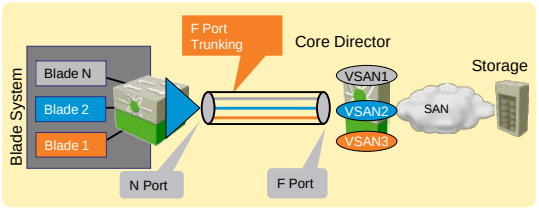
F Port Port Channel



F Port Trunking:

- Partition F Port to carry traffic for multiple VSANs
- Extend VSAN benefits to blade servers
- Separate management domains
- Separate fault isolation domains
- Differentiated services: QoS, security

F Port Trunking



© 2012 Cisco and/or its affiliates. All rights reserved.

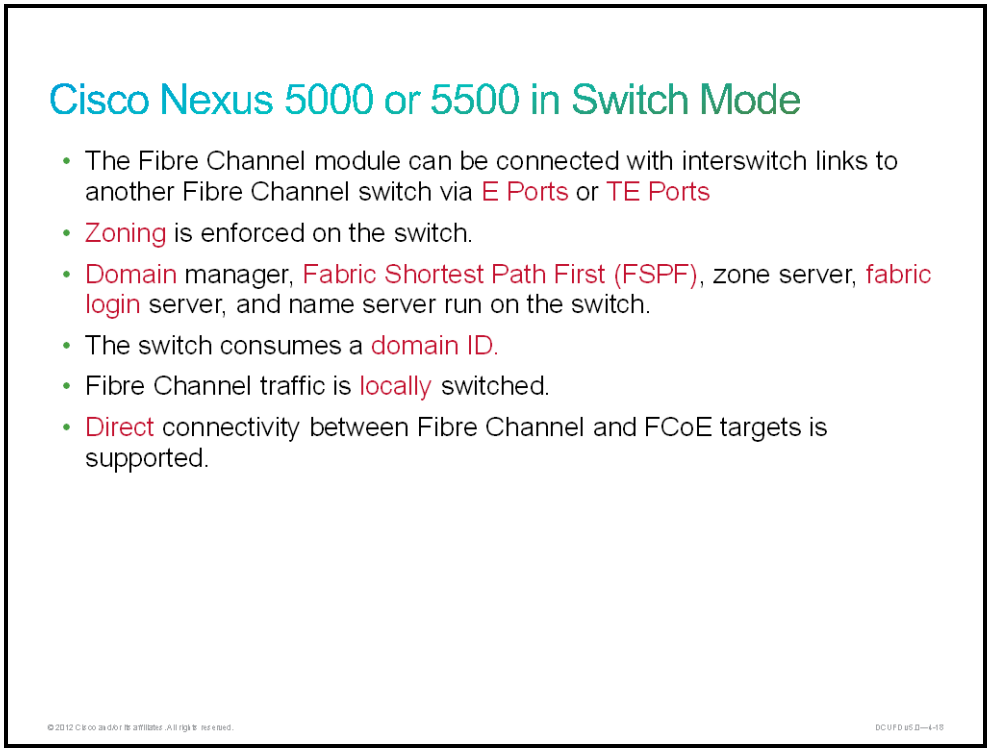
Several physical links between the NPV edge and the NPV core switches can be combined into a single logical port channel. While all links are operational, load balancing provides higher bandwidth between switches. If there is a single link failure, the port channel will continue to operate on the remaining links that provide high availability.

The F Port-facing NPV switch can also be a trunk port that carries traffic for several VSANs between switches. Each VSAN can have its own security and quality of service (QoS) configuration.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

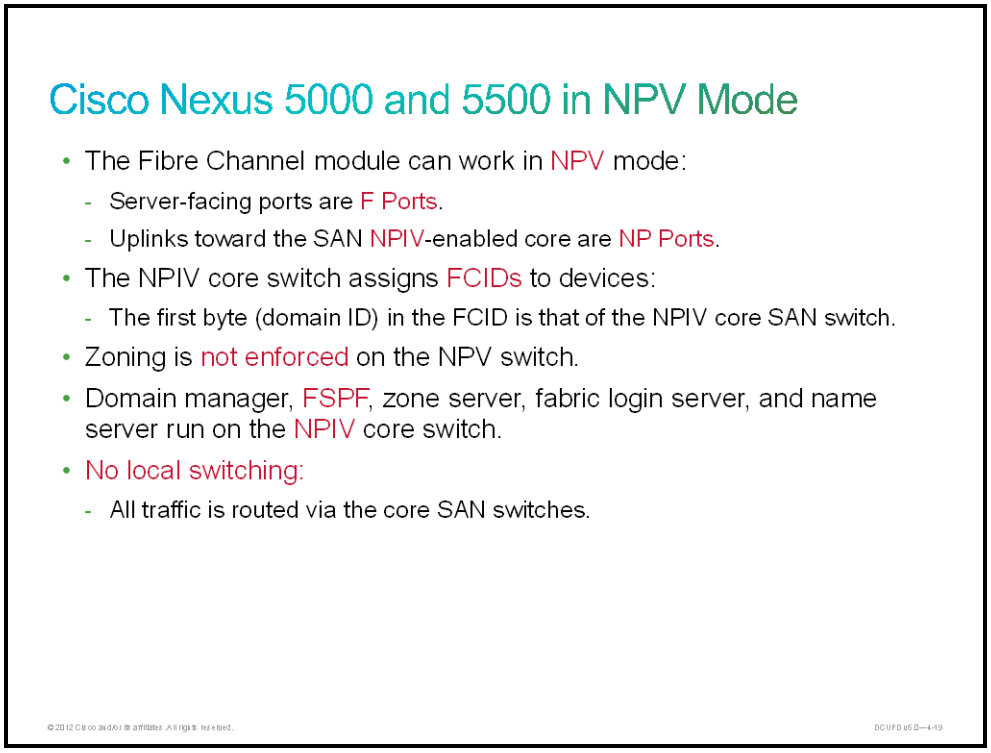
certcollection.net



When Cisco Nexus 5000 and 5500 switches operate as regular Fibre Channel switches, all Fibre Channel services run on them. The switch consumes the domain ID and provides fabric services to hosts that are connected to it. Fibre Channel traffic is locally switched.

https://t.me/learningnets

certcollection.net



While operating in NPV mode, Cisco Nexus 5000 and 5500 switches do not consume the domain ID and their ports are either F Ports toward the hosts or NP ports toward the NPV core switch. Traffic is not locally switched, but it is forwarded toward the NPV core switch.

When the switch configuration is changed from Fibre Channel to NPV, the configuration is erased and the switch is reloaded.

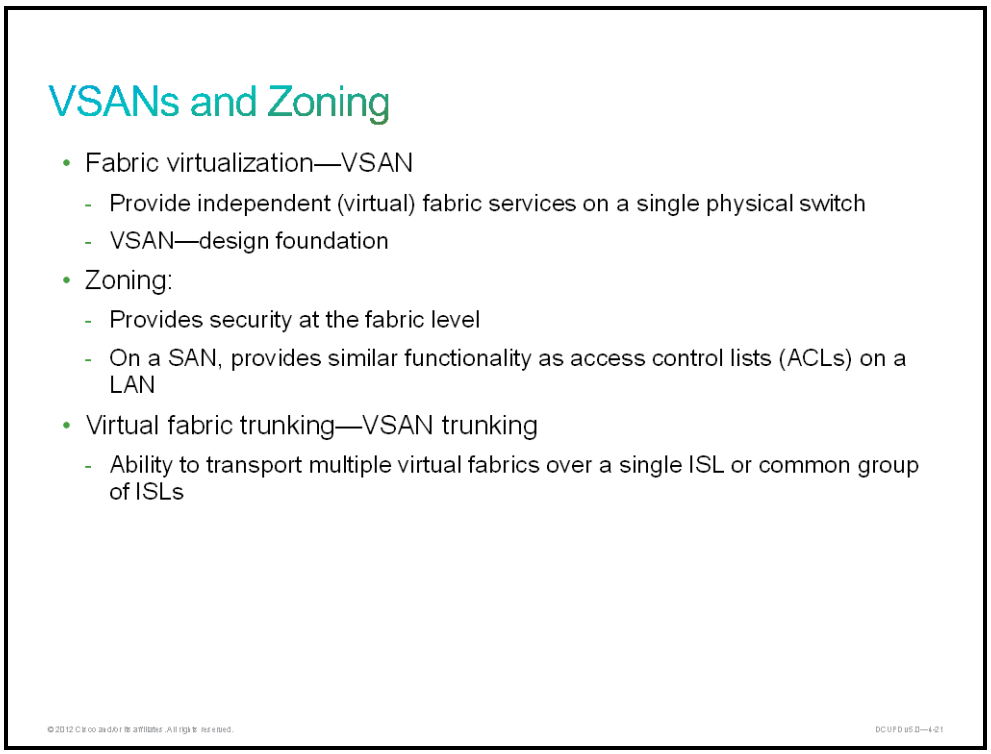
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

Multitenant SANs

This topic explains how to design scalable SANs with provisions for multitenancy.



VSANs improve consolidation and simplify management by allowing for more efficient SAN utilization by creating hardware-based isolated fabrics within a physical infrastructure.

All fabric services are contained within the VSAN and provide independent fabrics while using common hardware.

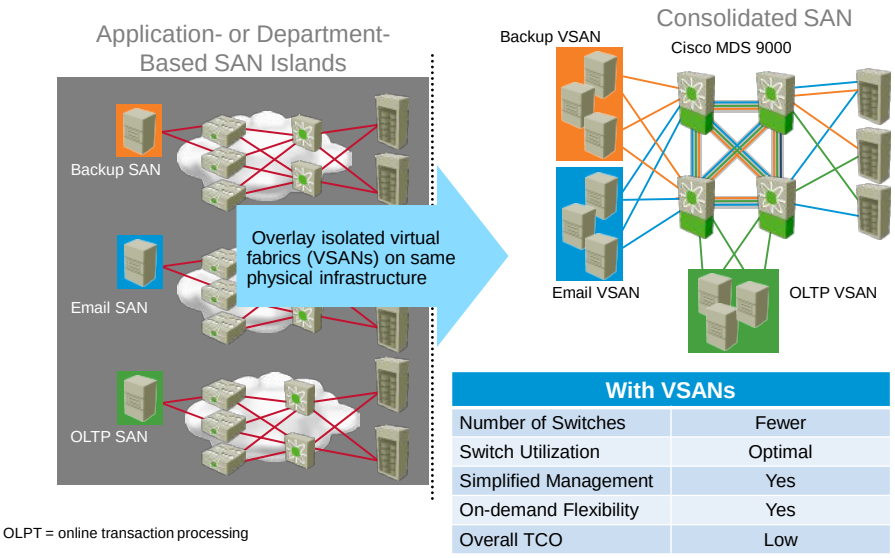
Zoning provides security on the fabric level by defining groups of devices that can communicate with each other. Zoning on the SAN provides functionally that is similar to access lists on a LAN.

VSAN trunking provides the possibility of carrying traffic from different VSANs over a common physical link.

https://t.me/learningnets

certcollection.net

VSANs—Essential for SAN Consolidation



With VSANs	
Number of Switches	Fewer
Switch Utilization	Optimal
Simplified Management	Yes
On-demand Flexibility	Yes
Overall TCO	Low

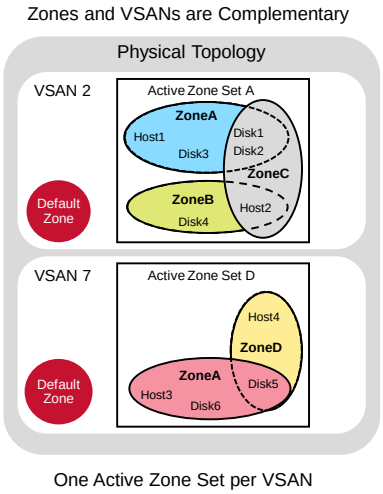
OLPT = online transaction processing

© 2012 Cisco and/or its affiliates. All rights reserved.

VSANs are the cornerstone of SAN consolidation by providing a common physical topology for several SAN fabrics. Fabric utilization becomes better, management is simpler, and total cost of ownership (TCO) is lowered.

Zoning and VSANs

- Hierarchical relationship
  - First, assign physical ports to VSANs.
  - Then configure independent zones per VSAN.
  - VSANs only change when ports are needed per virtual fabric.
  - Zones can change frequently (such as backup).
- Zones provide added security and allow sharing of device ports.
- Zone membership is configured as follows:
  - Port world wide name (pWWN)—device
  - Fabric world wide name (fWWN)—fabric
  - Fibre Channel ID (FCID)
  - Fibre Channel alias (FC alias)
  - IP address
  - Domain ID or port number
  - Interface



© 2012 Cisco and/or its affiliates. All rights reserved.

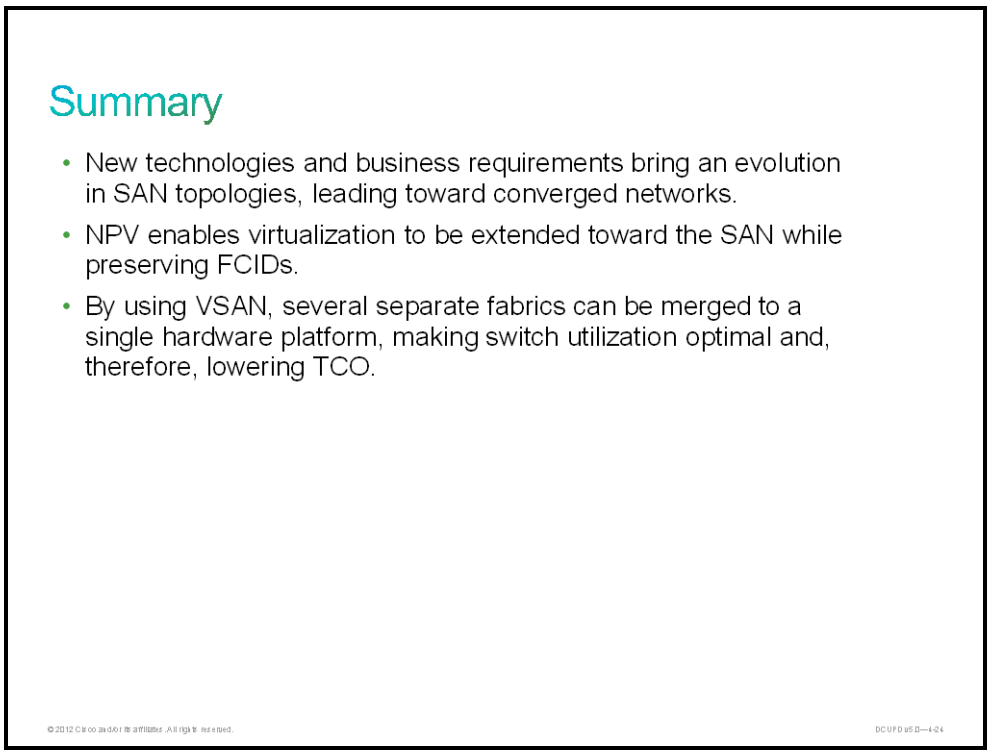
Zoning provides security on the SAN by limiting reachability among different devices on a single fabric. Zoning is contained within the VSAN and it must be configured on each VSAN independently.

https://t.me/learningnets

certcollection.net

Summary

This topic summarizes the primary points that were discussed in this lesson.



References

For additional information, refer to these resources:

- Cisco Validated Design Program at http://www.cisco.com/go/cvd
- Storage Networking at http://www.cisco.com/go/storage

https://t.me/learningnets

certcollection.net

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

Lesson 3

Designing Unified Fabric

Overview

This lesson explains how Fibre Channel traffic fits in unified fabric topology by using the Fibre Channel over Ethernet (FCoE) protocol.

Objectives

Upon completing this lesson, you will be able to design unified fabric. This ability includes being able to meet these objectives:

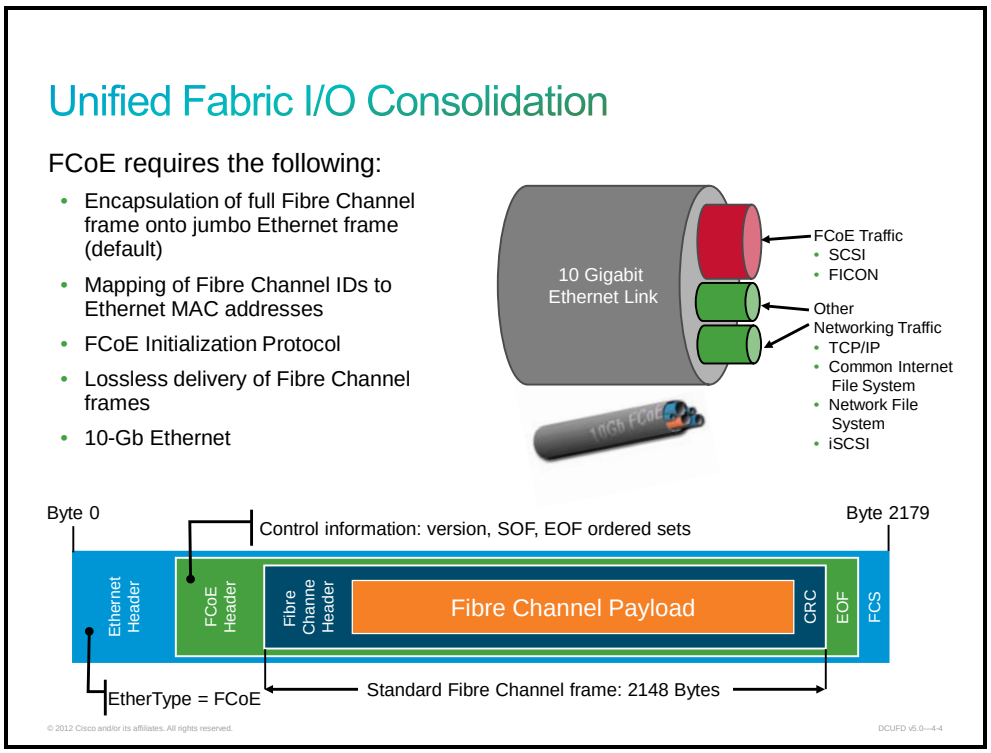
- Explain flow control when using FCoE
- Explain the use of FIP
- Describe different design options with unified fabric networks
- Design unified fabric deployments with FEXs

https://t.me/learningnets

certcollection.net

Unified Fabric

This topic describes flow control when using FCoE.



FCoE is a protocol that is based upon the Fibre Channel layers that are defined by the ANSI T11 committee. It replaces the lower layers of Fibre Channel with unified I/O.

There are several minimum requirements for FCoE:

- Jumbo frames, so that an entire Fibre Channel frame (length of 2180 bytes) can be carried in the payload of a single Ethernet frame
- The mapping of Fibre Channel port world wide name (pWWN) addresses to Ethernet MAC addresses
- An FCoE Initialization Protocol (FIP) that provides login for Fibre Channel devices across a unified fabric
- Lossless delivery of Fibre Channel frames
- A minimum 10-Gb/s Ethernet platform

FCoE traffic consists of a Fibre Channel frame that is encapsulated within an Ethernet frame. The Fibre Chanel frame payload may in turn carry Small Computer Systems Interface (SCSI) messages and data, or in the future, fiber connectivity (FICON) for mainframe traffic.

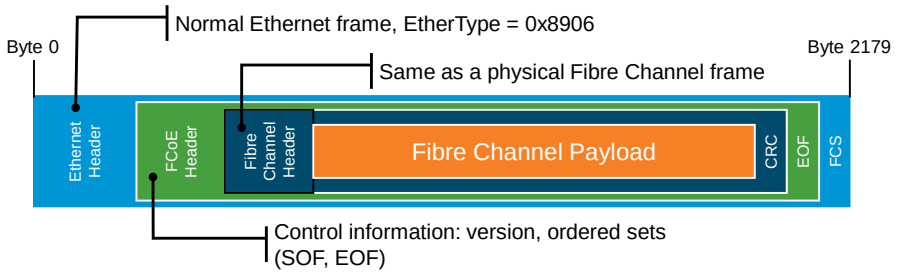
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

Fibre Channel over Ethernet

- Requires 10 Gigabit Ethernet
- Lossless Ethernet
  - Matches the lossless behavior guaranteed in Fibre Channel
- Ethernet jumbo frames
  - Maximum Fibre Channel frame = 2112 bytes



© 2012 Cisco and/or its affiliates. All rights reserved.

FCoE is an extension of Fibre Channel (and its operating model) onto a lossless Ethernet fabric. FCoE requires 10 Gigabit Ethernet and maintains the Fibre Channel operation model, which provides seamless connectivity between two networks.

FCoE positions Fibre Channel as the storage networking protocol of choice and extends the reach of Fibre Channel throughout the data center to all servers. Fibre Channel frames are encapsulated into Ethernet frames with no fragmentation, which eliminates the need for higher-level protocols to reassemble packets.

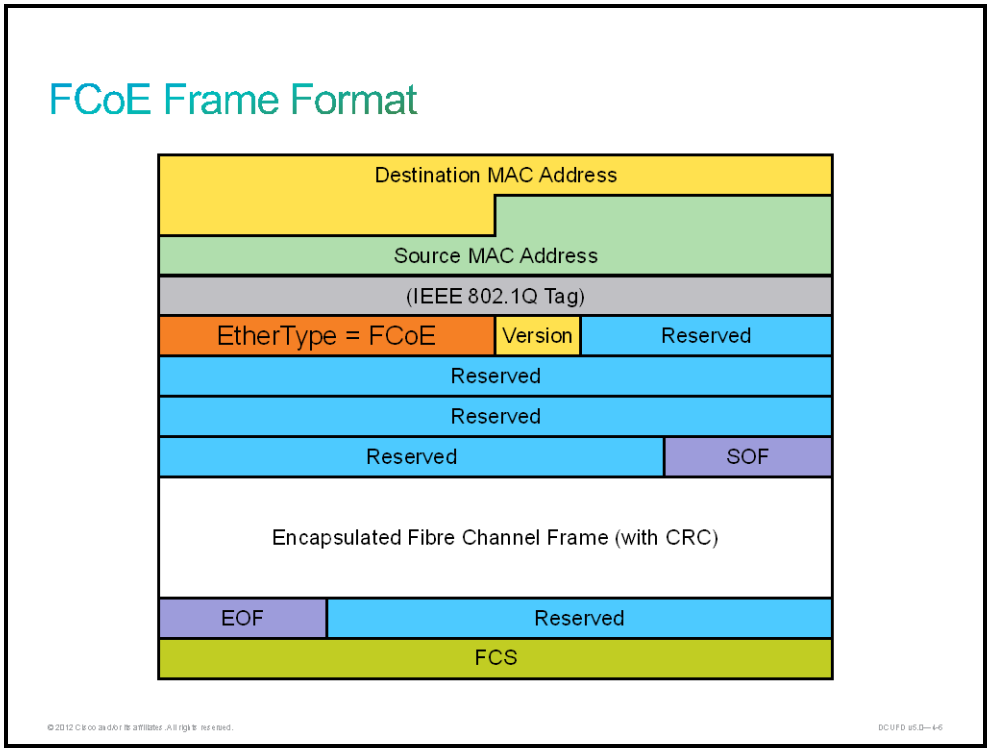
Fibre Channel overcomes the distance and switching limitations that are inherent in SCSI. Fibre Channel carries SCSI as its higher-level protocol. SCSI does not respond well to lost frames, which can result in significant delays when recovering from a loss. Because Fibre Channel carries SCSI, it inherits the requirement for an underlying lossless network.

FCoE transports native Fibre Channel frames over an Ethernet infrastructure, which allows existing Fibre Channel management modes to stay intact. One FCoE prerequisite is for the underlying network fabric to be lossless.

Frame size is a factor in FCoE. A typical Fibre Channel data frame has a 2112 byte payload, a header, and a frame check sequence (FCS). A classic Ethernet frame is typically 1.5 KB or less. To maintain good performance, FCoE must utilize jumbo frames (or the 2.5 KB “baby jumbo”) to prevent a Fibre Channel frame from being split into two Ethernet frames.

https://t.me/learningnets

certcollection.net



FCoE encapsulates a Fibre Channel frame within an Ethernet frame.

The first 48 bits in the frame are used to specify the destination MAC address and the next 48 bits specify the source MAC address. The 32-bit IEEE 802.1Q tag provides the same function as for virtual LANs, allowing multiple virtual networks across a single physical infrastructure.

FCoE has its own EtherType value, as designated by the next 16 bits, followed by the 4-bit version field. The next 100 bits are reserved and are followed by the 8-bit start of frame (SOF) and then the actual Fibre Channel frame. The 8-bit end of frame (EOF) delimiter is followed by 24 reserved bits. The frame ends with the final 32 bits, which are dedicated to the FCS function that provides error detection for the Ethernet frame.

Note	The source MAC address and destination MAC address change on every hop.

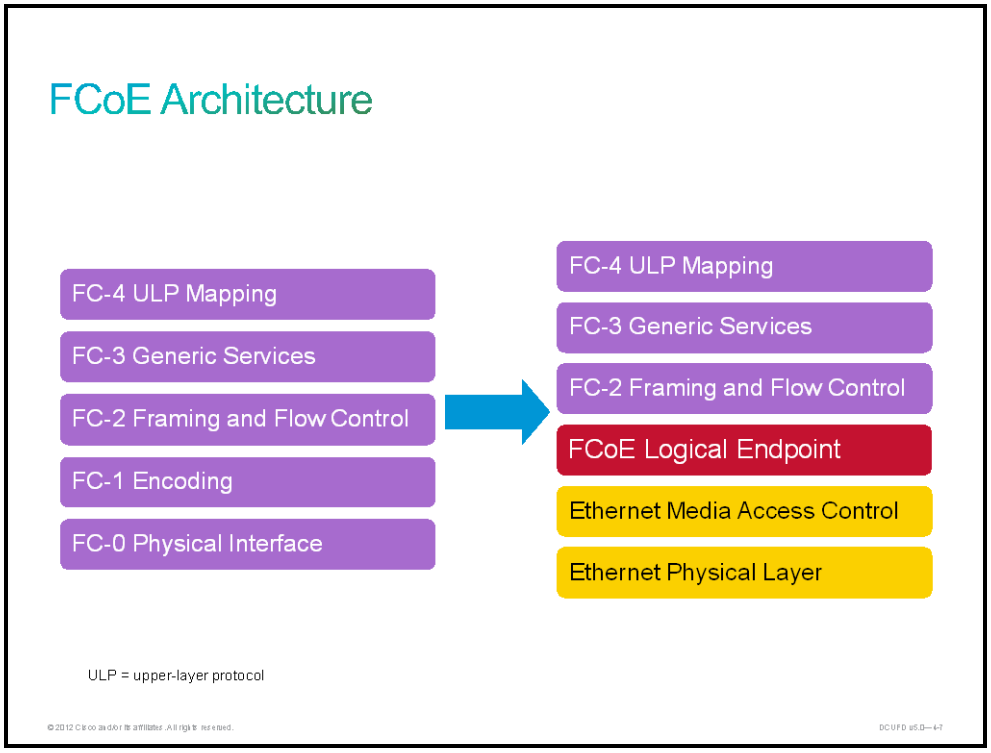
The encapsulated Fibre Channel frame comprises the original 24-byte Fibre Channel header and the data that is being transported (including the Fibre Channel cyclic redundancy check [CRC]). The CRC is used for error detection. The Fibre Channel header is maintained so that when a traditional Fibre Channel SAN is connected to an FCoE-capable switch, the frame is de-encapsulated and handed off seamlessly. This capability enables FCoE to integrate with existing Fibre Channel SANs without the need of a gateway.

Using IEEE 802.1Q tags, Ethernet can be configured with multiple virtual LANs (VLANs) that partition the physical network into multiple separate and secure virtual networks. Using VLANs, FCoE traffic can be separated from IP traffic so that the two domains are isolated, and one network cannot be used to view traffic on the other.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

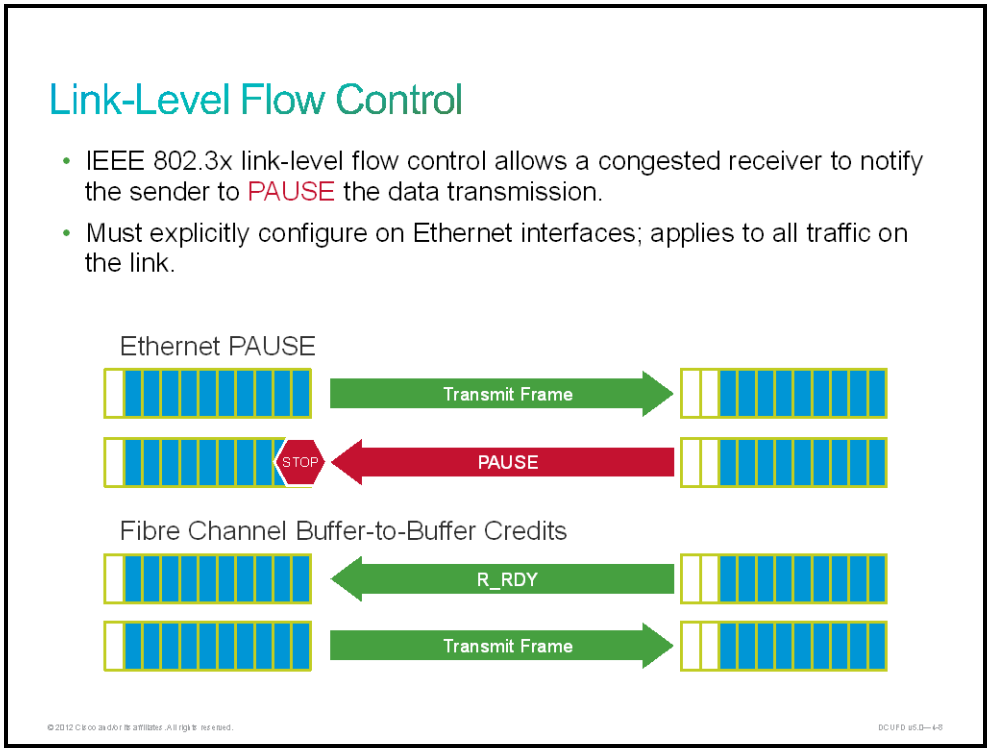
certcollection.net



The FCoE Logical Endpoint (FCoE LEP) is responsible for the encapsulation and de-encapsulation that is necessary to transport Fibre Channel frames over Ethernet. The figure shows that the FCoE LEP has the standard Fibre Channel layers starting with FC-2 and continuing up the Fibre Channel Protocol (FCP) stack. This gives the appearance to the higher-level system functions that the FCoE network is, in fact, a standard Fibre Channel network. This allows all of the same tools that are used in native Fibre Channel to be used in an FCoE environment. Below the FCoE LEP is the standard Ethernet media and physical layers for 10 Gigabit Ethernet with enhancements that allow Ethernet to be lossless. Using the Ethernet standards allows FCoE to take full advantage of a significant amount of existing technology.

https://t.me/learningnets

certcollection.net



Link-level flow control is required for a lossless fabric. Ethernet and Fibre Channel already contain mechanisms for link-level flow control with the Ethernet PAUSE feature and buffer-to-buffer flow control by Fibre Channel. Each of these techniques stops traffic on the entire link, which limits their usefulness in a unified fabric. It should also be noted that not all upper-level protocols require or desire a lossless fabric. TCP, for example, requires packet loss for congestion management.

The IEEE 802.3x link-level flow control capability allows a congested receiver to communicate to the far end, asking it to pause its data transmission for a short period of time. The link-level flow control feature applies to all traffic on the link.

Transmit and receive directions are separately configurable. By default, link-level flow control is disabled for both directions.

On the Cisco Nexus 5000 and 5500 switches, Ethernet interfaces do not automatically detect the link-level flow control capability. You must explicitly configure the capability on the Ethernet interfaces.

On each Ethernet interface, the switch can either enable priority-based flow control (PFC) or link-level flow control, but not both.

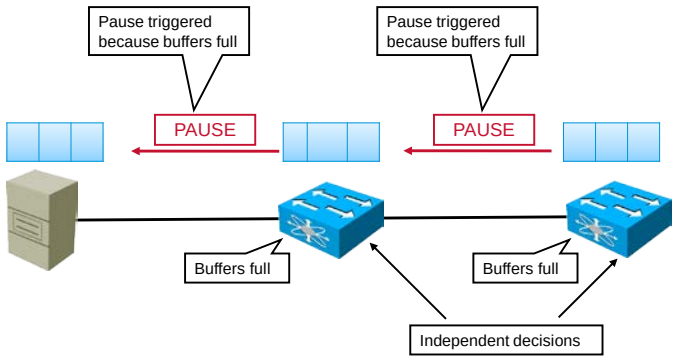
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

Link-Level Flow Control: Pause Propagation

- PAUSE is a hop-by-hop mechanism.
- Each hop accepts pause frames and sends pause frames independently based on available buffer space.



© 2012 Cisco and/or its affiliates. All rights reserved.

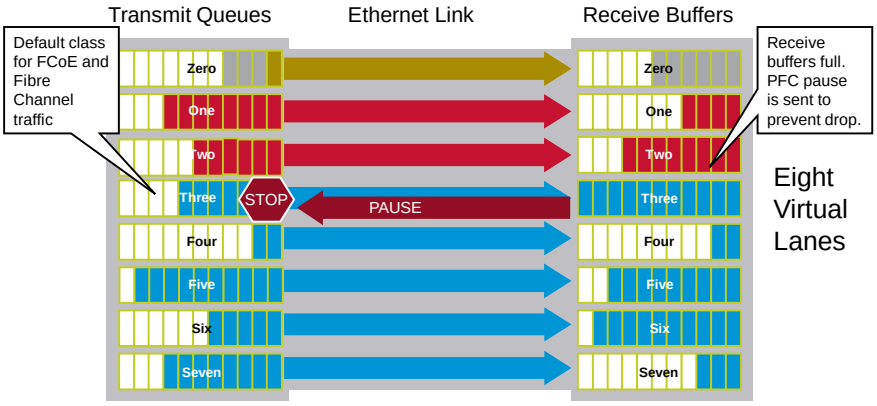
When link-level flow control is enabled, it works on a hop-by-hop basis, where each device initiates a pause that is based on its available buffers. This pause is eventually pushed back to the end devices as buffers begin to fill.

The goal is to suspend the transmission of frames so that the receiver does not drop them due to congestion. When the receiving queue reaches the threshold, the switch sends a pause message back to the sender. The pause is propagated back to the servers that are using the same congested link.

https://t.me/learningnets

certcollection.net

Priority Flow Control



- PAUSE control is used on virtual lane (flow) level.

© 2012 Cisco and/or its affiliates. All rights reserved.

PFC allows you to apply the pause functionality to specific classes of traffic on a link instead of all traffic on the link. PFC applies pause functionality that is based on the IEEE 802.1p class of service (CoS) value. When the switch enables PFC, it communicates to the adapter to which CoS values to apply the pause.

Ethernet interfaces use PFC to provide a lossless server to no-drop system classes. PFC implements pause frames on a per-class basis and uses the IEEE 802.1p CoS value to identify the classes that require lossless services.

In the switch, each system class has an associated 802.1p CoS value that is assigned by default or is configured. If you enable PFC, the switch sends the no-drop CoS values to the adapter, which then applies PFC to these CoS values.

The default CoS value for the FCoE system class is “3.” This value is configurable.

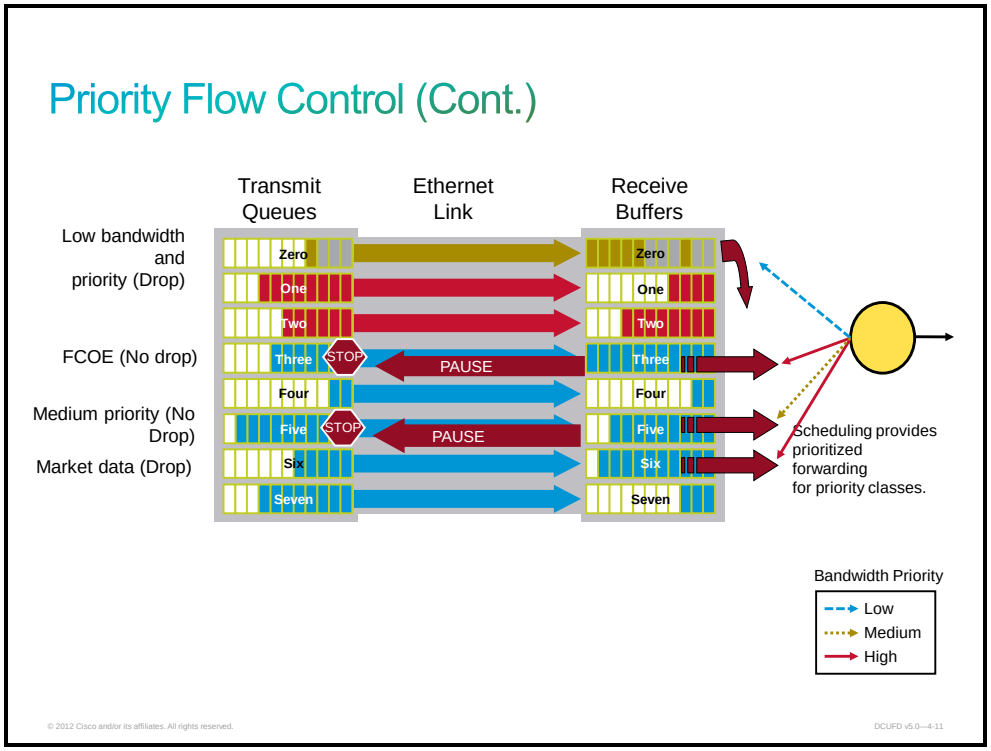
By default, the switch negotiates to enable the PFC capability. If the negotiation succeeds, PFC is enabled and link-level flow control remains disabled regardless of its configuration settings. If the PFC negotiation fails, you can either force PFC to be enabled on the interface or you can enable IEEE 802.3x link-level flow control.

If you do not enable PFC on an interface, you can enable 802.3x link-level pause, which by default is disabled.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

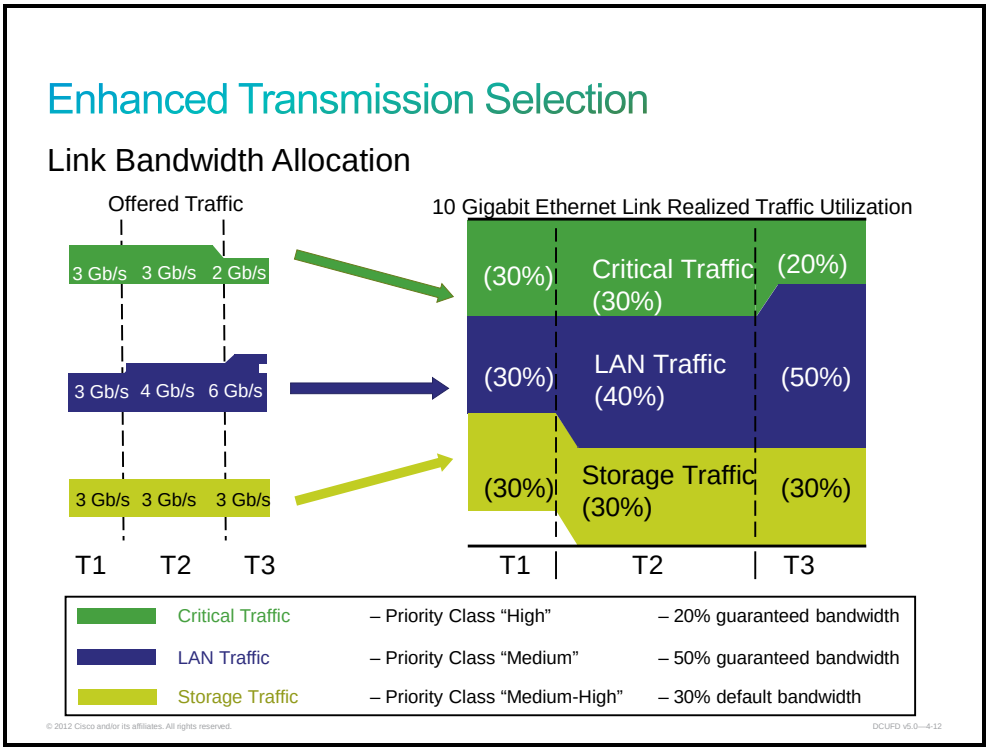
certcollection.net



This figure shows how multiple classes of traffic can use the PFC mechanism. Low-priority traffic, which is also low bandwidth, can be dropped if buffers fill, relying on TCP retransmissions for recovery. FCoE and Internet Small Computer Systems Interface (iSCSI) traffic can be set to the no-drop FCoE class. This setting prevents retransmissions and applies suitable flow control that is required for storage traffic. Medium-priority data traffic can be set to a no-drop class to pause when buffers reach high levels. When combined with low-medium bandwidth allocation, it pauses to allow more bandwidth for latency-sensitive applications. Such applications include voice or market data, as shown in the figure. There is no reason to pause voice, because jitter or latency is not tolerated by these applications. Pausing other traffic streams when buffers fill benefits these latency-sensitive applications.

https://t.me/learningnets

certcollection.net



Enhanced Transmission Selection (ETS) is an IEEE 802.1Qaz standard that enables optimal bandwidth management of virtual links. ETS allows differentiation among traffic of the same priority class, therefore creating priority groups.

ETS is also called "priority grouping." Eight distinct virtual link types can be created by implementing PFC. It can be advantageous to have different traffic classes that are defined within the different PFC types.

ETS enables these differentiated treatments within the same priority class of PFC. This provides prioritized processing that is based on bandwidth allocation, low latency, or best effort. This results in per-group traffic class allocation.

For example, an Ethernet class of traffic may have a high-priority designation and a best effort within that same class. ETS allows differentiation between traffic of the same priority class, therefore creating priority groups.

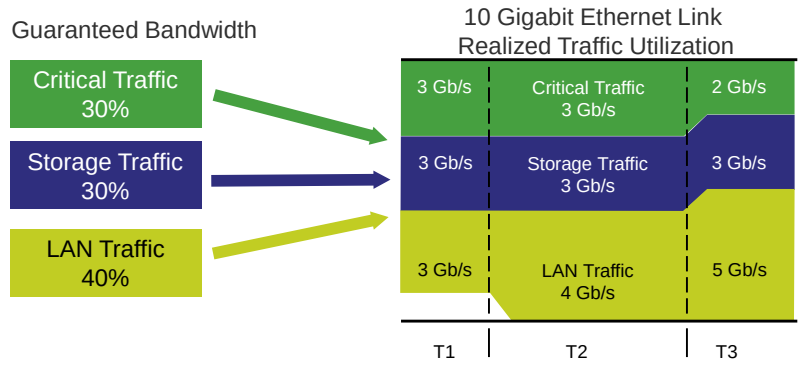
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

Enhanced Transmission Selection (Cont.)

- Bandwidth is guaranteed but can be used by other classes if not in use.
- Enables intelligent sharing of bandwidth between traffic classes.
- Proposed standard 802.1Qaz ETS.
- Bursty traffic in managed classes can exist alongside strict priority traffic classes.



© 2012 Cisco and/or its affiliates. All rights reserved.

Bandwidth management is an important requirement when consolidating I/O. IEEE 802.1Qaz is a standard that specifies ETS to support the allocation of bandwidth among different traffic classes. When a given load in a traffic class does not fully utilize its allocated bandwidth, this standard allows other traffic classes to use the available bandwidth. This feature helps accommodate the bursty nature of some traffic classes while maintaining bandwidth guarantees. Managed bandwidth classes exist alongside strict priority classes. It is desirable to be able to share bandwidth between priorities carrying bursty high-offered loads rather than servicing them with strict priority, at the same time allowing strict priority for time-sensitive and management traffic requiring minimum latency.

In any switch fabric, traffic from multiple ingress ports can compete for the limited bandwidth in a single egress port. For example, bursts of data traffic could reduce the bandwidth that is available to storage traffic, causing congestion and increased latency. To solve this problem, the IEEE developed ETS, which provides advanced traffic scheduling. This scheduling includes features such as guaranteed minimum bandwidth for certain traffic classes like storage or business-critical traffic.

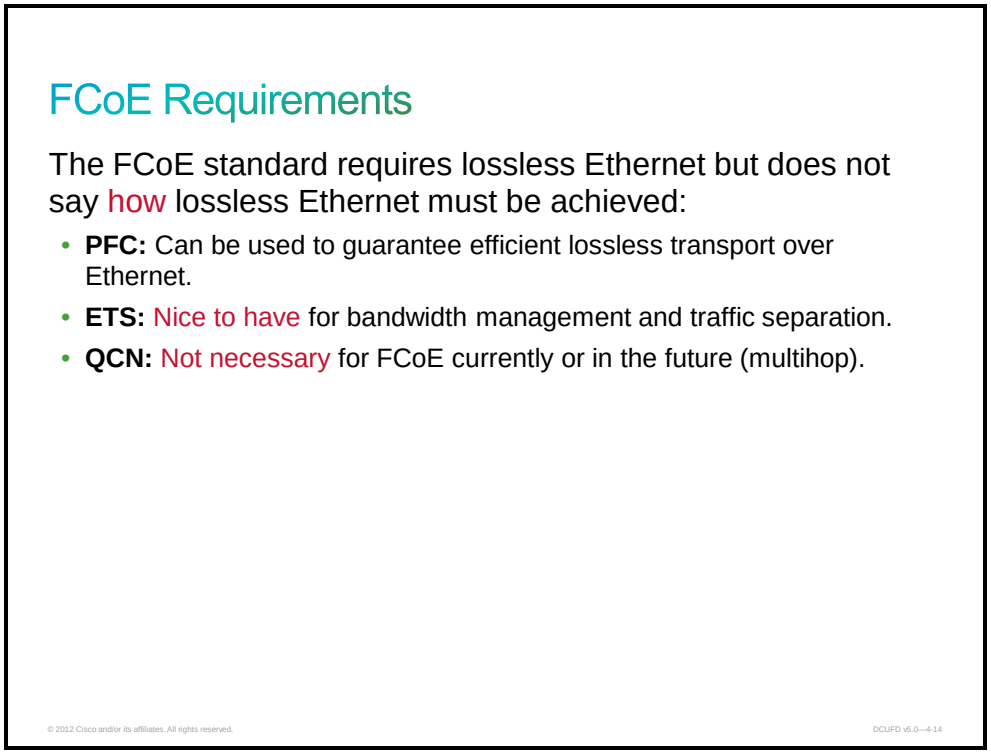
Consider an example with three classes of traffic being offered to a 10-Gb/s link. The guaranteed bandwidth is 30 percent for critical traffic, 30 percent for storage traffic, and 40 percent for LAN traffic.

- At T1, critical, storage, and LAN traffic use 3 Gb/s. At this point, each type of traffic is using its guaranteed bandwidth and the link bandwidth is not yet fully utilized.
- At T2, critical and storage traffic remain the same while LAN traffic increases to 4 Gb/s. At this point, link bandwidth is fully utilized.
- At T3, critical traffic drops to 2 Gb/s, storage traffic remains at 3 Gb/s, and LAN traffic increases to 5 Gb/s. Because one class of traffic (critical) is not using all its reserved bandwidth, another class of traffic (LAN) can use the available bandwidth.

The reserved bandwidth implementation is based on deficit weighted round robin (DWRR) and strict priority can also be configured for a class of service.

https://t.me/learningnets

certcollection.net



Several protocols can be used for providing a non-drop condition on Ethernet:

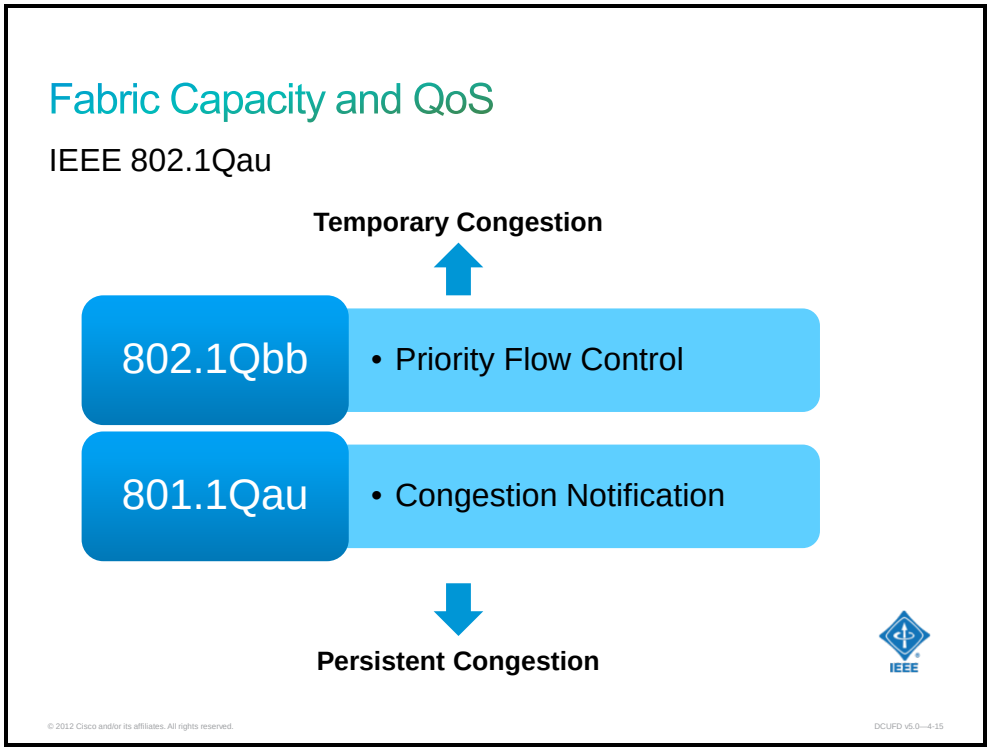
- **Priority-based flow control (PFC):** IEEE 802.1Qbb provides a link-level flow control mechanism that can be controlled independently for each frame priority. The goal of this mechanism is to ensure zero loss under congestion in Data Center Bridging (DCB) networks.
- **Enhanced Transmission Selection (ETS):** IEEE 802.1Qaz provides a common management framework for assignment of bandwidth to frame priorities.
- **Quantized congestion notification (QCN):** IEEE 802.1Qau provides end-to-end congestion management for protocols that are capable of transmission rate limits to avoid frame loss. It should benefit protocols such as TCP that do have native congestion management, because it reacts to congestion in a more timely manner.

For each of these protocols, separate IEEE task groups maintain development of them as part of the DCB set of Ethernet enhancements.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net



The goal of QCN is for switches to have the ability to notify end hosts of any congestion within the network. The hosts can then respond by decreasing the transmission of packets and therefore alleviate the congestion. QCN must be enabled throughout the entire Layer 2 fabric (including the hosts) in order to be effective. While it may work in Layer 2 Ethernet networks, FCoE networks require that FCoE traffic traverse a Fibre Channel Forwarder (FCF) where source and destination MAC addresses are rewritten. This process make it impossible to send QCN messages to an end host MAC after traffic has passed through an FCF.

Because of this limitation, QCN becomes ineffective in current FCoE networks.

https://t.me/learningnets

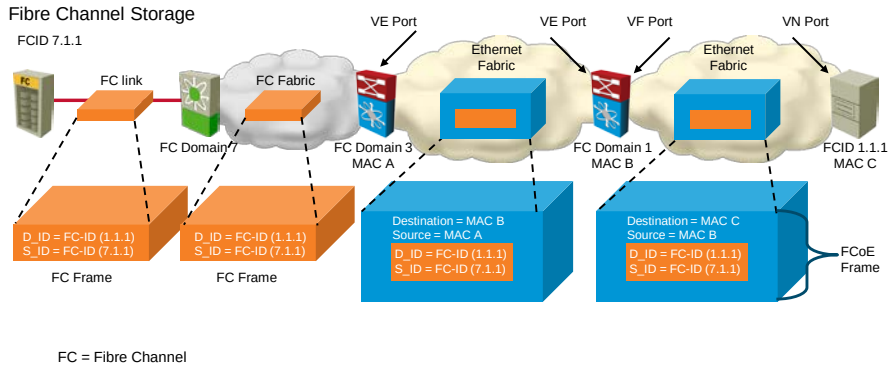
certcollection.net

FCoE Forwarding

FCF

Intermediate switches in the Ethernet cloud

- All are Fibre Channel-aware



© 2012 Cisco and/or its affiliates. All rights reserved.

DCUFD v5.0—4-16

This figure shows a Fibre Channel frame traversing the Fibre Channel and Ethernet cloud.

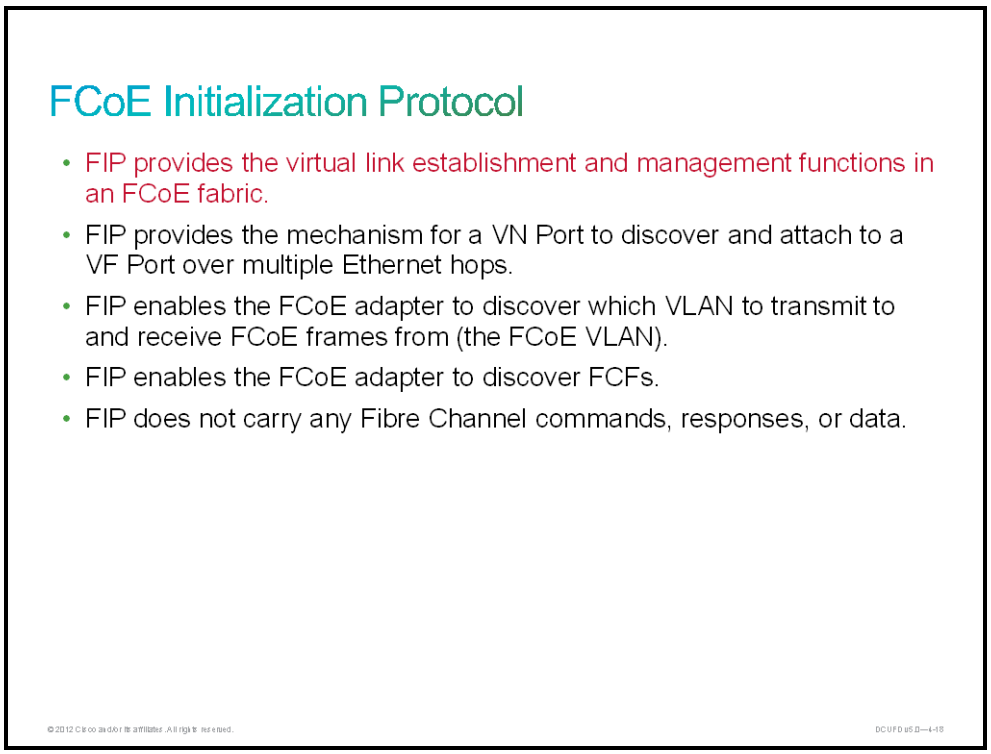
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

FCoE Initialization Protocol

This topic describes the use of FIP.



FIP provides the virtual link establishment and management functions in an FCoE fabric.

Initially, FIP provides the mechanism for a VN Port to discover and attach to a VF Port over a single or multiple Ethernet hops. This is done by the Converged Network Adapter (CNA) that discovers which VLAN is used to transmit and receive FCoE frames (known as the FCoE VLAN).

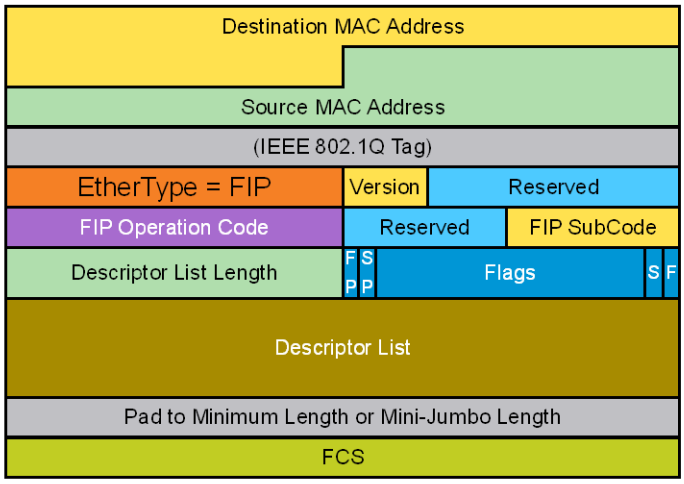
After the FCoE VLAN discovery is complete, the CNA discovers, by using FIP, the FCoE FCFs that are present in the FCoE fabric, and attempts to log in to the SAN fabric through the discovered FCF.

Note	FIP does not carry any Fibre Channel commands, responses, or data. It is used only to establish the FCoE session between the CNA and the FCF. From there onward, the Fibre Channel protocol stack performs the fabric login (FLOGI) and subsequent steps.

https://t.me/learningnets

certcollection.net

FIP Frame Format



FIP is encapsulated in an Ethernet packet with a dedicated EtherType, 0x8914. The packet has a 4-bit version field, the source and destination MAC addresses, a FIP operation code, and a FIP operation subcode. The following describes the FIP operation codes:

- 0x0001:
  - 0x01: Discovery Solicitation
  - 0x02: Discovery Advertisement
- 0x0002:
  - 0x01: Virtual Link Instantiation Request
  - 0x02: Virtual Link Instantiation Reply
- 0x0003:
  - 0x01: FIP Keepalive
  - 0x02: FIP Clear Virtual Links
- 0x0004:
  - 0x01: FIP VLAN Request
  - 0x02: FIP VLAN Notification

Pre-FIP virtual link instantiation consists of two phases—link discovery using the Data Center Bridging Exchange Protocol (DCBX), which is then followed by the FLOGI.

The Cisco Nexus 5000 and 5500 switches are backward-compatible with first-generation CNAs that operate in pre-FIP mode.

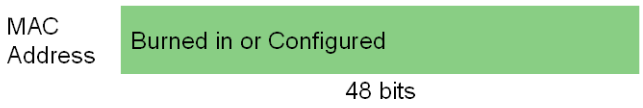
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

FCoE Node Addressing

- The adapter uses the burned-in or configured MAC address.
  - Consistent with the Ethernet model
- The FCF needs to maintain a table to map between MAC addresses and FCIDs.
- The SPMA does not allow for the use of a unique FC-MAP value to designate individual fabrics.
- SPMA support is not required by the FCoE standard.



© 2012 Cisco and/or its affiliates. All rights reserved.

FCoE Nodes (ENodes) use Fibre Channel and Ethernet addressing schemes for the networks to which they attach. There must be some correlation between these addressing schemes.

Server-provided MAC addresses (SPMAs) use the burned-in MAC address or a configured MAC address as the station address for all traffic that is generated or received by an ENode. This technique has some implications for the FCF. It must keep state information mapping the Fibre Channel IDs (FCIDs) to the Ethernet MAC addresses to properly encapsulate Fibre Channel traffic that is destined to an ENode. SPMAs do not allow the use of unique identification within the MAC address to designate independent fabrics operating on the same Ethernet cloud.

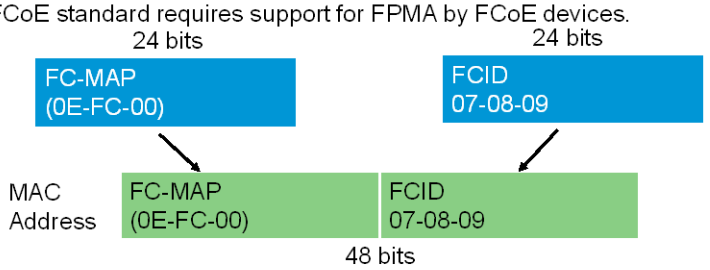
SPMA support is not required by the FCoE standard.

https://t.me/learningnets

certcollection.net

FCoE Node Addressing (Cont.)

- A MAC address is assigned for each FCID.
  - Consistent with the Fibre Channel model
- Uses Organizationally Unique Identifiers (OUIs) with Universal/Local (U/L) bit = 1 (local addressing), called FC-MAPs.
- Multiple FC-MAPs are supported (one per physical SAN).
- No table is needed for translation.
- Multiple MACs may be needed for N-Port ID Virtualization (NPIV).
- The FCoE standard requires support for FPMA by FCoE devices.



Cisco Nexus 5000 and 5500 switches only support FPMA addressing.

© 2012 Cisco and/or its affiliates. All rights reserved.

Fabric-provided MAC addresses (FPMAs) create a direct mapping between the FCID that is assigned by the Fibre Channel fabric services in the FCF, and the Ethernet MAC address that is used as the ENode station address. The 48-bit Ethernet MAC address consists of a fabric-wide FCoE MAC Address Prefix (FC-MAP) value in the high-order 24 bits, with the assigned FCID in the lower-order 24 bits. Fibre Channel traffic can be encapsulated directly in FCoE frames with no table lookup, because the FC-MAP is a known quantity. The destination ID in the Fibre Channel frame (which is sent by the FCF during the FLOGI process) supplies the FCID.

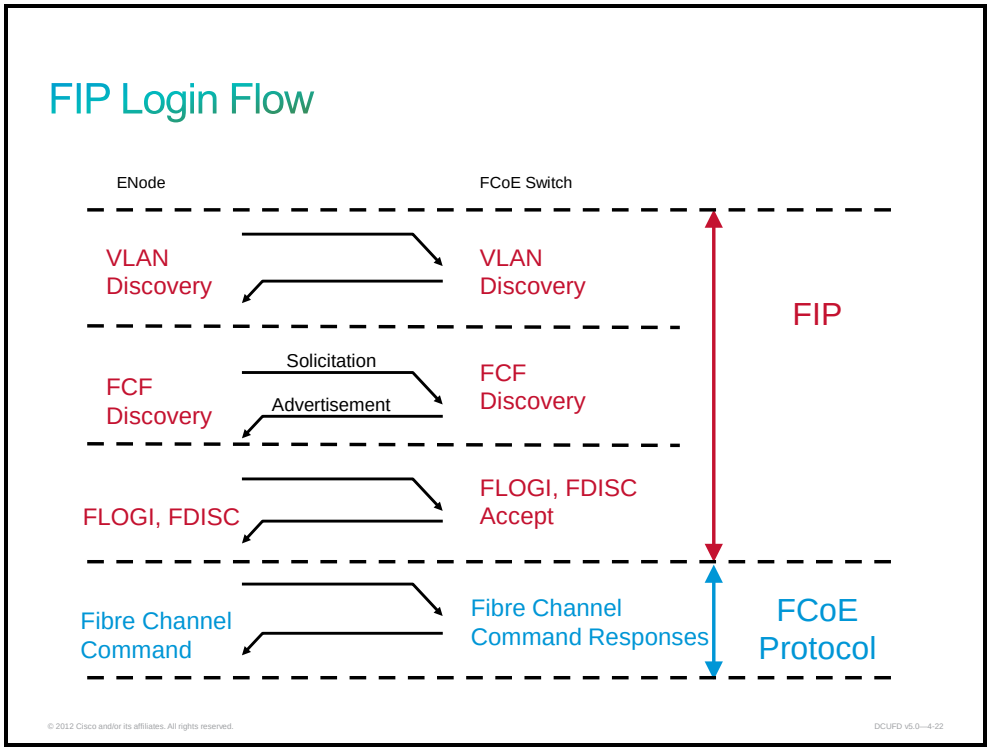
Unfortunately, the FCoE MAC address to be used by the station cannot be determined until a Fibre Channel FLOGI is sent. This address is not available for use as the source MAC address during the FLOGI itself. In addition, a mechanism must be identified to determine the MAC address of the FCF so that the destination MAC address for the FLOGI is known.

An FCoE initialization process is therefore required.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net



FIP defines two discovery protocols as well as a protocol to establish virtual links between VN Ports and VF Ports. This figure shows a typical FIP exchange. The exchange results in the establishment of a virtual link between the VN Port of an ENode and the VF Port of an FCF.

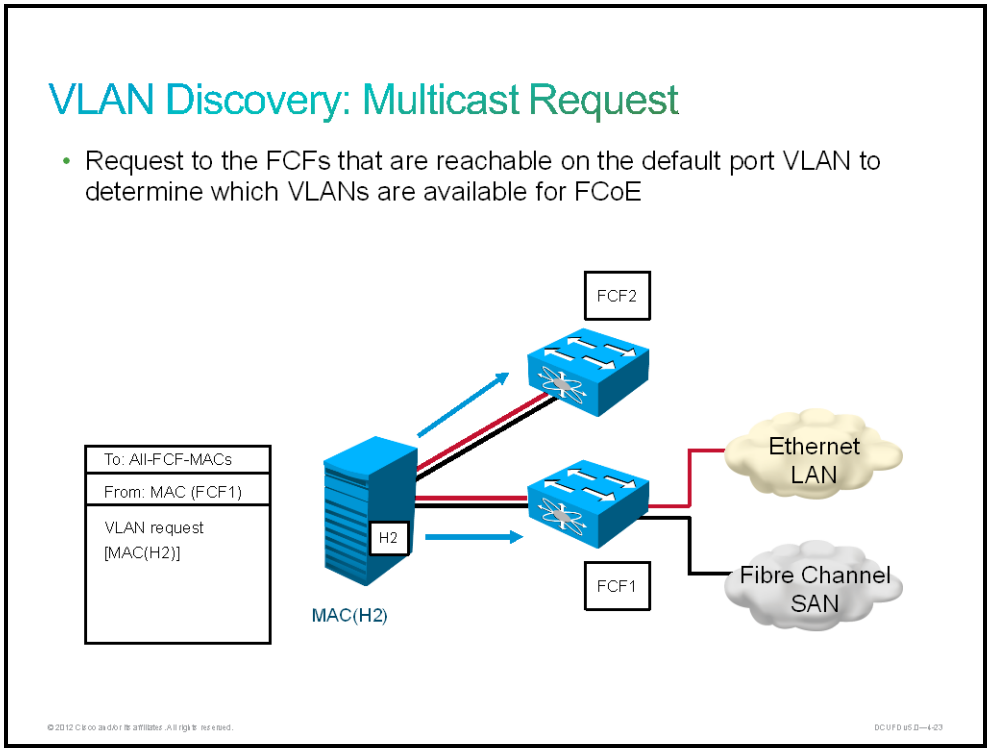
All the protocols are usually initiated by ENodes, although FCFs can generate unsolicited FIP advertisements.

The following should be noted:

- The FIP frames at the top of the figure and the FCoE frames at the bottom of the figure use different EtherTypes and encapsulations.
- FCoE frames encapsulate native Fibre Channel payloads.
- FIP frames describe a new set of protocols that have no reason to exist in native Fibre Channel definitions.
- FIP packets are built using a globally unique MAC address that is assigned to the CNA at manufacturing (called the ENode MAC address).
- FCoE packets are encapsulated using a locally unique MAC address that is unique within the boundaries of the local Ethernet subnet. It is dynamically assigned to the ENode by the FCF as part of the FIP virtual link establishment process.

https://t.me/learningnets

certcollection.net



FIP VLAN discovery determines the FCoE VLAN to use by the FIP protocols, as well as by the FCoE encapsulation for Fibre Channel payloads on the established virtual link. FIP VLAN discovery is performed on the native VLAN. All other FIP protocols run on the discovered FCoE VLAN.

The ENode sends a FIP VLAN discovery request to a multicast MAC address that is called “All-FCF-MACs.” All-FCF-MACs is a multicast MAC address to which all FCFs listen.

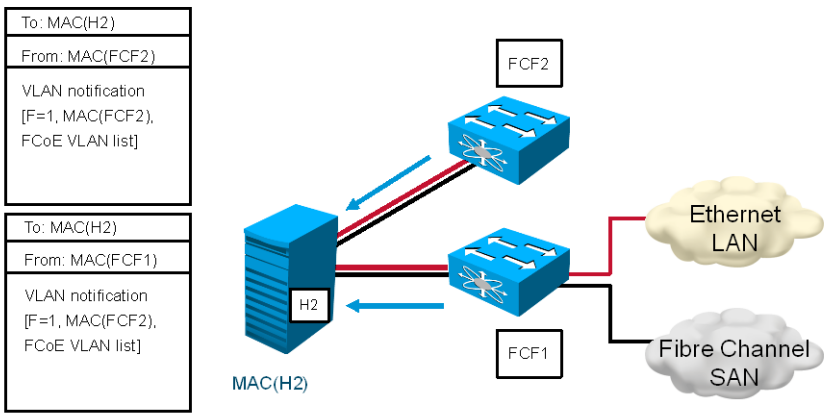
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

VLAN Discovery: Unicast Notification

- One or more FCoE VLANs are provided.

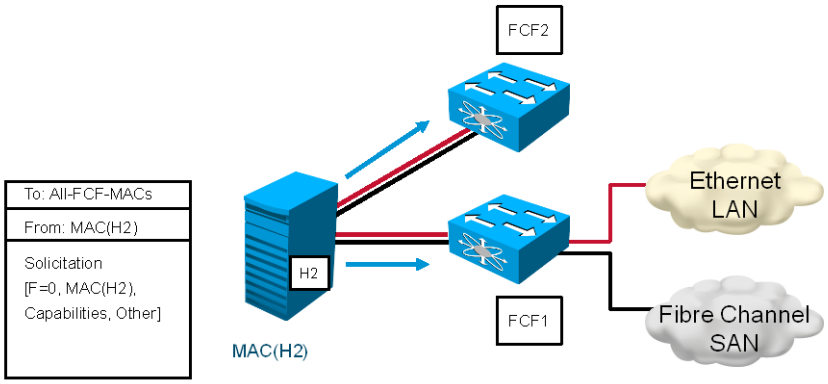


© 2012 Cisco and/or its affiliates. All rights reserved.

After receiving a multicast request, all FCFs respond to the sender with available FCoE VLANs, using unicast traffic.

VLAN Discovery: Multicast Solicitation

- Solicitation identifies VF Port-capable FCF-MACs with compatible addressing capabilities.



© 2012 Cisco and/or its affiliates. All rights reserved.

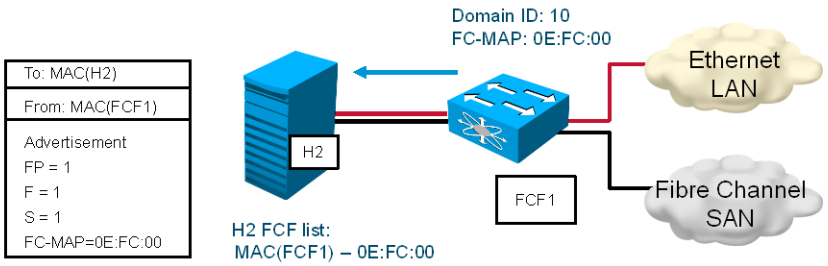
After receiving the list of available FCoE VLANs, the host sends a solicitation request for available capabilities from the fabric.

https://t.me/learningnets

certcollection.net

FCF Discovery Process—Advertisement

- Switch provides the fabric the unique FC-MAP (that is used for the first 3 bytes of the FPMA).
- FP = 1 indicates FPMA-capable.
- F = 1 indicates the frame was generated by an FCF.
- EtherType = FIP

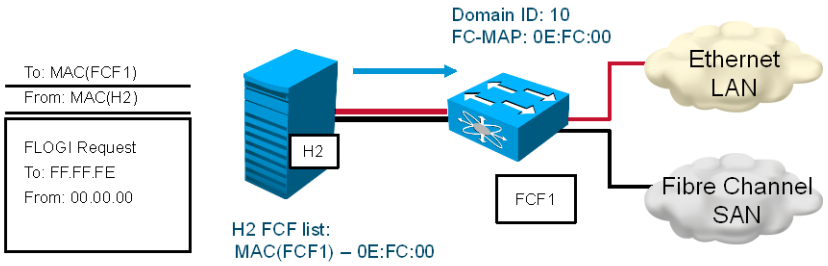


© 2012 Cisco and/or its affiliates. All rights reserved.

The fabric responds with its available capabilities.

Fibre Channel FLOGI Process—Request

- Host continues FIP with FLOGI.
- At this point, host is still using the burned-in MAC address.
- EtherType = FIP



© 2012 Cisco and/or its affiliates. All rights reserved.

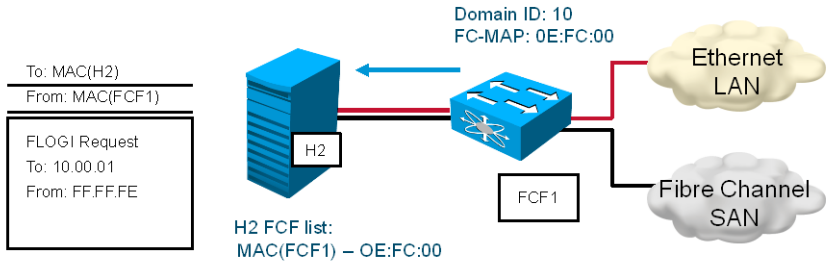
After the ENode discovers all the FCFs and selects one for login, the final step is to inform the selected FCF of the intention to create a virtual link with its VF Port. After this step has been performed, Fibre Channel payloads can start being exchanged on the new virtual link that was just established.

https://t.me/learningnets

certcollection.net

FLOGI Process—Accept

- FCF provides an FCID .
- FCF replies using the host burned-in MAC address.
- EtherType = FIP

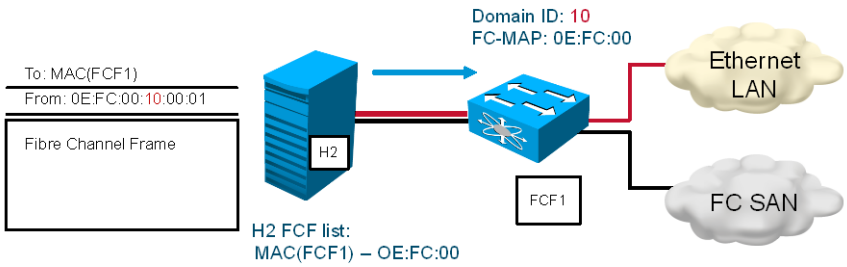


© 2012 Cisco and/or its affiliates. All rights reserved.

The switch assigns an FCID and responds to the host. This FCID is appended to the previously acquired FC-MAP to create the FPMA. The FPMA is used for future communication.

Subsequent Frames

- Host uses the FPMA for subsequent transmissions.
- FPMA is the FCID that is obtained through FLOGI, appended to the FC-MAP that is provided by the FCF.
- EtherType = FCoE



© 2012 Cisco and/or its affiliates. All rights reserved.

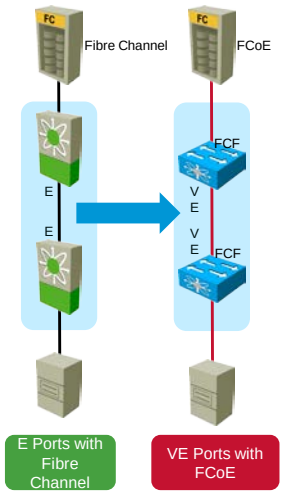
Now that the device has a complete FPMA (FC-MAP + FCID), it can communicate on the fabric using FCoE frames.

https://t.me/learningnets

certcollection.net

Fibre Channel Forwarder

FCF and Virtual Expansion Port (VE Port)



- FCFs
  - Allows switching of FCoE frames across multiple hops
- Creates standards-based FCoE ISL
- Necessary for multihop FCoE
- Nothing further required
- (No TRILL, vPC or spanning tree)
- FCF supports all Fibre Channel functionalities:
  - Supports up to 7 hops
  - Supports up to 10,000 logins per fabric
  - Supports up to 8000 zones per switch
  - Supports up to 500 zone sets per switch

It is Fibre Channel.

Same Fibre Channel CLI available on the Ethernet switch.

© 2012 Cisco and/or its affiliates. All rights reserved.

FIP can also create virtual expansion-to-virtual expansion links. This simple migration follows traditional Fibre Channel Inter-Switch Link (ISL) trunk design while using Ethernet as the transport layer. There is no need for any additional Ethernet protocol.

Virtual expansion-to-virtual expansion links and FCFs are necessary for multihop FCoE.

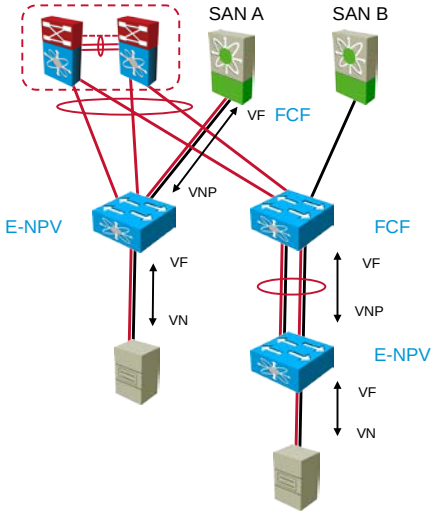
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

Ethernet NPV

- On the control plane (FIP EtherType), an Ethernet NPV bridge improves over a FIP snooping bridge by intelligently proxying FIP functions between a CNA and an FCF.
- E-NPV load-balance logins from CNAs are spread evenly across the available FCF uplink ports.
- An E-NPV bridge is VSAN-aware and capable of assigning VSANs to the CNAs.
- E-NPV will take the VSAN into account when mapping the pinned logins from a CNA to an uplink towards the FCF.



© 2012 Cisco and/or its affiliates. All rights reserved.

For Ethernet N-Port Virtualizer (E-NPV), server-side FLOGI is passed as is to the core switch instead of translating it into a fabric discover (FDISC). This is because FIP supports multiple FLOGIs on the same physical ports if they are from different ENode MAC addresses.

An important aspect of enabling E-NPV is that the FIP parameters from the FCF must be replicated on the server ports, because NPV is acting as a proxy FCF.

Spanning Tree Protocol (STP) is automatically disabled on the FCoE VLANs for virtual Fibre Channel (vFC)-bound interfaces.

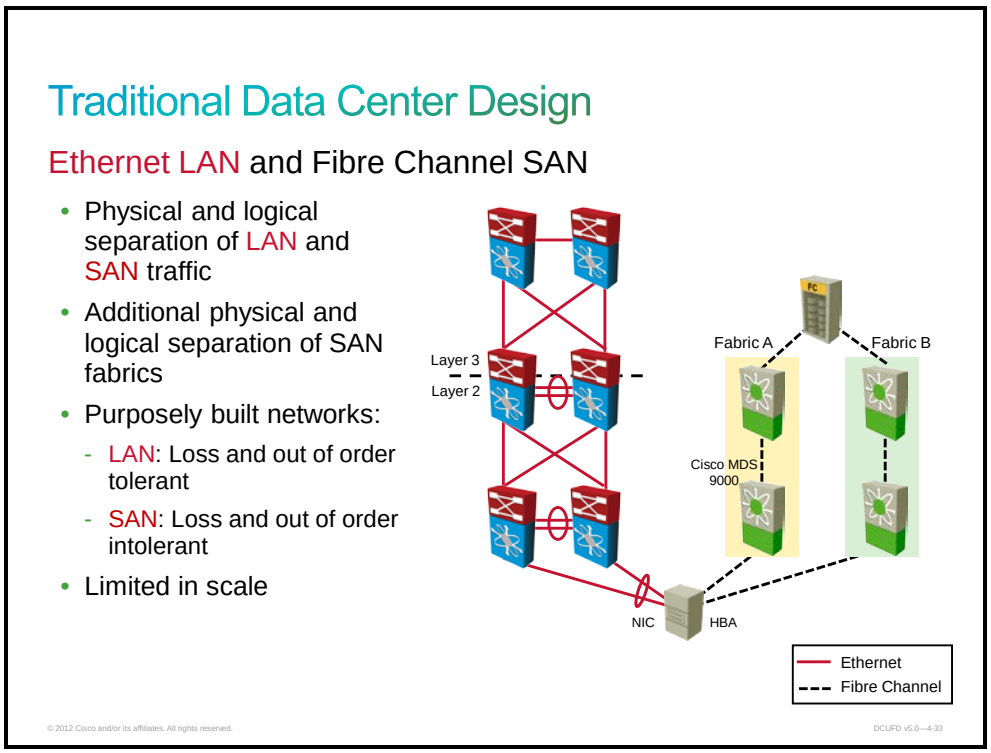
All traffic except FCoE and FIP is discarded on FCoE VLANs in order to prevent loops.

https://t.me/learningnets

certcollection.net

Unified Fabric Designs

This topic describes different design options for unified fabric networks.



Traditional data center design is a combination of LAN and SAN networks. While redundancy on a LAN is achieved by using redundant links between network devices, SAN redundancy is achieved by splitting a SAN into two individual, separate SANs.

The reason for splitting the SAN into two fabrics is because, in case the principal switch fails, the whole fabric fails. This design is the only one that provides high availability by adding an additional fabric.

 © 2012 Cisco Systems, Inc.

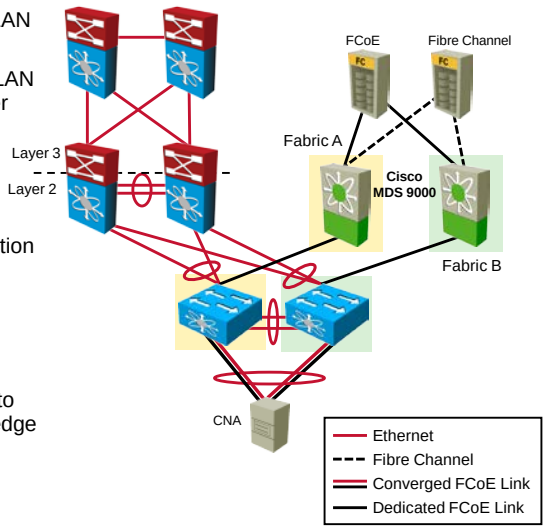
https://t.me/learningnets

certcollection.net

Converged Access

Sharing access layer for LAN and SAN

- Shared physical, separate logical LAN and SAN traffic at access layer
- Physical and logical separation of LAN and SAN traffic at aggregation layer
- Additional physical and logical separation of SAN fabrics
- Storage virtual device context (Cisco Nexus 7000 only) for additional management and operation separation
- Higher I/O, high availability, fast reconvergence for host LAN traffic
- Edge-core topology
- Use where core switch is required to provide storage services to many edge devices



© 2012 Cisco and/or its affiliates. All rights reserved.

This is the best model for interoperability with an existing Fibre Channel SAN with convergence at the access layer. This design enables better utilization from Fibre Channel storage arrays.

Usage of FCoE on the access layer reduces first-hop cabling and simplifies data rack deployment.

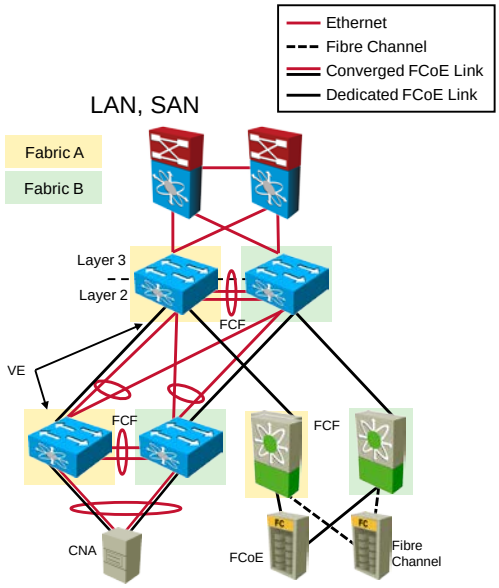
This design allows connectivity to an existing Fibre Channel SAN infrastructure without redesigning the SAN.

https://t.me/learningnets

certcollection.net

Converged Network Fabrics with Dedicated Links

- LAN and SAN traffic share physical switches, and traffic uses dedicated links between switches
- All access and aggregation switches are FCoE FCF switches
- Storage virtual device context (Cisco Nexus 7000 only) for additional operation separation at high-function aggregation and core
- Improved high availability, load sharing, and scale for LAN vs. traditional STP topologies
- SAN can utilize higher performance, higher density, lower cost Ethernet switches for the aggregation and core
- Edge-core-edge topology connectivity to existing SAN
- Use where future growth has number of storage devices exceeding ports in the core



© 2012 Cisco and/or its affiliates. All rights reserved.

In this figure, the network is converged up to the aggregation layer while still providing SAN separation.

On the LAN segment, a virtual port channel (vPC) can be used to provide higher availability and better load balancing compared to traditional STP.

This design allows connectivity to an existing Fibre Channel SAN infrastructure without redesigning the SAN.

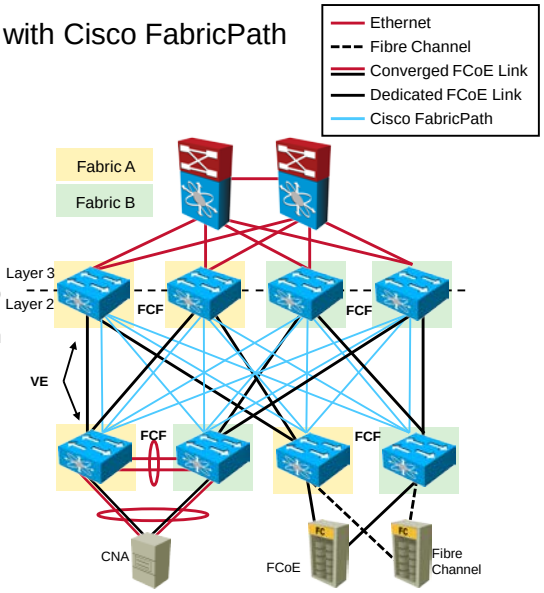
© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Converged Network with Dedicated Links

Maintaining dual SAN fabrics with Cisco FabricPath

- Cisco FabricPath enabled for LAN traffic
- Dual switch core for SAN A and SAN B
- All access and aggregation switches are FCoE FCF switches
- Dedicated links between switches are VE Ports
- Storage virtual device context (Cisco Nexus 7000 only) for additional operation separation at high-function aggregation and core
  - Improved high availability and scale over vPC (IS-IS, RPF, and N+1 redundancy)
  - SAN can utilize higher performance, higher density, lower cost Ethernet switches
  - Fibre Channel connectivity only available on Cisco Nexus 5000



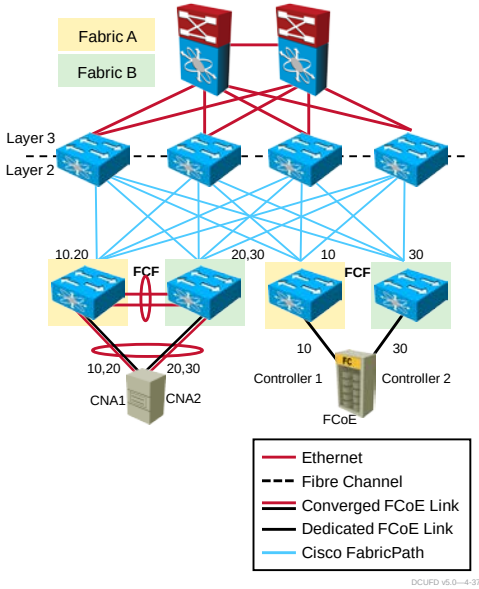
© 2012 Cisco and/or its affiliates. All rights reserved.

On the LAN segment, Cisco FabricPath is used for improved high-availability and scalability. Clear separation of Fabric A and Fabric B is maintained by using dedicated links and FCFs.

Converged Network—Single Fabric

View in the future

- LAN and SAN traffic share physical switches and links
- Cisco FabricPath enabled
- All access switches are FCoE FCF switches
- VE Ports to each neighbor access switch
- Single process and database (Cisco FabricPath) for forwarding
- Improved (N + 1) redundancy for LAN and SAN
- Sharing links increases fabric flexibility and scalability
- Distinct SAN A and SAN B for zoning isolation and multipathing redundancy
- Fibre Channel connectivity currently only available on Cisco Nexus 5000 and 5500



© 2012 Cisco and/or its affiliates. All rights reserved.

This figure represents a future goal of Cisco Data Center design. Future Cisco FabricPath releases will be capable of carrying FCoE traffic while using a single forwarding engine.

https://t.me/learningnets

certcollection.net

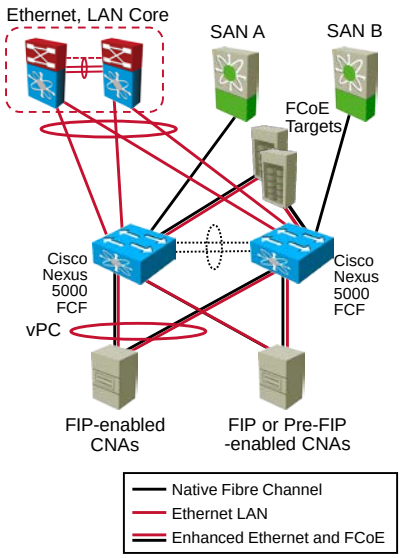
Unified Fabric Designs with FEXs

This topic describes how to design unified fabric deployments with fabric extenders (FEXs).

Single-Hop FCoE Deployments

Direct-Attached Topology

- Servers and FCoE targets are directly connected to the Cisco Nexus 5000 over 10-Gb FCoE
  - Cisco Nexus 5000 operates as the FCF
- Support for up to 52 10-Gb FCoE-attached hosts or FCoE targets per Cisco Nexus 5000
- Native Ethernet LAN network and native Fibre Channel network break off at the Cisco Nexus 5000 access layer



© 2012 Cisco and/or its affiliates. All rights reserved.

This figure represents a classic, traditional single-hop topology with directly attached hosts using FCoE on the access layer, and Cisco Nexus 5000 switches separating the LAN and SAN traffic. One of the Cisco Nexus 5000 switches belongs to SAN A and another belongs to SAN B, providing for dual-fabric separation.

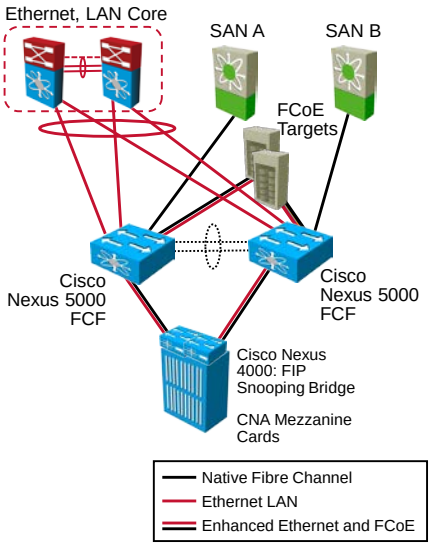
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

Multitier FCoE Deployments

Example:

- Blade servers connect to Nexus 4000 over 10-Gb FCoE
  - Cisco Nexus 4000 is a FIP snooping bridge
- Cisco Nexus 4000 connects to Cisco Nexus 5000 over 10-Gb FCoE
  - Cisco Nexus 5000 operates as the FCF
- Support for up to 192 FCoE-attached blade servers per Cisco Nexus 5000
- Native Ethernet LAN network and native Fibre Channel network break off at the Cisco Nexus 5000



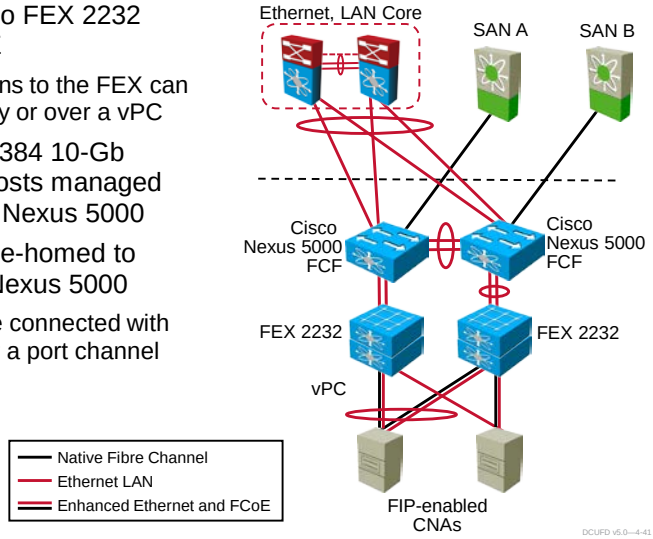
© 2012 Cisco and/or its affiliates. All rights reserved.

In multihop FCoE, the Cisco Nexus 4000 is a FIP snooping bridge that is aware of FCoE traffic. The Cisco Nexus 5000 or 5500 Series switch is necessary for providing the first Fibre Channel hop (the FCF), to where all the hosts register.

Extending the FCoE Access Layer

Supported Directly Attached Topologies

- Servers connect to FEX 2232 over 10-Gb FCoE
  - Server connections to the FEX can be Active/Standby or over a vPC
- Support for up to 384 10-Gb FCoE-attached hosts managed by a single Cisco Nexus 5000
- FEX 2232 is single-homed to upstream Cisco Nexus 5000
  - FEX 2232 can be connected with individual links or a port channel



© 2012 Cisco and/or its affiliates. All rights reserved.

A directly connected topology can easily be extended using a FEX, enabling higher port density per Cisco Nexus 5000 or 5500 switch and maintaining top-of-rack (ToR) or end-of-row (EoR) cabling topology.

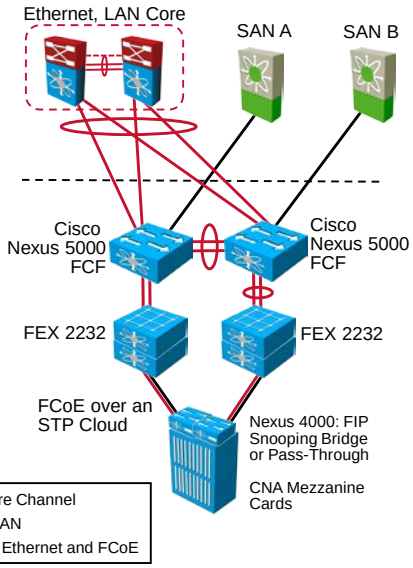
https://t.me/learningnets

certcollection.net

Extending the FCoE Access Layer (Cont.)

Multihop Topologies

- Server connection to the Cisco Nexus 4000 is Active/Standby
- Servers connect to Cisco Nexus 4000 over 10-Gb FCoE
- PFC support at every hop to achieve flow control
- Support for up to 640 10-Gb FCoE-attached hosts managed by a single Cisco Nexus 5000
- FEX 2232 is single-homed to upstream Cisco Nexus 5000 with single links or a port channel



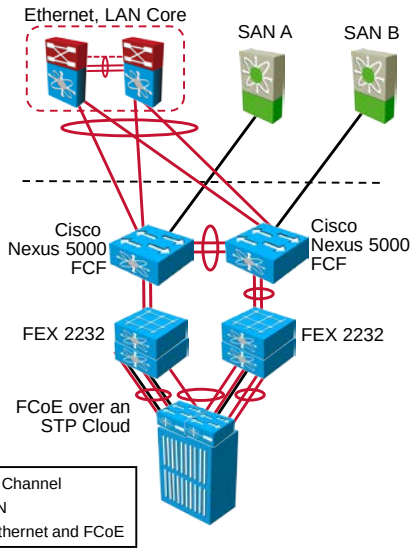
© 2012 Cisco and/or its affiliates. All rights reserved.

In this topology, the FEX can be connected to the Cisco Nexus 5000 switch either by using static or dynamic pinning. By using the FEX, a single Cisco Nexus 5000 switch can be expanded to 640 attached hosts without adding any new management points.

Extending the FCoE Access Layer (Cont.)

Multihop Topologies (Cont.)

- Server connection to the Cisco Nexus 4000 is Active/Standby
- Servers connect to Cisco Nexus 4000 over 10-Gb FCoE
  - Separate 10-Gb links for native Ethernet and FCoE
  - Native Ethernet links can connect to the FEX using single link, port channel, or virtual port-channel
  - FCoE links can be connected to FEX using single link or port channel
- Support for up to 640 10-Gb FCoE-attached hosts managed by a single Cisco Nexus 5000



© 2012 Cisco and/or its affiliates. All rights reserved.

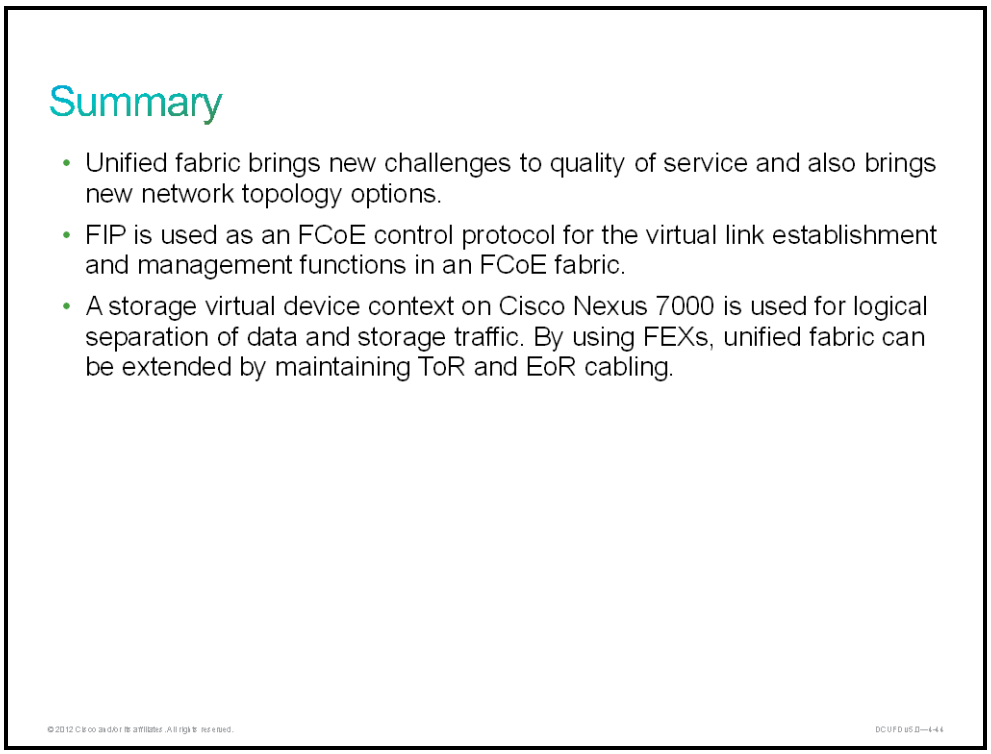
This figure shows the usage of FEX 2232 in a multihop topology with a blade server system attached to it. The FEX provides a point of physical attachment and is not involved in the connection between the Cisco Nexus 4000 as the FIP snooping bridge and the Cisco Nexus 5000 or 5500 as the FCF.

https://t.me/learningnets

certcollection.net

Summary

This topic summarizes the primary points that were discussed in this lesson.



References

For additional information, refer to these resources:

- Cisco FCoE page: http://www.cisco.com/go/fcoe
- Cisco Nexus portfolio page: http://www.cisco.com/go/nexus

https://t.me/learningnets

certcollection.net

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

Lesson 4

Designing SAN Services

Overview

This lesson explains the SAN-based services that are used during SAN design, when various other services cannot be performed by hosts or storage.

Objectives

Upon completing this lesson, you will be able to design SAN-based Fibre Channel services. This ability includes being able to meet these objectives:

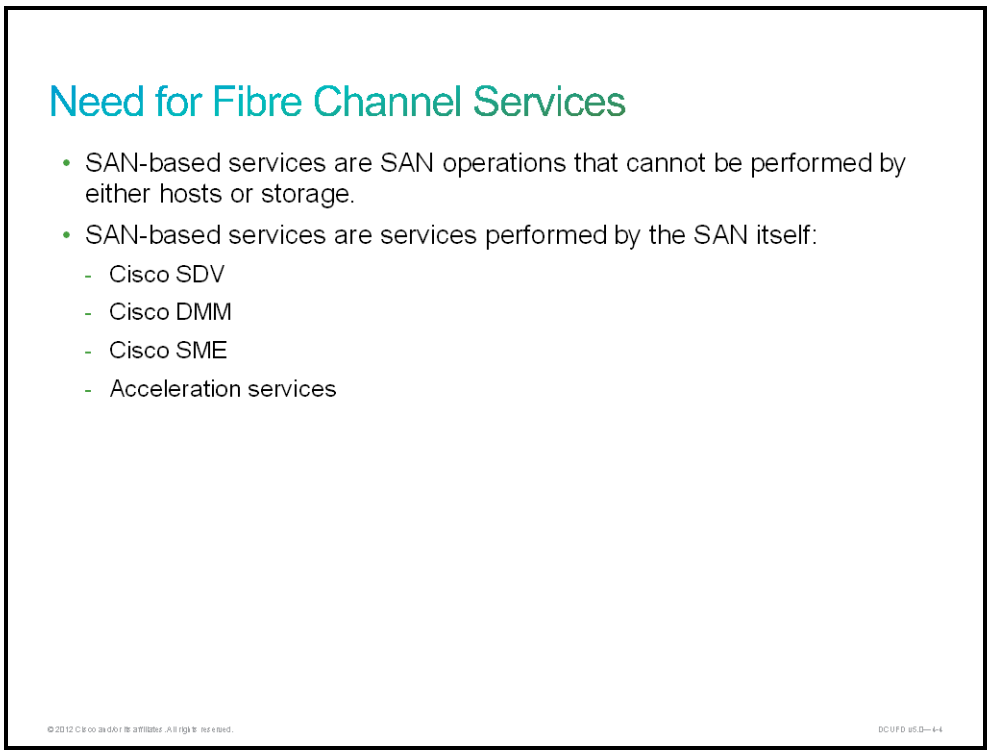
- Identify the need for SAN-based Fibre Channel services
- Design SAN-based Fibre Channel services
- Explain SAN replication
- Design long-distance Fibre Channel interconnects
- Present design examples and use cases for various SAN long-distance acceleration solutions

https://t.me/learningnets

certcollection.net

SAN-Based Services

This topic describes how to identify the need for SAN-based Fibre Channel services.



SAN-based services are performed by the SAN itself. Hosts and storage are not included or aware of the actions that are performed by SAN services.

 © 2012 Cisco Systems, Inc.

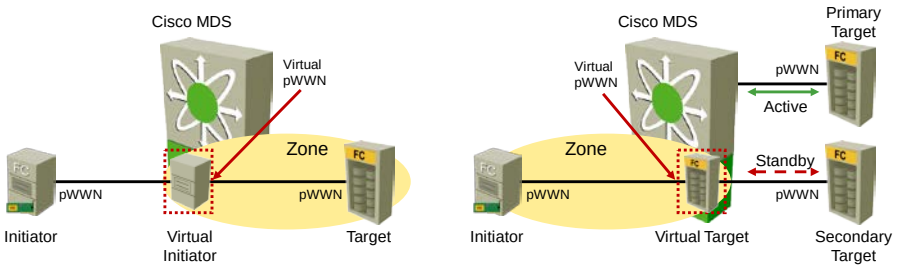
https://t.me/learningnets

certcollection.net

Cisco SDV

Use Cisco SDV to create virtual devices that represent physical end devices:

- Accelerates swapout of failed device interfaces
- Only one entry to change when end device is replaced
- No need to reconfigure zoning or security
- Can create virtual initiators or virtual targets
- Requires the Enterprise license



© 2012 Cisco and/or its affiliates. All rights reserved.

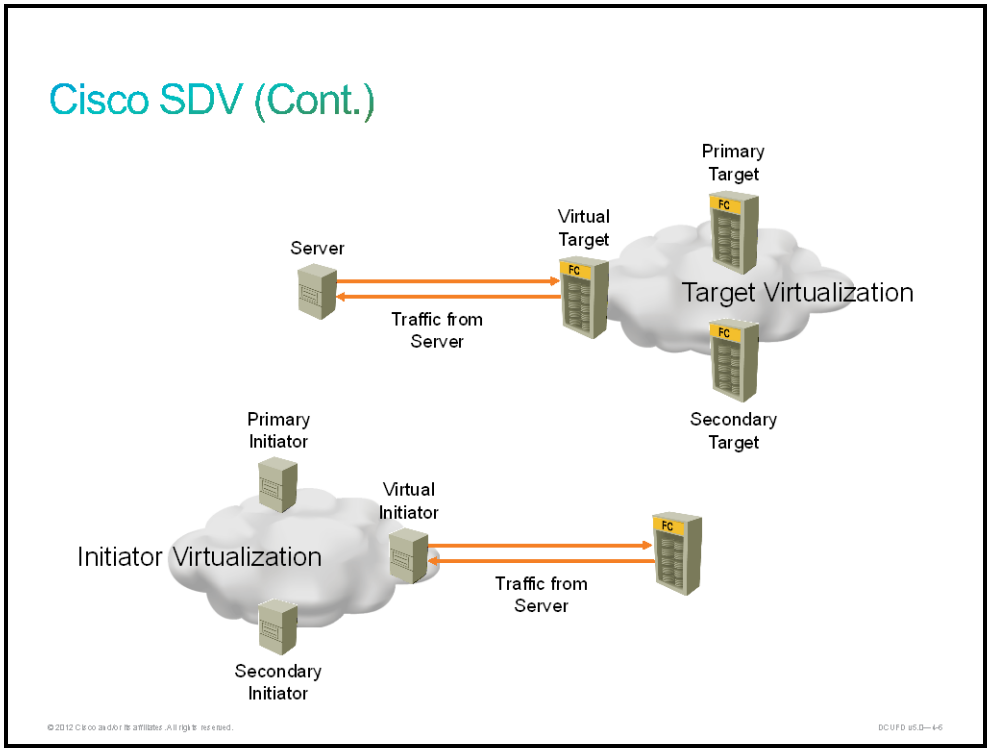
DCUFD v5.0—4-5

Cisco SAN device virtualization (SDV) provides a virtual port world wide name (pWWN) that represents a physical pWWN on a connected device.

SAN devices that are virtualized can be either initiators or targets. You can virtualize targets to create a virtual target, and you can also virtualize initiators to create a virtual initiator.

https://t.me/learningnets

certcollection.net



Virtualization of SAN devices accelerates swapout or failover to a replacement disk array and minimizes downtime when replacing host bus adapters (HBAs) or when rehosting an application on a different server. The Cisco SDV feature allows you to create virtual devices to represent physical end devices. Cisco SDV has been available since Cisco MDS SAN-OS Release 3.1(2) and Cisco Nexus Operating System (NX-OS) Release 4.1(1a).

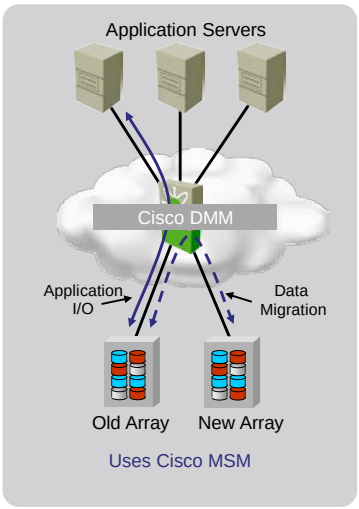
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

Cisco DMM

- Migrates data between storage arrays for the following:
  - Technology refreshes
  - Workload balancing
  - Storage consolidation
- Cisco DMM offerings:
  - Online migration of heterogeneous arrays
  - Simultaneous migration of multiple LUNs
  - Unequal size LUN migration
  - Rate adjusted migration
  - Verification of migrated data
  - Secure erase
  - Dual-fabric support
  - CLI and wizard-based management with Cisco Fabric Manager
- Requires no SAN reconfiguration or rewiring



© 2012 Cisco and/or its affiliates. All rights reserved.

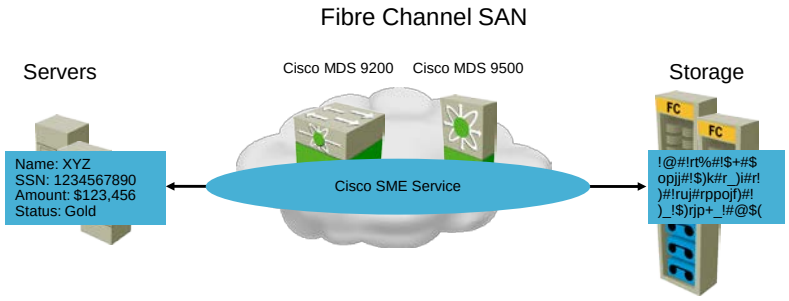
While it is designed to support various SAN topologies, the Cisco Data Mobility Manager (DMM) feature is also influenced by the topologies. Similarly, the location of the Cisco Storage Services Module (SSM) or Cisco Multiprotocol Services Module (MSM) is also affected by the SAN topology. Cisco DMM supports homogeneous and heterogeneous SANs, as well as single-fabric and dual-fabric SAN topologies. Dual-fabric and single-fabric topologies both support single-path and multipath configurations. In a single-path configuration, a migration job includes only the one path, represented as an initiator and target port pair. In a multipath configuration, a migration job must include all paths, represented as two initiator and target port pairs.

https://t.me/learningnets

certcollection.net

Cisco SME

- Enables Cisco SME service globally within the SAN; no additional appliances or cabling
- Encryption for specific disk and tape storage devices
- Supported by Cisco MDS 9222i, MPS 18/4, or SSN-16 modules



© 2012 Cisco and/or its affiliates. All rights reserved.

The Cisco Storage Media Encryption (SME) solution is a comprehensive network-integrated encryption service with enterprise-class management that works transparently with existing and new SANs.

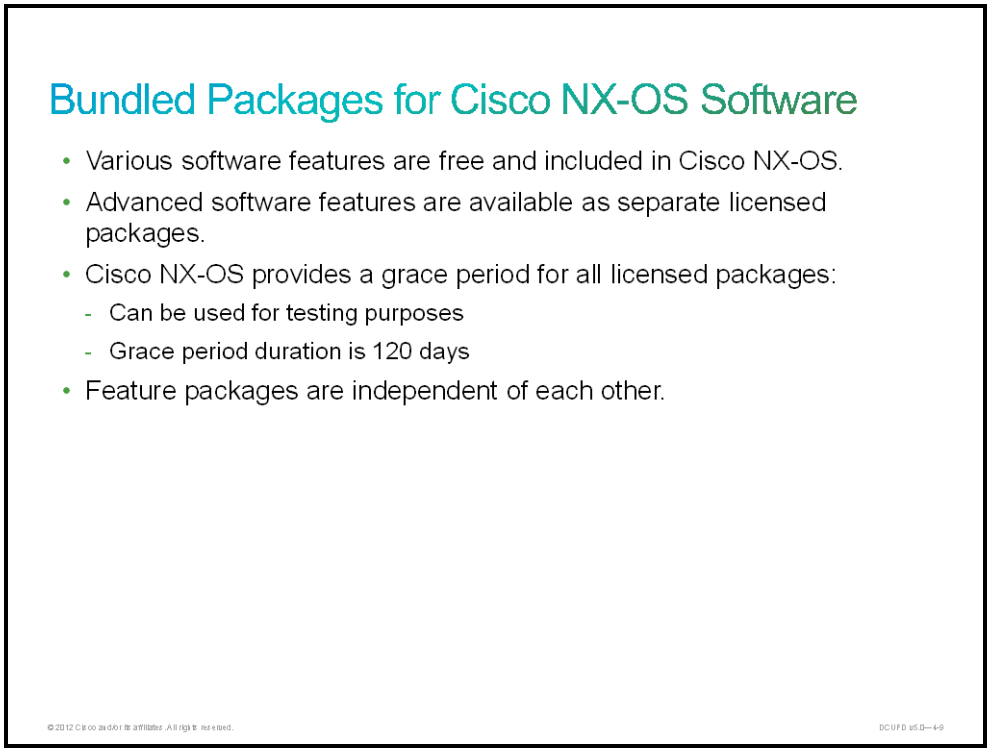
- Cisco SME installation and provisioning are simple and nondisruptive.
- Encryption engines are integrated on the Cisco MDS 9000 18/4-Port Multiservice Module (MSM-18/4), the Cisco MDS 9222i Multiservice Modular Switch, and the Cisco MDS 9000 16-Port Gigabit Ethernet Storage Services Node (SSN-16).
- Traffic from any virtual storage area network (VSAN) can be encrypted using Cisco SME, enabling flexible, automated load balancing through network traffic management across multiple SANs.
- Cisco SME is integrated into Cisco Fabric Manager and requires no additional software.

Cisco SME is a standards-based encryption solution for heterogeneous and virtual tape libraries. Cisco SME is managed with Cisco Fabric Manager and the CLI.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net



The Cisco NX-OS is a plentiful operating system. Numerous features are included in the base image. Advanced software features are available as separate licensed packages. These are the available packages:

- Enterprise Package
- SAN Extension over IP Package for IPS-8
- SAN Extension over IP Package for MPS-14/2 and MSP-18/4 modules
- Mainframe Package
- Storage Service Enabler Package
- Fabric Manager Service Package
- On-demand Port Activation Licensing Package
- 10-Gb/s Port Activation Package
- Storage Media Encryption (SME)
- Data Mobility Manager (DMM)
- Cisco I/O Acceleration (IOA)
- Extended Remote Copy (XCR) Acceleration

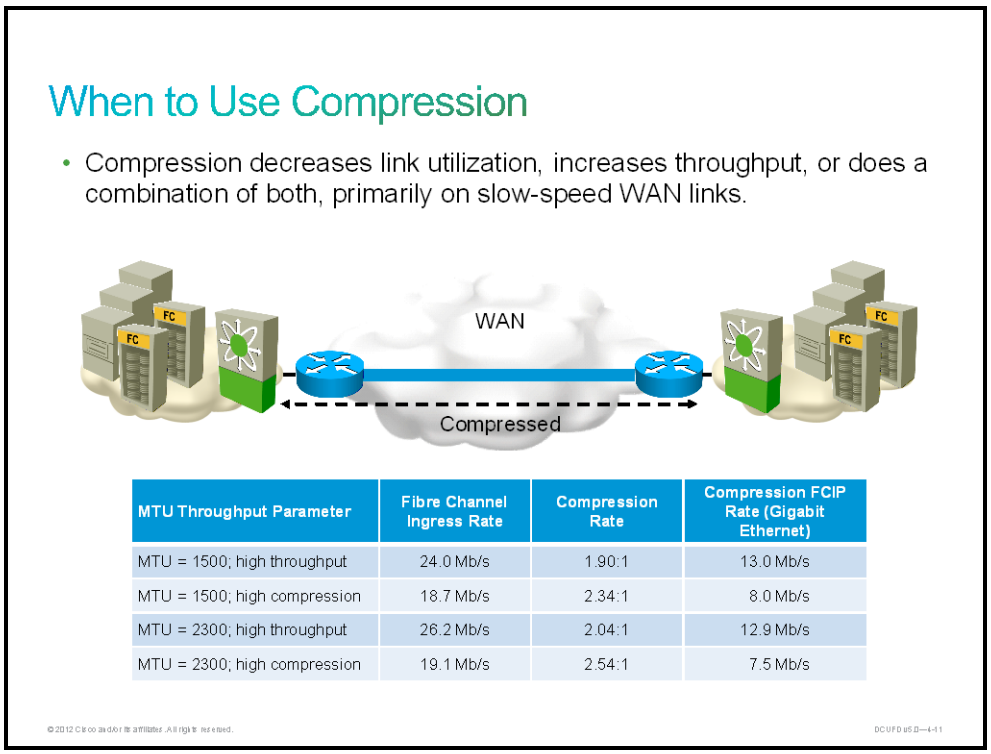
For the latest list of Cisco NX-OS packages, refer to *Licensing Cisco MDS 9000 Family NX-OS Software Features*.

https://t.me/learningnets

certcollection.net

SAN-Based Services Design Considerations

This topic describes how to design SAN-based Fibre Channel services.



MTU Throughput Parameter	Fibre Channel Ingress Rate	Compression Rate	Compression FCIP Rate (Gigabit Ethernet)
MTU = 1500; high throughput	24.0 Mb/s	1.90:1	13.0 Mb/s
MTU = 1500; high compression	18.7 Mb/s	2.34:1	8.0 Mb/s
MTU = 2300; high throughput	26.2 Mb/s	2.04:1	12.9 Mb/s
MTU = 2300; high compression	19.1 Mb/s	2.54:1	7.5 Mb/s

This figure shows the effective compression ratios and resulting throughput when using different maximum transmission unit (MTU) values.

An encrypted data stream is not compressible, because it results in a bit stream that appears random. If encryption and compression are required together, it is important to compress the data before encrypting it. The receiver should first decrypt the data and then uncompress it.

You can configure Fibre Channel over IP (FCIP) compression using one of the following modes:

- **Mode 1:** A fast compression mode for high-bandwidth links (more than 25 Mb/s).
- **Mode 2:** A moderate compression mode for moderately low-bandwidth links (between 10 and 25 Mb/s).
- **Mode 3:** A high compression mode for low-bandwidth links (less than 10 Mb/s).
- **Auto:** Picks the appropriate compression scheme, based on the bandwidth of the link that is configured in the TCP parameters of the FCIP profile. This is the default mode.

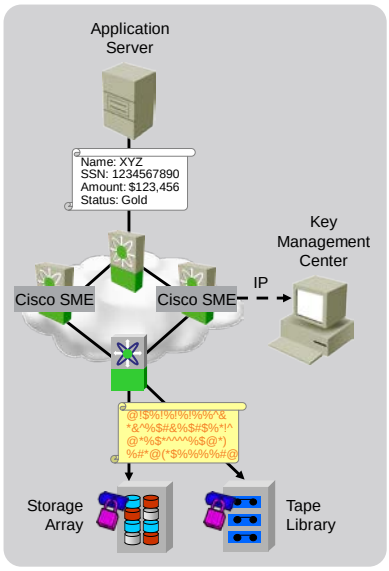
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

Cisco SME

- Encrypts storage media (data at rest)
  - IEEE-compliant AES-256 encryption
  - Integrated as transparent fabric service
- Transparent fabric service
- Supports heterogeneous storage arrays, tape devices, and virtual tape libraries
- Compresses tape data
- Offers secure, comprehensive key management
- Allows offline media recovery
- Built upon the Federal Information Processing Standard (FIPS) Level 3 system architecture

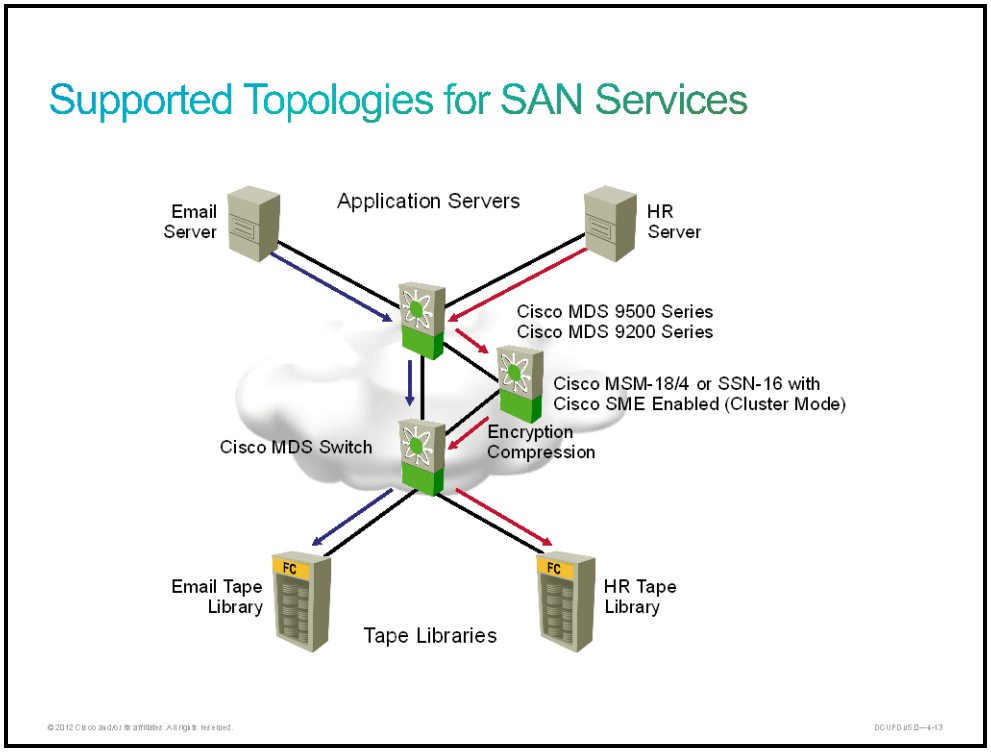


© 2012 Cisco and/or its affiliates. All rights reserved.

Cisco SME supports a single-fabric topology where the Cisco MSM-18/4, the Cisco MDS 9222i switch, and the Cisco SSN-16 provide the storage media encryption engines that are used by Cisco SME to encrypt and compress the data at rest. To easily scale-up performance, to simplify load balancing, and to increase availability, multiple modules can be deployed in the Fibre Channel fabric. In a typical configuration, one Cisco MSM-18/4 is required in each Cisco SME cluster.

https://t.me/learningnets

certcollection.net



In this figure, the data from the human resources (HR) server is forwarded to the Cisco MSM-18/4 or SSN-16. The Cisco MSM-18/4 or SSN-16 can be anywhere in the fabric. Cisco SME performs one-to-one mapping of the information from the host to the target and forwards the encrypted data to the dedicated HR tape. Cisco SME tracks the bar codes on each encrypted tape and associates the bar codes with the host servers.

The encrypted data from the HR server is compressed and stored in the HR tape library. Data from the email server is not encrypted when it is backed up to the dedicated email tape library.

The encryption and compression services are transparent to the hosts and storage device. These services are available for devices in any VSAN in a physical fabric and can be used without rezoning.

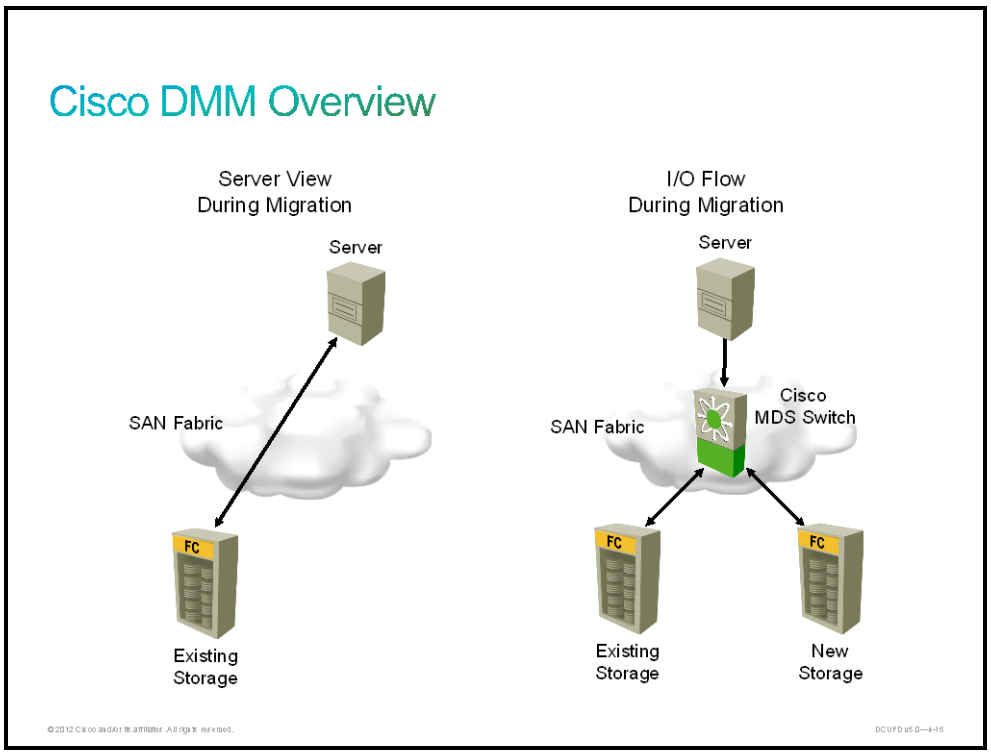
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

SAN-Based Data Replication

This topic explains SAN replication.



Traditional data migration methods can be complex and disruptive, often requiring extensive rewiring and reconfiguration of the SAN infrastructure. Configuration changes to servers and storage subsystems require coordination among different IT groups and storage vendor service representatives. Server downtime requires advanced scheduling with potentially long lead times.

Cisco MDS DMM is an intelligent software application that runs on the Cisco SSM of a Cisco MDS 9000 Series switch, the Cisco MDS 9222i Multiservice Module (MDS 9222i), or the Cisco MDS 9000 18/4-Port Multiservice Module (MSM-18/4). With Cisco MDS DMM, no rewiring or reconfiguration is required for the server, the existing storage, or the SAN fabric. Data migrations can be enabled and disabled by software control from the Cisco Fabric Manager.

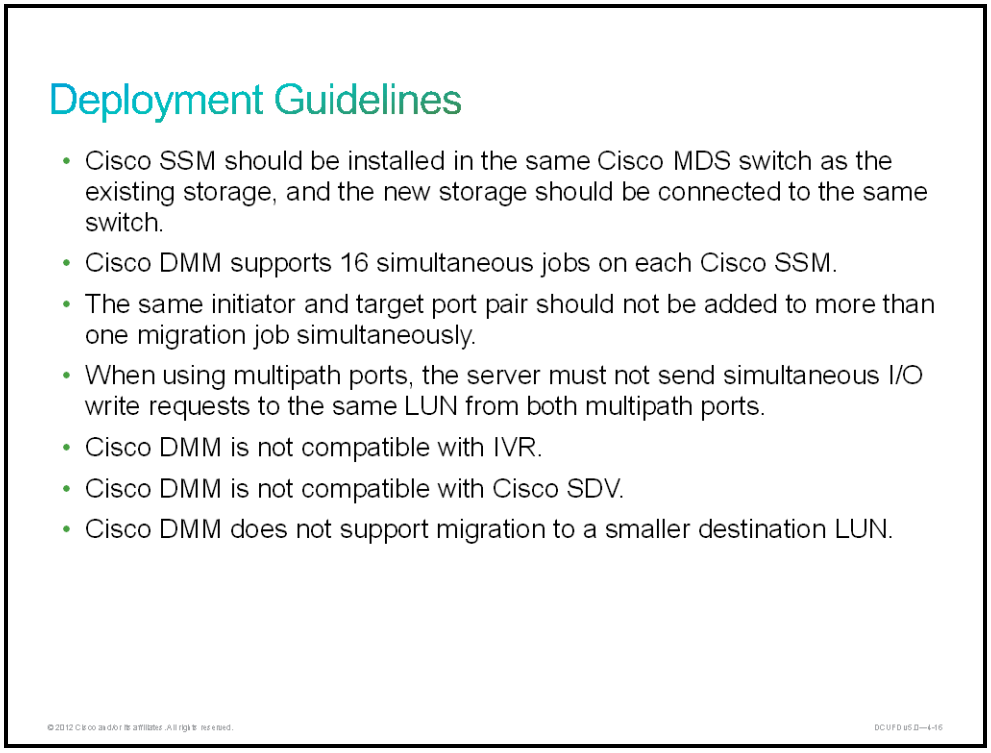
Cisco MDS DMM provides a GUI that is integrated into Cisco Fabric Manager for configuring and executing data migrations. It also provides a CLI that is suitable for creating scripts.

Application downtime is a critical factor in data migration, because prolonged periods of downtime are difficult to schedule. Cisco MDS DMM minimizes application downtime by making the existing data available to the applications while the migration is performed. Cisco MDS DMM uses hardware and software resources on the Cisco SSM, the Cisco MDS 9222i, or the Cisco MSM-18/4 to move data to the new storage. This approach ensures that data migration adds no processing overhead to the servers.

Cisco MDS DMM can be enabled when data needs to be migrated and then disabled after the migration is complete.

https://t.me/learningnets

certcollection.net



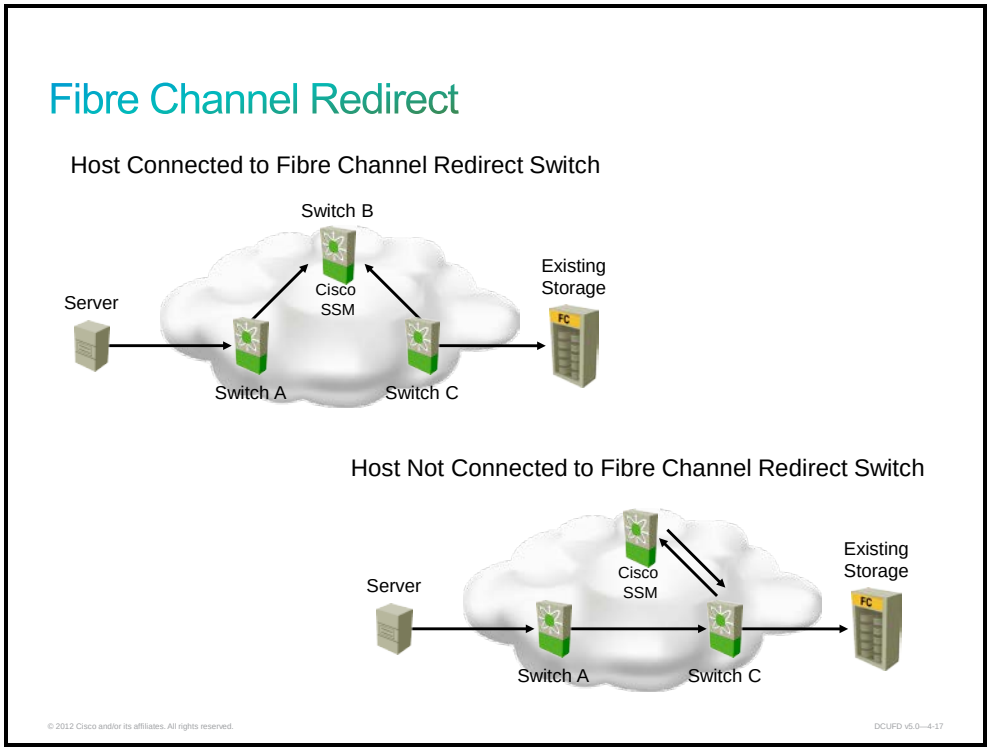
The following deployment guidelines should be considered when planning and configuring data migration using Cisco MDS DMM:

- Cisco SSM should be installed in the same Cisco MDS 9000 Series switch as the existing storage, and the new storage should be connected to the same switch. Data migration causes increased Inter-Switch Link (ISL) traffic if the existing storage or new storage devices are connected to different switches than Cisco SSM.
- Cisco MDS DMM supports 16 simultaneous jobs on Cisco SSM.
- The same initiator and target port pair should not be added to more than one migration job simultaneously.
- When using multipath ports, you must ensure that the server does not send simultaneous I/O write requests to the same logical unit number (LUN) from both multipath ports. The first I/O request must be acknowledged as completed before initiating the second I/O request.
- Cisco DMM is not compatible with LUN zoning.
- Cisco DMM is not compatible with Inter-VSAN Routing (IVR). The server and storage ports must be included in the same virtual storage area network (VSAN).
- Cisco DMM is not compatible with Cisco SDV. The server and storage ports cannot be virtual devices, or physical devices that are associated with a virtual device.
- Cisco DMM does not support migration to a smaller destination LUN.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

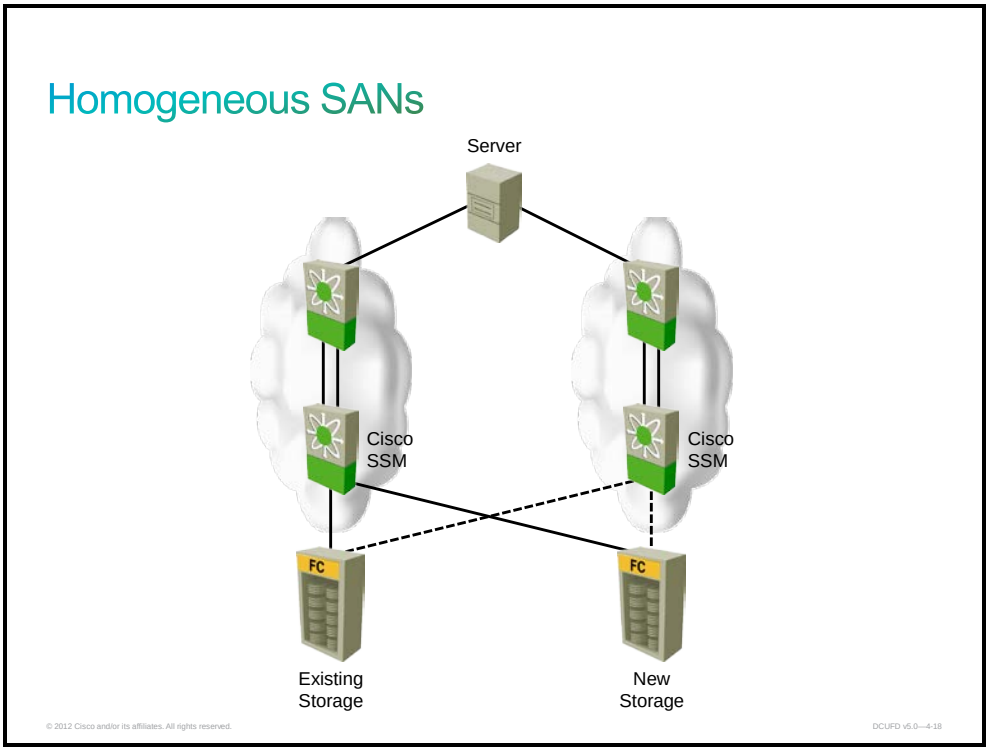


In the first diagram of this figure, the server HBA port is connected to Switch A, and the existing storage is connected to Switch C. Both switches have Fibre Channel Redirect Capability (RDC). The Cisco SSM or MSM is installed in Switch B. When the data migration job is started, Fibre Channel RDC is configured on Switch A to divert the server traffic to Cisco SSM or MSM. Fibre Channel RDC is configured on Switch C to redirect the storage traffic to Cisco SSM or MSM.

In the second diagram of this figure, the server HBA port is connected to Switch A, which either does not have Fibre Channel RDC or is not running Cisco MDS SAN-OS Software Release 3.2(1) or Cisco NX-OS Software Release 4.1(1b) or later. The existing storage is connected to Switch C, which has Fibre Channel RDC. Cisco SSM or MSM is installed on Switch B. Switches B and C are running Cisco MDS SAN-OS Software Release 3.2(1) or Cisco NX-OS Software Release 4.1(1b) or later. When the data migration job is started, Fibre Channel RDC is configured on Switch C to redirect the server and storage traffic to Cisco SSM or MSM. This configuration introduces additional network latency and consumes additional bandwidth, because traffic from the server travels an extra network hop (A to C, C to B, B to C). The recommended configuration of placing Cisco SSM or MSM in Switch C avoids the increase in network latency and bandwidth.

https://t.me/learningnets

certcollection.net



A homogeneous SAN contains only Cisco MDS 9000 Series switches. Most topologies fit in one of the following categories:

- **Core-edge:** Hosts at the edge of the network, and storage at the core.
- **Edge-core:** Hosts and storage at the edge of the network, and ISLs between the core switches.
- **Edge-core-edge:** Hosts and storage are connected to opposite edges of the network and to the core switches with ISLs.

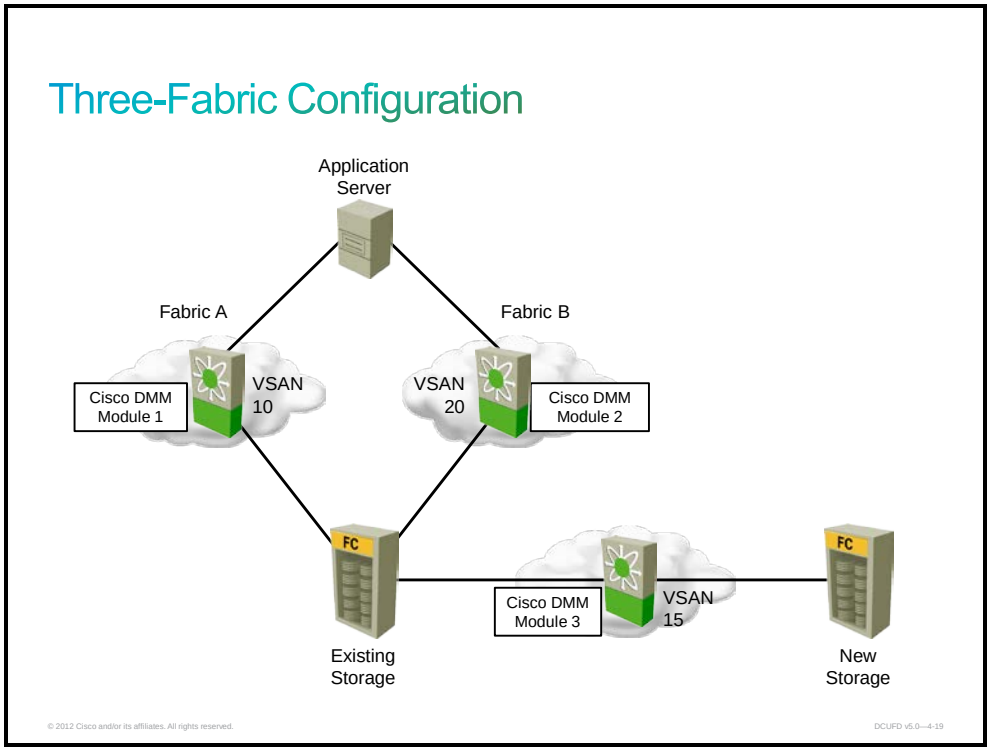
It is recommended for any of these topologies that Cisco SSM or MSM is located in the switch that is closest to the storage devices so that no additional network traffic is introduced by Cisco DMM during data migration.

In a homogeneous network, Cisco SSM or MSM can be located on any Cisco MDS 9000 Series switch that supports Cisco DMM in the fabric where the existing storage is attached. The new storage should be connected to the same switch as the existing storage.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net



This Cisco DMM Method 3 supports the dedicated migration fabric and is designed to address the problem of migrating data from an array port that is connected to a dedicated SAN that is different from the product SAN.

Many IT organizations require data migration to a remote data center. Some organizations prefer to use a dedicated storage port (on the existing storage array) that is connected to a separate physical fabric. This fabric is called the migration or replication fabric because it is used for data migration as well as continuous data replication services.

In Cisco DMM Method 3, Cisco SSM or MSM in the migration SAN is responsible for executing the sessions in the Cisco DMM job and copying the data from the existing storage to the new storage. Cisco SSM or MSM in the production SANs is responsible for tracking the server writes to the existing storage. No server writes are expected in the migration SAN.

Server writes in the production SAN are logged by Cisco SSM or MSM in that fabric by maintaining a Modified Region Log (MRL) for each LUN that is migrated. Cisco SSM or MSM in the migration SAN does not maintain any MRL for the LUN because no server writes to the existing storage LUN are expected in the migration SAN. Cisco SSM or MSM in the migration SAN is responsible for retrieving the MRLs for a LUN from both the production SANs and for performing a union of the MRLs to create a superset of all modified blocks on the LUN via paths from both production SANs. Cisco SSM or MSM then copies all the modified regions from the existing storage LUN to the new storage LUN in the migration SAN. This process is repeated until the administrator is ready to finish the Cisco DMM job and perform a cut-over. The finishing operation in this method places all LUNs in offline mode and performs a final pass over the combined MRL to synchronize the existing and new storage LUN in each session.

The three-fabric topology supports two production fabrics and one migration fabric. Each of the fabrics has one VSAN per fabric.

https://t.me/learningnets

certcollection.net

The production fabric consists of the following:

- Two fabrics—Fabric A and Fabric B
- Two VSANs in each of the fabrics—VSAN 10 in Fabric A, and VSAN 20 in Fabric B
- Two Cisco DMM instances in each of the fabrics—DMM module 1 and DMM module 2
- Ports for the application server and the existing storage
- Application server port and storage port in the same VSAN for each fabric

The VSANs in both the fabrics can have different numbers.

The migration fabric consists of the following:

- One fabric—Fabric C
- One VSAN—VSAN 15
- One Cisco DMM instance—DMM module 3
- Existing storage port and new storage port in the same VSAN

The migration fabric VSAN can have a different number from the production fabric VSAN.

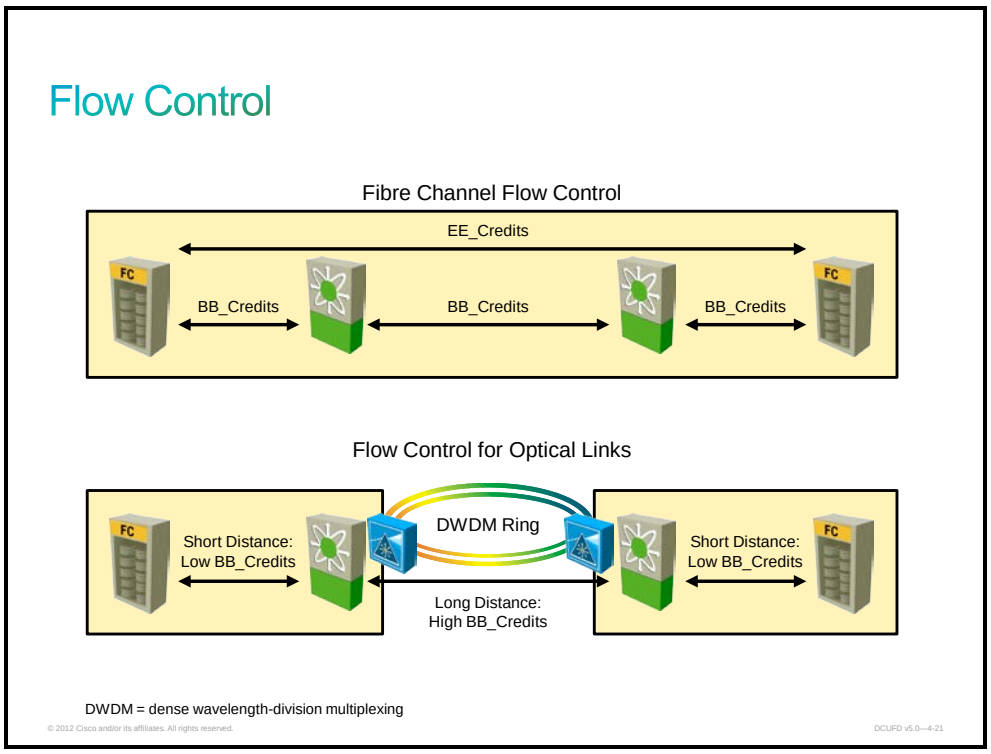
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

Long-Distance Fibre Channel Interconnects

This topic explains how to design long-distance Fibre Channel interconnects.



All data networks employ flow control to prevent data overruns in intermediate and end devices.

Fibre Channel networks use buffer-to-buffer credits (BB_Credits) on a hop-by-hop basis with Class 3 storage traffic. Senders are permitted to send up to the negotiated number of frames (equal to the BB_Credit value) to the receiver before waiting for receiver ready (R_RDY) responses to return from the receiver to replenish the BB_Credits for the sender. As distance increases, so does latency, therefore the number of BB_Credits that are required to maintain the flow of data increases.

Fibre Channel Flow Control

Fibre Channel implements two levels of flow control:

- **Port-to-port (BB_Credit):** A link-level protocol that is used in Class 2 and Class 3 services between a node and a switch.
- **Source-to-destination (end-to-end credits [EE_Credits]):** Used in Class 1 and Class 2 services between two end nodes, regardless of the number of switches in the network.

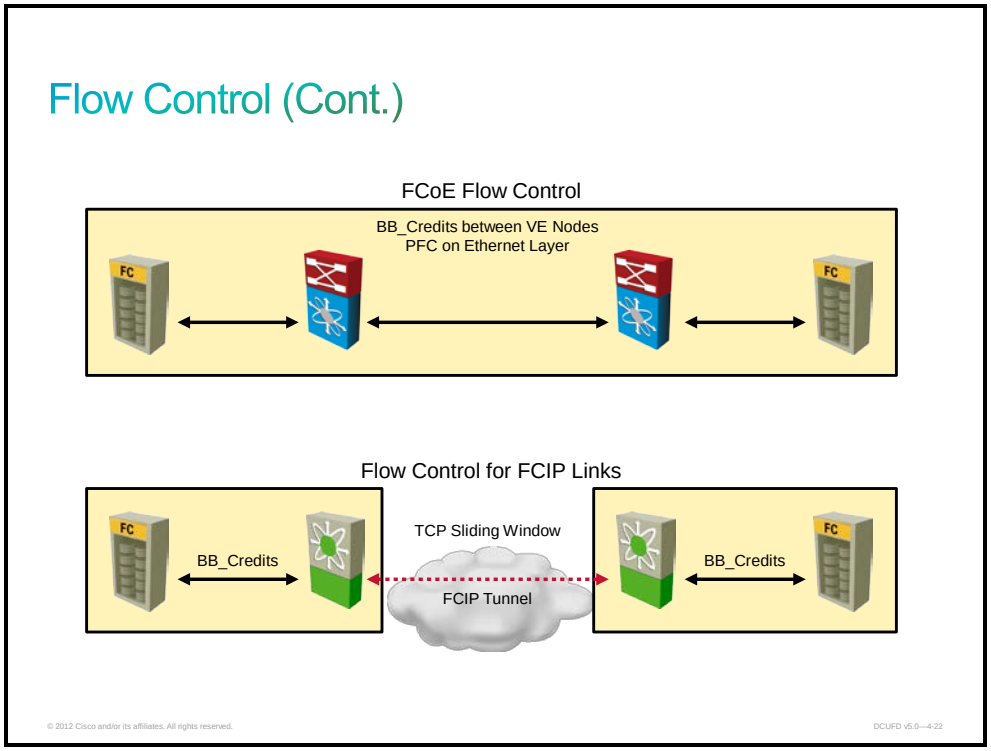
Fibre Channel Flow Control over Optical Links

Optical links, when they are available and affordable, allow for remote connectivity for distances in the hundreds of kilometers. However, longer distances increase latency (which is the time that a frame or R_RDY response can be in transit). More BB_Credits are required for longer distances. This is the formula to calculate the minimum required BB_Credits based on distance:

- 1 BB_Credit per 0.62 miles (1 km) at 2 Gb/s
- 1 BB_Credit per 1.24 miles (2 km) at 1 Gb/s

https://t.me/learningnets

certcollection.net



FCoE Flow Control

The Fibre Channel over Ethernet (FCoE) flow control protocol utilizes Ethernet as a replacement for the FC0 and FC1 layers of the Fibre Channel stack. FCoE introduces an enhancement to Ethernet with priority-based flow control (PFC). PFC is used to prevent frame loss between nodes. With PFC, a new Fibre Channel nodes type is created (called “virtual extension”). Virtual extension nodes between each other utilize BB_Credit flow control.

FCIP Flow Control

FCIP links are slightly different from Fibre Channel links. Although FCIP links carry Fibre Channel traffic, they do not use, and are therefore not constrained by, BB_Credits. Instead, TCP flow control is used. TCP is an end-to-end sliding window flow control mechanism. The ability to keep the pipeline full or maintain the data flow is governed by the window size.

In Fibre Channel networks, BB_Credits apply to every hop except FCIP hops. Fibre Channel flow control terminates at the virtual expansion port of an FCIP tunnel. To adhere to the Fibre Channel standard flow control, R_RDY responses can be spoofed on the local SAN. Cisco ONS 15454 and Cisco ONS 15530 Fibre Channel SL-Series cards (version 5.0+) spoof R_RDY responses.

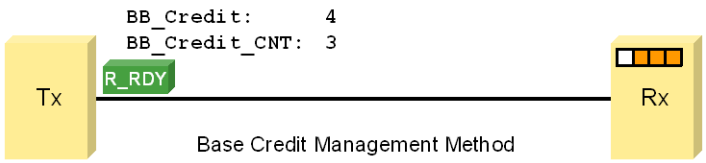
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

Base Credit Management Method

1. At login, Rx tells Tx how many buffers Rx has (BB_Credit).
2. Tx sets BB_Credit_CNT to 0 at login.
3. Tx increments BB_Credit_CNT when it sends a frame.
4. Rx sends an R_RDY notification when it processes the frame.
5. Tx decrements BB_Credit_CNT when it receives the R_RDY notification. Tx sends a frame only when BB_Credit_CNT is less than BB_Credit.



The base credit management method works as follows:

- When the transmit (Tx) port sends a port login request, the receive (Rx) port responds with an accept (ACC) frame that includes information about the size and number of frame buffers it has (BB_Credit). The Tx port stores the BB_Credit value in a table.
- The Tx port also stores another value called “BB_Credit_CNT,” which represents the number of used buffer credits. BB_Credit_CNT is set to zero after the ports complete the login process.
- Each time the Tx port sends a frame, it increments BB_Credit_CNT.
- Upon receiving the frame, the Rx port processes the frame and moves it to upper-layer protocol (ULP) buffer space. The Rx port then sends an R_RDY acknowledgment signal back to the Tx port, informing it that a buffer is available.
- When the Tx port receives the R_RDY signal, it then decrements its BB_Credit_CNT.

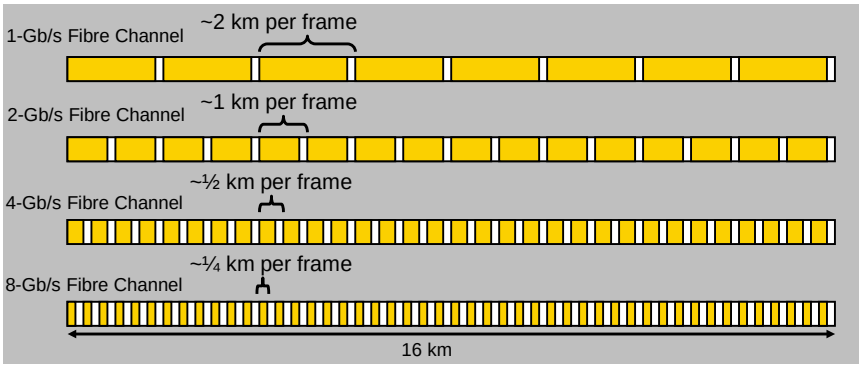
To prevent overrunning the Rx port buffers, the Tx port can never allow BB_Credit_CNT (the number of frames that have not yet been acknowledged) to exceed BB_Credit (the total number of buffers in the Rx port). In other words, if it cannot confirm that the Rx port has a free buffer, it does not send any more frames.

https://t.me/learningnets

certcollection.net

Extending Optical SAN Extension

- BB_Credits are used to ensure that enough Fibre Channel frames are in flight.
- A full (2112 byte) Fibre Channel frame is approximately 1.24 miles (2 km) long at 1 Gb/s, 0.62 miles (1 km) long at 2 Gb/s, and 0.31 miles (0.5 km) long at 4 Gb/s.
- As the distance increases, the number of available BB_Credits need to increase as well.
- Insufficient BB_Credits will throttle performance—no data will be transmitted until an R_RDY notification is returned.



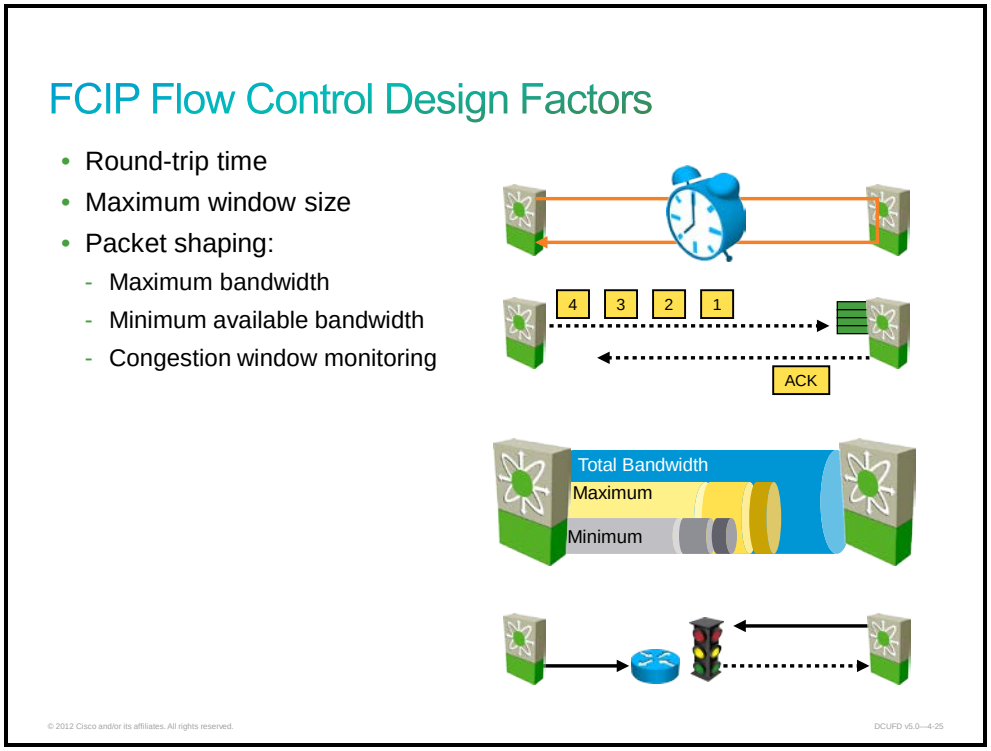
© 2012 Cisco and/or its affiliates. All rights reserved.

This figure shows the relationship between fiber length, Fibre Channel speed, and the number of BB_Credits that are needed. Failure to adjust the number of BB_Credits can lead to nonoptimal performance.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net



Controlling TCP behavior is important to optimizing your FCIP tunnel. The following design factors are outlined here:

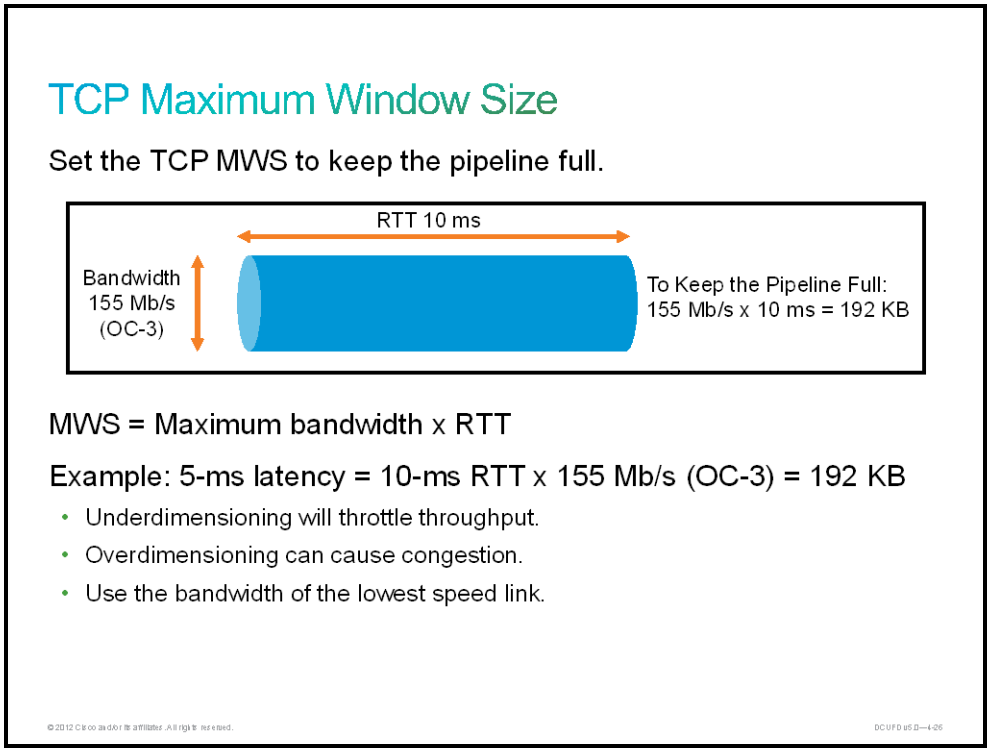
- **Round-trip time (RTT):** The estimated RTT across the IP network to reach the FCIP peer endpoint.
- **Maximum window size:** The amount of unacknowledged data in flight between the sender and the receiver.

Packet shaping includes the following parameters:

- **Maximum bandwidth:** The maximum available end-to-end bandwidth in the path. On a dedicated path, this could equal 100 percent of total bandwidth.
- **Minimum available bandwidth:** The amount of bandwidth that is guaranteed to be available or the minimum slow-start threshold. On a dedicated path, this could equal the maximum bandwidth.
- **Congestion window monitoring:** This allows TCP to monitor congestion and determine the maximum burst size that is allowed after each idle period.

https://t.me/learningnets

certcollection.net



The TCP window is the amount of unacknowledged data in flight between the sender and the receiver. In order to improve throughput, the sender transmits multiple segments without waiting for the next acknowledgment from the receiver. The TCP window is an estimate of the upper bound on the number of segments that can fit in the length of the pipeline between the sender and receiver. The window size is increased during a TCP transfer until the end-to-end path becomes too full (which is indicated by a segment being dropped somewhere in the network). Then, the window size is backed off and increased slowly again until the limit is reached.

This cycle of shrinking and slowly expanding the window size continues throughout the TCP connection. In this way, TCP tries to optimize the transmit window to maximize throughput over the lifetime of the connection. The receiver advertises its maximum window size to give the sender an idea of how much buffer space the receiver has available. This puts a firm limit on the size of the window, even if more bandwidth is available in the network.

If the pipeline is somewhat large, and the round-trip delay is long, many segments might fit in the network between the sender and receiver, and the window size needs to be somewhat large to keep the pipeline full. The formula to determine how large it should be is as follows:

window size = bandwidth * delay (round-trip-time parameter)

For example, the 155-Mb/s bandwidth with an RTT parameter of 10 ms requires a window size of approximately 192 KB.

The TCP maximum window size (MWS) for the Cisco Intrusion Prevention System (IPS) ports has the following characteristics:

- Scales up to 32 MB
- Is automatically calculated
- Varies with compression

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

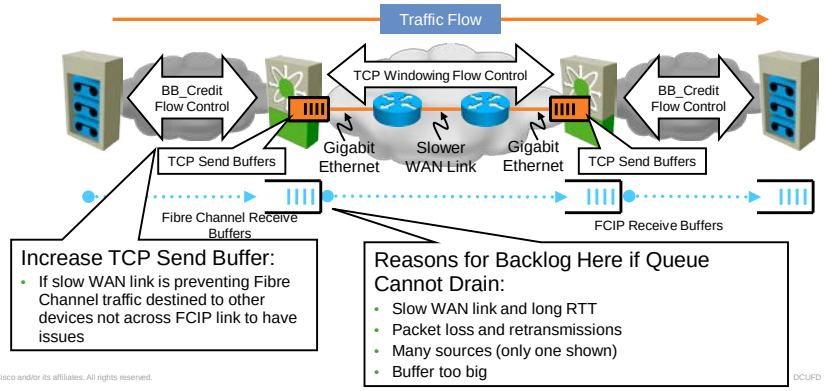
Frame Buffering: FCIP and Fibre Channel

FCIP presents a lower bandwidth pipe (if WAN link):

- Drain rate (send rate) depends upon bandwidth and congestion.

Slow ramp-up of traditional TCP can cause Fibre Channel frame expiry in some conditions:

- Mixture of slow link (such as DS3 or E3, retransmissions, many sources, big buffers)



When on a Fibre Channel network, the BB_Credit mechanism controls flow control, but flow over a WAN link uses regular TCP windowing flow control.

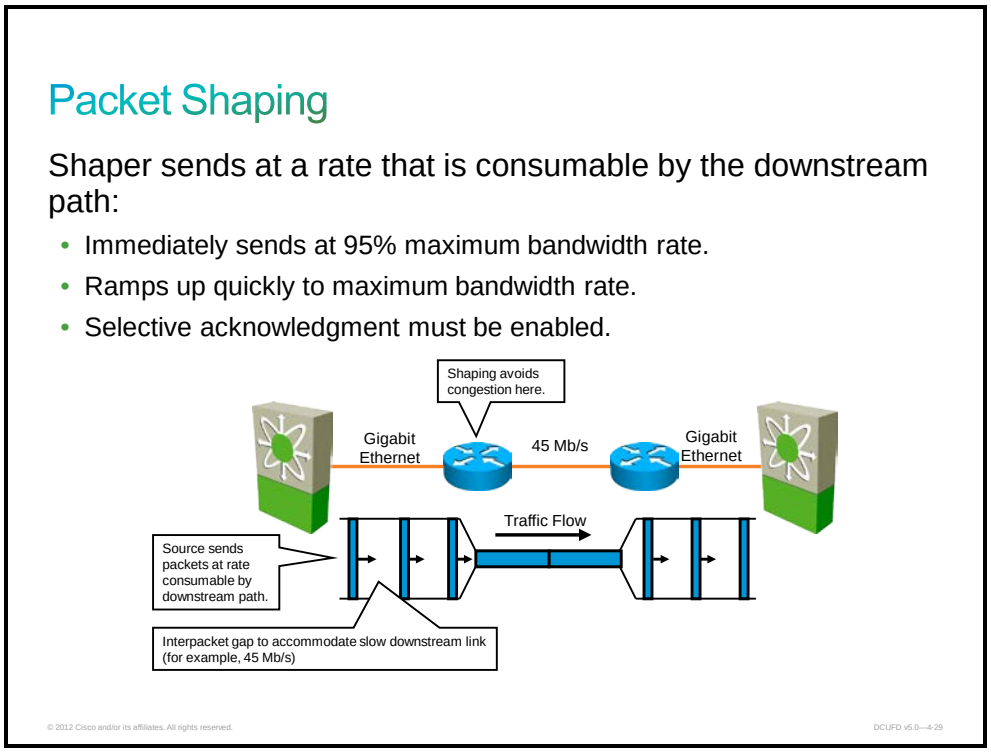
Buffer depth is controlled by Rx BB_Credit configuration on switches at Fibre Channel and FCIP boundaries. There is a frame expiration timer limit that is set to 500 ms and it is not configurable. Any frame that is waiting in the buffer for longer than 500 ms will be marked as expired and discarded. In that case, retransmission of the FCIP frame must occur.

https://t.me/learningnets

certcollection.net

Fibre Channel Long-Distance Acceleration Solutions

This topic describes design examples and use cases for various SAN long-distance acceleration solutions.



The TCP implementation on the Cisco IPS ports is slightly different from typical TCP. The TCP implementation on Cisco IPS ports employs a traffic-shaping function that sends traffic during the first round-trip period after an idle at a rate that is equivalent to the minimum available bandwidth of the path. This mechanism allows the Cisco IPS ports to ramp up more quickly and recover from retransmissions more effectively than normal TCP implementations.

Packet shaping results in sending packets at a consumable rate for downstream routers and switches, which is determined by the minimum guaranteed available bandwidth of the path. Shaping is operative only during the first RTT. After that, returning acknowledgments pace the transmission to determine the interpacket gap to accommodate slow downstream links.

For example, consider an FCIP link without shaping capability over a network where the maximum path bandwidth between the two FCIP endpoints is 45 Mb/s. If the FCIP endpoint bursts the data out of the Gigabit Ethernet interface, then the downstream router has to buffer the packets while serializing them over the slower 45-Mb/s link.

When packet shaping is correctly configured, packets are sent over the Gigabit Ethernet interface with sufficient spacing so that they can be forwarded with minimal or no buffering at each intermediate point in the FCIP path.

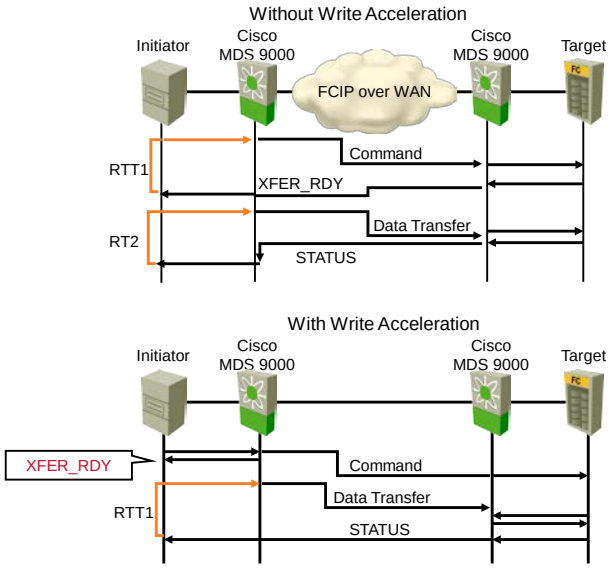
© 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

FCIP Write Acceleration

- Normal SCSI write requires two round trips:
  - WRITE > XFER_RDY
  - DATA > STATUS
- Write accelerator spoofs XFER_RDY.
- Write accelerator allows a single round trip over the WAN.



© 2012 Cisco and/or its affiliates. All rights reserved.

When FCIP write acceleration is enabled, WAN throughput is maximized by minimizing the impact of WAN latency for write operations. The figure shows that a Small Computer Systems Interface (SCSI) write command without write acceleration requires two round trips, while a write command with write acceleration requires only one round trip.

With write acceleration, the SCSI transfer ready (XFER_RDY) message is sent from the host side of the FCIP link back to the host before the write command reaches the target. This allows the host to start sending the write data without waiting for the long latency over the FCIP link of the write command and the returning XFER_RDY message. It also eliminates the delay that is caused by multiple XFER_RDY messages that are needed for the exchange going over the FCIP link.

https://t.me/learningnets

certcollection.net

FCIP Tape Acceleration

Tape drives cannot manage high WAN latencies:

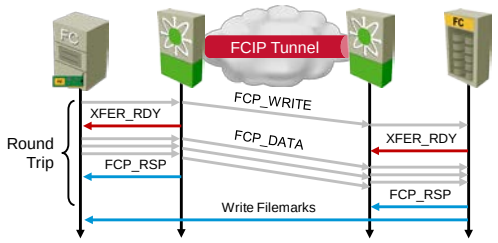
- Cannot keep the tape streaming, which causes shoe-shining.

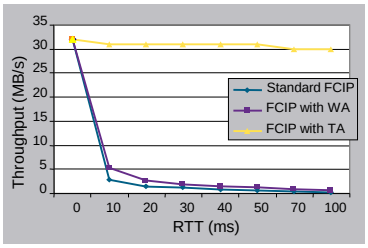
Write acceleration alone cannot keep the tape streaming:

- Tape drives allow only one outstanding I/O.

Tape acceleration is an enhancement to write acceleration:

- Spoofs FCP response so next write operation is not delayed.
- Extends tape buffering onto the IPS modules.
- IPS modules act as proxy tape device and backup host.





WA = write acceleration
TA = tape acceleration

© 2012 Cisco and/or its affiliates. All rights reserved.

More customers are realizing the benefits of tape backup over WAN in terms of centralizing tape libraries and maintaining central control over backups. With increasing regulatory oversight of data retention, these benefits are growing in importance.

One issue that customers often face is that tape drives have limited buffering that is often not sufficient to process WAN latencies.

Even with write acceleration, each drive can support only one outstanding I/O.

When the tape drive writes a block, it issues an FCP response frames only (FCP_RSP) status command to tell the initiator to send more data. The initiator then responds with another FCP write command. If the latency is too high, the tape drive will not receive the next data block in time and must stop and rewind the tape. This “shoe-shining” effect not only increases the time that it takes to complete the backup job (potentially preventing it from completing within a reasonable time), but it also decreases the life of the tape drive.

Write acceleration alone is not sufficient to keep the tape streaming. It halves the total RTT for an I/O, but the initiator must still wait to receive the FCP_RSP message before sending the next FCP write.

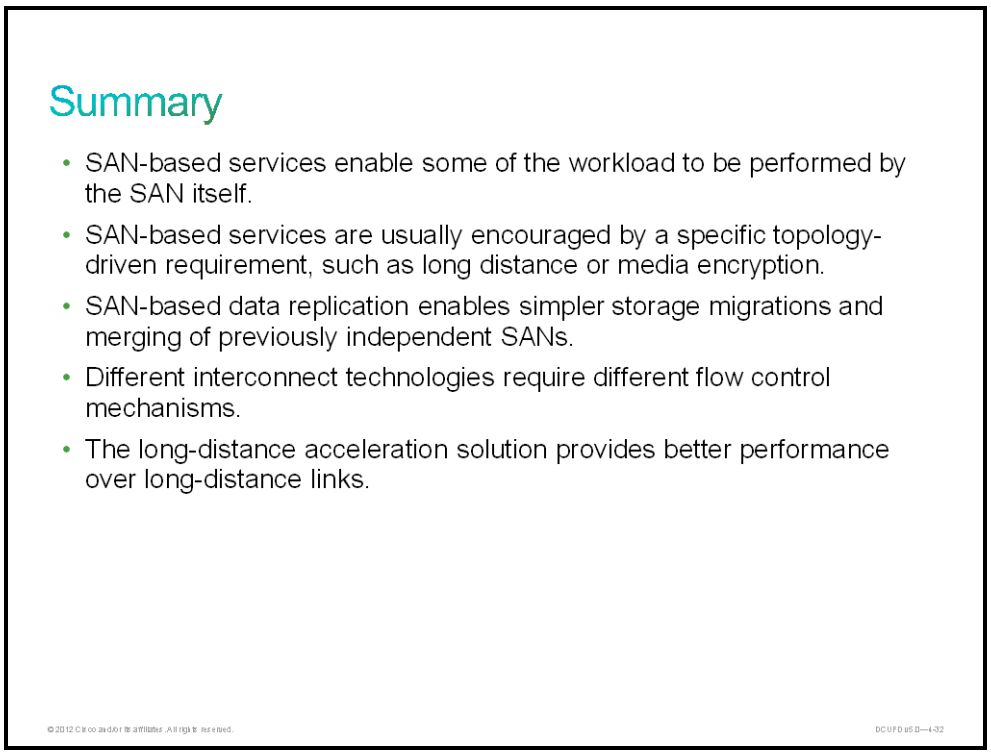
FCIP tape acceleration is an enhancement to write acceleration that extends tape buffering onto the Cisco IPS-capable modules. The local Cisco IPS-capable module proxies as a tape library, while the remote Cisco IPS-capable module proxies as a backup server. The local Cisco IPS-capable module sends an FCP_RSP message back to the host immediately after receiving each block, and data is buffered on both Cisco IPS-capable modules to keep the tape streaming. It includes a flow control scheme to avoid overflowing the buffers, which allows the Cisco IPS-capable module to compensate for changes in WAN latencies or tape speed.

https://t.me/learningnets

certcollection.net

Summary

This topic summarizes the primary points that were discussed in this lesson.



https://t.me/learningnets

certcollection.net

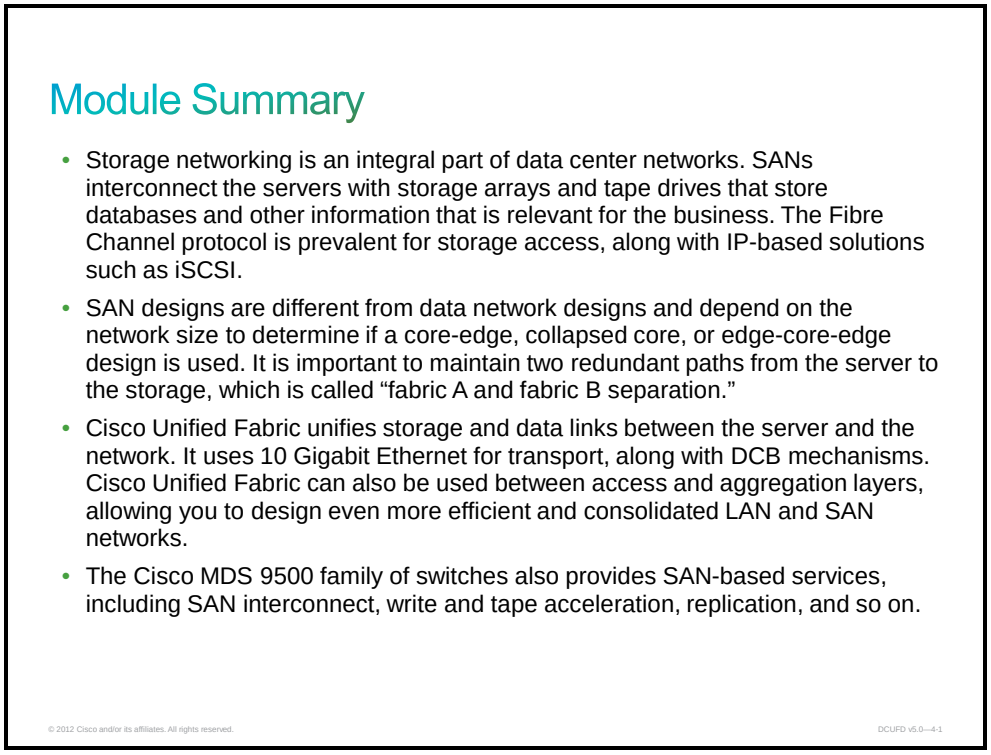
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

Module Summary

This topic summarizes the primary points that were discussed in this module.



SAN networks form an important part of data center designs. The main protocol is Fibre Channel Protocol (FCP), which has its own logic and design. Native Fibre Channel fabrics operate using Cisco MDS switches and always provide two independent paths from servers to storage, which is a design principle that must always be followed to achieve required redundancy in designs.

https://t.me/learningnets

certcollection.net

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

Module Self-Check

Use these questions to review what you learned in this module. The correct answers and solutions are found in the Module Self-Check Answer Key.

Q1) What are the three most common storage access methods that are used in data centers? (Choose three.) (Source: Introducing SAN)

A) NAS
B) iSCSI
C) FTP
D) SAN
E) NFS

Q2) What are the two most important services in a Fibre Channel SAN? (Choose two.) (Source: Introducing SAN)

A) login server
B) name server
C) security server
D) syslog server

Q3) Which kind of a flow control is used in Fibre Channel networks? (Source: Introducing SAN)

A) TCP sliding window flow control
B) buffer-to-buffer credits flow control
C) priority flow control
D) hardware flow control

Q4) What are the three most common topologies in SAN fabrics? (Choose three.) (Source: Designing SAN)

A) core-edge
B) edge-core-edge
C) core-aggregation-access
D) collapsed core

Q5) What is the primary motivation for using NPV in modern SAN fabrics? (Source: Designing SAN)

A) to enable every switch to get a Fibre Channel domain ID
B) to allow all servers to log in to the fabric simultaneously
C) to allow for easy scalability without running out of domain IDs
D) to fine-tune traffic paths

Q6) What are the two features that allow you to design a true multitenant SAN fabric? (Choose two.) (Source: Designing SAN)

A) VSANs
B) NPV
C) zoning
D) data encryption

https://t.me/learningnets

certcollection.net

Q7) What must be enabled on the network in order to transport an FCoE frame? (Source: Designing Unified Fabric)

A) OSPF
B) jumbo frames
C) FCIP
D) PoE

Q8) Which three elements are used by DCB-enabled Ethernet? (Choose three.) (Source: Designing Unified Fabric)

A) priority-based flow control
B) Enhanced Transmission Selection
C) Enhanced Fast Software Upgrade
D) per-priority pause
E) Data Center Bridging Protocol

Q9) What are the two roles of FIP? (Choose two.) (Source: Designing Unified Fabric)

A) Allow the end node to boot from the SAN.
B) Assign a MAC address to the end node.
C) Allow for seamless communication between servers.
D) Allow the end node to find the FCoE VLAN and the closest FCF.

Q10) In which two cases would you have the need for SAN-based services? (Choose two.) (Source: Designing SAN Services)

A) when performance would be affected if storage-based services would be used
B) when host-based services are available
C) when a host-based or a storage-based feature is impractical to use
D) when a host-based or a storage-based feature has licensing restrictions

Q11) What is the main benefit of SAN-based data migration? (Source: Designing SAN Services)

A) selective migration of data
B) migration of data between LUNs without host intervention
C) migration of data between LUNs using the storage array
D) migration of data between hosts using the storage array

Q12) What is the most popular protocol to implement long-distance Fibre Channel interconnects? (Source: Designing SAN Services)

A) FCIP
B) iSCSI
C) FIP
D) dark fiber

Q13) What is the flow control mechanism that is used on FCIP tunnels? (Source: Designing SAN Services)

A) priority-based flow control
B) selective acknowledgement
C) TCP sliding window
D) LZW

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

Module Self-Check Answer Key

Q1) A, B, D

Q2) A, B

Q3) B

Q4) A, B, D

Q5) C

Q6) A, C

Q7) B

Q8) A, B, D

Q9) B, D

Q10) C, D

Q11) B

Q12) A

Q13) C

https://t.me/learningnets

certcollection.net

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

Module 5

Data Center Security

Overview

In this module, you will learn how application and data security is managed in data center networks. Application security is provided using firewalls or security appliances that can filter, check, and inspect the traffic going toward servers that are running an application. In this manner, servers are protected from attacks and intrusions from within the traffic flow.

Link security is another important subject for security-conscious environments or for where regulatory requirements exist for data to flow encrypted on the link. Cisco switches and directors feature link encryption to satisfy these needs.

Module Objectives

Upon completing this module, you will be able to design secure data centers that are protected from application-based threats, network security threats, and physical security threats. This ability includes being able to meet these objectives:

- Design secure data center networks on the application level
- Design secure data center networks on the network and device level
- Design secure data center SANs

https://t.me/learningnets

certcollection.net

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

Lesson 1

Designing Data Center Application Security

Overview

This lesson enables you to design firewall security features to reduce the security risks in the data center environment. The importance of the firewall and its position in the data center is explained. The lesson also describes the role of the load balancer in the data center environment.

Objectives

Upon completing this lesson, you will be able to design secure data center networks on the application level. This ability includes being able to meet these objectives:

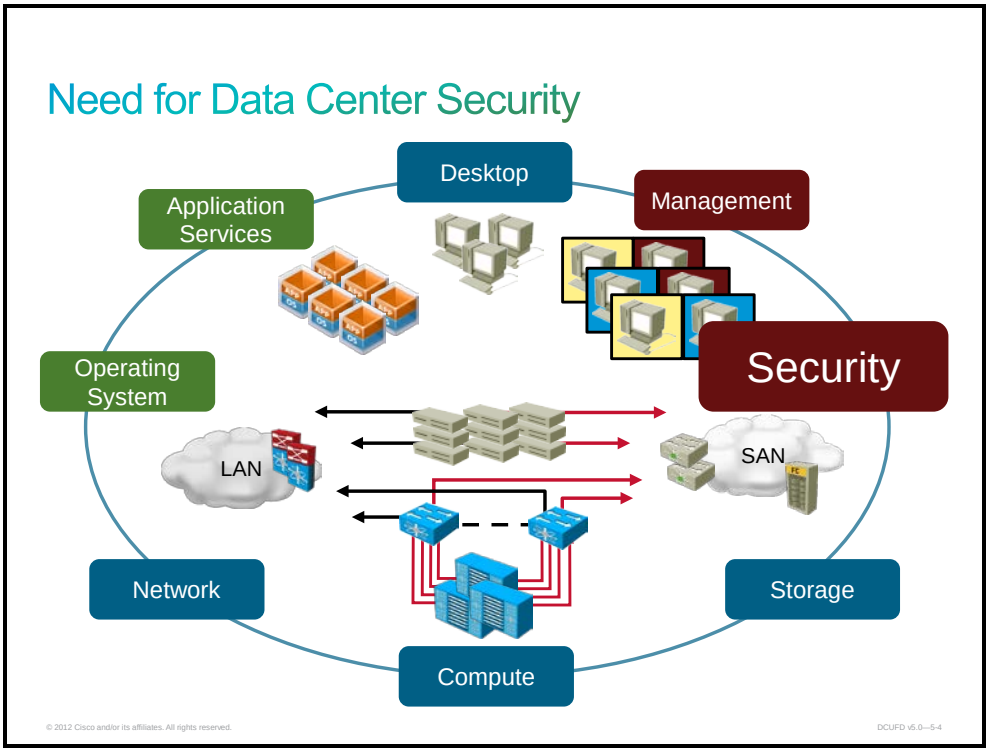
- Identify the need for data center security technologies
- Describe characteristics of firewalls
- Position security appliances within data center networks
- Design secure communication on multiple layers

https://t.me/learningnets

certcollection.net

Need for Data Center Security

This topic describes how to identify the need for data center security technologies.



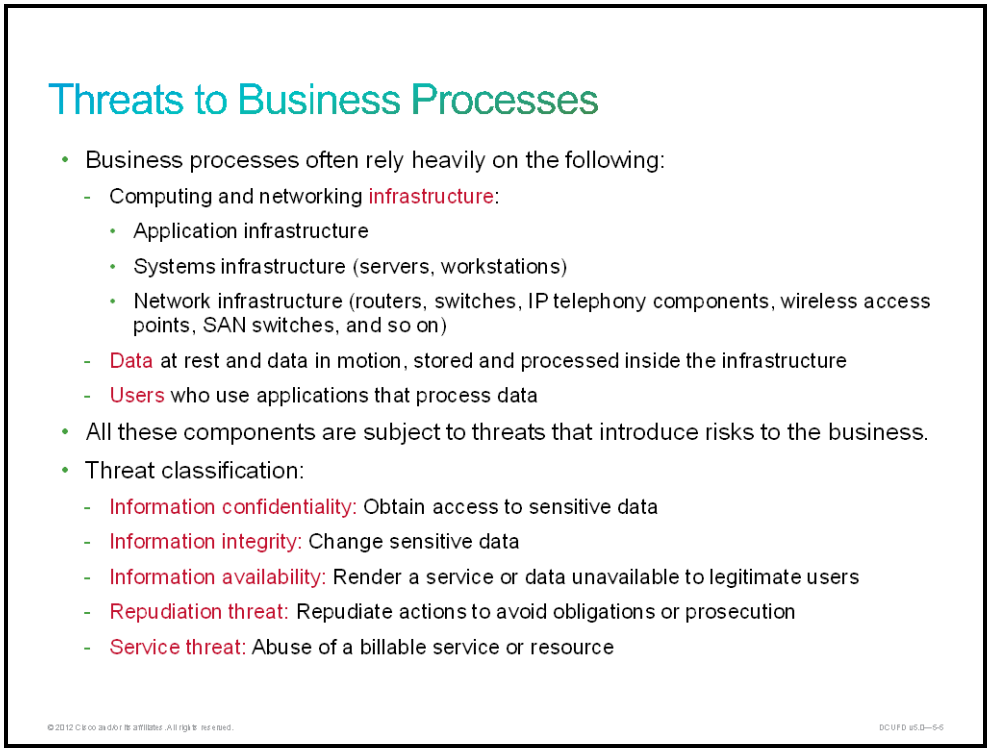
Contemporary data center computing solutions encompass multiple aspects and technologies, as shown in this figure. Security technologies and equipment are employed to ensure confidentiality and security to sensitive data and systems.

Data center design must provide physical and logical security. All critical components of server and data security are within the data center itself. All entry points to the data center must be controlled and monitored. Data center servers, storage devices, and fabric operation can be attacked over the network. Almost all valuable information of an organization is stored online. Therefore, it is imperative to protect data and servers, from people with malicious intent, by using different restricted access policies, authentication, and authorization mechanisms.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net



Modern business processes rely heavily on the underlying information and communications infrastructure that comprises the following components:

- Applications, systems (such as workstations, servers, laptops, handheld devices, mobile phones, and IP phones), and the network infrastructure (such as routers, switches, telephony components, wireless access points, and SAN switches) that provide the foundation for data processing and data transfer between various components and external parties
- Data at rest (stored on systems and SANs) and in motion (transferred over networks) that is processed within the foundation infrastructure
- Infrastructure users and administrators who use and manage systems and applications that store, process, and transfer data over the foundation infrastructure

All of these components are integral parts of an organization and are subject to threats that are caused by active malicious agents that introduce risk to the business and its processes.

Here is some basic security-related terminology:

- **Asset:** An asset is anything that has value to an organization. An asset can be a process, a user, a database record, a USB flash drive, a network device or link, or a mainframe computer.
- **Threat:** A threat is any circumstance or event with the potential to cause harm to an information system in the form of destruction, disclosure, adverse modification of data, or the denial of service. Examples of threats are application layer network attacks against exposed application servers, malware targeting workstations, or physical destruction of a server.

https://t.me/learningnets

certcollection.net

- **Vulnerability:** A vulnerability is a weakness in a system or its design that could be exploited by a threat. Examples of vulnerabilities are software defects in browsers that allow client systems to be compromised, misconfigurations of network devices that allow for routing protocol injection, or employees allowing strangers to enter sensitive locations. Closely related to a vulnerability, an exploit is a piece of software, a chunk of data, or a sequence of commands that takes advantage of a vulnerability in order to cause unintended or unanticipated behavior to occur on computer software or hardware.
- **Risk:** A risk is the likelihood that a particular threat using a specific attack will exploit a particular vulnerability of an asset that results in an undesirable consequence.

Threat Classification

Information security is about protecting information and information systems from unauthorized access, use, disclosure, disruption, modification, or destruction. Security practitioners usually divide information security threats into three to five main threat classes.

These are the three major threat classes:

- **Confidentiality:** Threats to information confidentiality, in which attackers attempt to disclose (that is, obtain “read” access to) sensitive data
- **Integrity:** Threats to information integrity, in which attackers attempt to change (that is, obtain “write” access to) sensitive data
- **Availability:** Threats to information availability, in which attackers attempt to render a service or data unavailable to legitimate users

There are two additional classes of threats, which may not be applicable in all environments:

- **Repudiation:** Threats in which an entity can repudiate their actions in a system in order to avoid obligations or prosecution
- **Theft of service:** Threats in which an attacker abuses a billable service or resource at the expense of the service and resource owner

These threat classes are specific to an environment. For example, a specific organization might be mostly concerned about integrity and availability threats in a specific context. Security practitioners must therefore always determine the importance of a particular threat class when designing a secure system.

 © 2012 Cisco Systems, Inc.

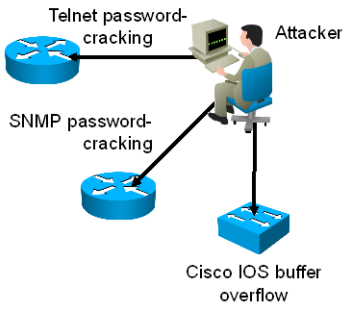
https://t.me/learningnets

certcollection.net

Threats to Business Processes (Cont.)

Device Compromise

Threat	Preventive Controls
Exploitation of software defects	Device software patching Minimized device network exposure (host firewall, disabling of services)
Installation of malicious software	Software authenticity verification, digitally signed software
Configuration errors	Secure management practices: templates, best practices, configuration integrity verification, administrator auditing, change control
Impersonation of trusted partners	Cryptographic protection of signaling and management protocols Strong authentication and access control to management functions



© 2012 Cisco and/or its affiliates. All rights reserved.

The first important variety of network infrastructure threats are those that result in network device compromise, in which the attacker is able to take full or partial control of a network infrastructure device and abuse the device to attack business processes that are supported by the network.

Such a compromise may be possible due to the following factors:

- Software defects in the operating system of the network device, which may allow exploitation and, therefore, unauthorized access to the device. You can reduce the risk of these attacks by minimizing the exposure of the device to network traffic (for example, by filtering traffic using its host firewalling functions or by disabling all unused services) and by timely patching of device software.
- Installation of malicious device software, either inadvertently by the administrator or maliciously by an attacker. You can reduce the risk of these attacks by verifying software authenticity, using manual or automatic (digital signature) verification.
- Misconfiguration of the device, which is especially likely if no security and provisioning standards are set within the organization and if the configuration of similar devices varies from device to device. You can reduce the risk of these attacks by using secure management practices, such as the use of configuration templates that ensure consistent device configuration, periodic verification of configuration correctness and integrity, auditing of administrator actions, and change control procedures.
- Impersonation of trusted partners in device management and signaling channels, such as network management sessions supporting administrative access, in which attackers can obtain unauthorized administrative access to devices or influence the operation of a device. You can reduce the risk of these attacks through strong cryptographic protection of device signaling and management protocols, and strong authentication and access control to the management functions of the device.

To illustrate the last point, spoofed Network Time Protocol (NTP) packets that are sent to a router could change its sense of time and change the way that time-based access control lists (ACLs), public key infrastructure (PKI), logging, and other functions work.

https://t.me/learningnets

certcollection.net

Many attacks are not limited to the device itself, because the device may have many trust channels (such as management protocols, NTP, and so on) with other partners. An attacker that is compromising a single device can use these channels to influence peer devices and possibly, due to a chain of trust, many other devices. Therefore, the attack range of even a single compromised device can easily include the entire network.

Attack Examples

Examples of device compromise attacks include the following:

- Exploitation of device software security defects, in which the attacker attempts to exploit a known software security issue on the device, either locally (by being logged on the device) or remotely (using the exposed network services of a device), or by sending malicious, vulnerability-triggering traffic through the device. The general consequence of such attacks is the attacker gaining partial (user-level) or full (administrator-level) control of the device operation.
- Guessing of user and administrator passwords to log on to the management interface of a device. These attacks are either active (in which the attacker attempts to log on using a dictionary or brute-force list of passwords) or passive (in which the attacker captures a valid authentication exchange and attempts to cryptographically break it to recover user or administrator credentials).
- Administrative session spoofing, in which the attacker attempts to spoof certain properties of administrative sessions (such as source IP address or other protocol-level credentials) to log on to a device.
- Administrative session hijacking, in which the attacker attempts to take over an existing administrative session, usually using mechanisms such as Address Resolution Protocol (ARP) spoofing.
- Network device rootkits, in which the attacker installs a rootkit (which is a piece of software that is designed to hide the presence of an attacker and the fact that the device has been compromised) on an already-compromised device to evade detection of the compromise and subsequent malicious activity.

 © 2012 Cisco Systems, Inc.

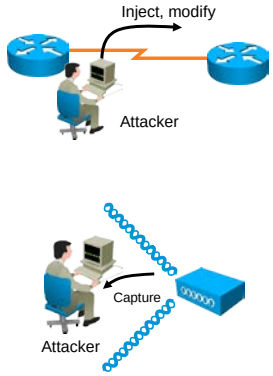
https://t.me/learningnets

certcollection.net

Threats to Business Processes (Cont.)

Traffic Capture and Injection

Threats	Preventive Controls
Traffic capture (passive or man-in-the-middle)	Protected routing and switching processes *or* End-to-end, gateway-to-gateway (VPN), endpoint-to-gateway (VPN), or link cryptographic protection methods *or* Both of these controls
Changing of network traffic (man-in-the-middle)	
Spoofing of network traffic (injection, replay, man-in-the-middle)	



© 2012 Cisco and/or its affiliates. All rights reserved.

Network links that can be physically accessed by attackers can often be very easily compromised to allow the attacker to either passively monitor or actively intercept or inject traffic on the compromised communications link. WAN links that are outside the physical control of the enterprise are generally considered to be subject to such threats. With wireless technologies, these attacks only require the attacker to be in proximity of wireless links and access points, possibly outside the enterprise physical perimeter.

You can reduce the risk of these attacks by protecting the network infrastructure routing and switching processes against malicious manipulation, and often by providing data transmission security by cryptographically protecting data in transit over the network using VPN, or end-to-end transmission protection.

Attack Examples

Examples of link compromise attacks include the following:

- Capturing of sensitive information over the link by snooping on cleartext data flowing through a physically compromised link. The consequence of such attacks is the disclosure of confidential information flowing over the compromised link.
- Active modification of sensitive information flowing over the link by intercepting legitimate traffic and changing it in real time before sending it to its destination. The consequence of such attacks is the violation of integrity of sensitive information flowing over the compromised link. Additionally, an attacker can spoof and replay traffic over the link by sending traffic with spoofed identity information or by resending previously seen legitimate traffic to achieve a particular goal.
- Traffic analysis, in which the attacker passively monitors data flowing over a link and obtains “meta-information” about an organization (for example, who is talking to whom, which hosts appear to be most important, and so on) by analyzing communicating addresses, the amount of data transferred, the times of communication, and so on. With these attacks, an attacker obtains information about the structure of the enterprise infrastructure and business processes.

https://t.me/learningnets

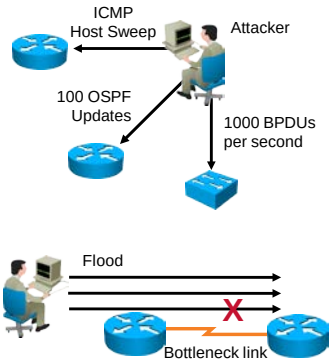
certcollection.net

- Denial of service (DoS), by intercepting communications over a compromised link and refusing to forward legitimate data. The consequence of such attacks is DoS to network applications that support important business processes.

Threats to Business Processes (Cont.)

Device and Link DoS

Threat	Preventive controls
Malformed traffic to and through devices	Device software patching Minimized device network exposure (host firewall, disabling of services)
Flooding targeting device slow paths	Controlled use of device resources (CPU, memory) Edge user authentication and compliance verification
Flooding targeting network links	Quality of service mechanisms Anomaly detection and response Edge user authentication and compliance verification



BPDUs = bridge protocol data units

© 2012 Cisco and/or its affiliates. All rights reserved.

Device-Focused DoS Attacks

Network devices can be targets of DoS attacks, with attackers attempting to slow down or stop device operation. One method to achieve this is by sending malformed packets to be switched across them or by sending malicious information to the exposed services of the device. You can reduce the risk of these attacks by patching device software in a timely fashion against known DoS vulnerabilities and minimizing the exposure of a device to malicious traffic.

The other form of device-focused DoS attacks involves flooding attacks that attempt to exhaust device resources. These attacks come in several varieties:

- Sending a high rate of (forged) signaling packets (such as Open Shortest Path First [OSPF] or Spanning Tree Protocol [STP]), forcing path recalculations, impacting the router CPU.
- Sending traffic that requires special processing to the device. Generally, devices are designed to process low levels of such special traffic and usually choke on it when under such an attack.
- Exhausting resources of the network stack. A classic example is a TCP SYN flood, in which the network stack must process a series of pending TCP synchronization (TCP SYN) requests. Many incomplete TCP handshakes can cause the stack to allocate an excessive amount of resources and eventually stop accepting new connections.

You can reduce the risk of these attacks by controlling the use of device resources, such as CPU and memory, and using features that can limit traffic rates or resource consumption. Additionally, by limiting access to only authorized users on the network edge, you can reduce the risk of network worms that could generate traffic patterns that might cause the network infrastructure to behave suboptimally.

https://t.me/learningnets

certcollection.net

Attack Examples

Examples of device DoS attacks include the following:

- Malformed packets and requests that are sent by the attacker, targeting the local TCP/IP stack and local applications, in which the attacker sends abnormal TCP/IP packets or abnormal requests to local applications (such as the Simple Network Management Protocol [SNMP] server, the Session Initiation Protocol [SIP] listener, or the HTTP server). The typical consequence of these attacks is to reload or freeze the attacked device.
- Malformed packets that are sent through the device (transit traffic) that try to poison and disable the forwarding engine of a device.
- Excessive traffic on the control plane, in which the attacker sends excessive packets to slow processing paths (such as demand switching and flow processing code) that are used by a device, or excessive request rates to local applications of a device. The typical consequences of these attacks are to reload or freeze the attacked device, or slow down its traffic forwarding or management functions.
- Excessive transit traffic, in which the attacker attempts to overload the normal (fast) forwarding paths of the device by simply exceeding the traffic forwarding capacity of the device. Again, the typical consequences of these attacks are to reload or freeze the attacked device, or slow down its traffic forwarding or management functions.

Link-Focused DoS Attacks

Network links can also be subject to remote DoS attacks, in which the attacker attempts to stop or slow down traffic flow through a link, by either sending excessive traffic rates over a bottleneck link that cannot support such throughput, by spoofing control messages that manage the link, or through interference attacks that involve overpowering normal communication signals through signal jamming.

You can reduce the risk of these attacks using quality of service (QoS) mechanisms on links or by using anomaly-based detection engines that examine traffic or traffic telemetry to identify such attacks.

Attack Examples

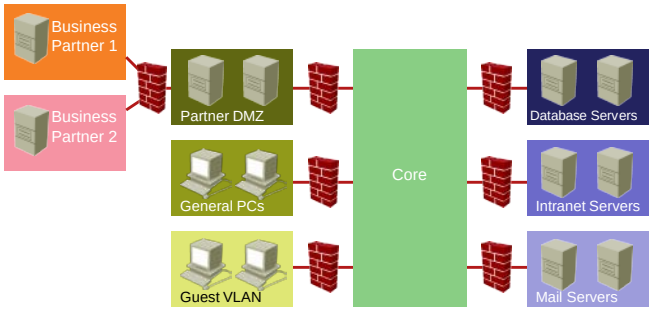
Examples of link DoS attacks include the following:

- Traffic flooding, in which the attacker sends a high rate of network traffic over a bottleneck link. The consequence of such an attack is DoS for all network applications using that particular link.
- Distributed traffic flooding (known as a distributed DoS [DDoS] attack), in which such traffic flooding is directed to a target from many (hundreds or thousands) sources, making it extremely difficult to respond to. The consequence of such an attack is generally a prolonged DoS condition for all network applications using that particular link.
- Spoofing control messages that manage the link, in which the attacker attempts to disconnect the logical link between legitimate devices using spoofed control packets. Such attacks were popular against wireless networks that did not authenticate control packets, which enabled attackers to disconnect wireless service to specific hosts and deny them wireless connectivity.
- Signal strength attacks, in which the attacker uses a powerful transmitter to overpower the legitimate signal source with a stronger signal, rendering the legitimate service unavailable. Again, this attack is typical of wireless networks.

https://t.me/learningnets

certcollection.net

Necessity for Firewalling the Data Center



- One common approach to protecting network resources involves partitioning of the network into security domains and implementation of boundary filtering (firewalling).
- Security domains use physical or logical (VLAN, VRF, MPLS, and so on) separation methods to ensure a single point of transit (that is, a chokepoint) between domains.

DMZ = demilitarized zone

© 2012 Cisco and/or its affiliates. All rights reserved.

A common approach to securing a system involves separating the system into individual parts and minimizing the interactions between these parts. This approach is commonly applied to enterprise networks, where security designers partition the network into security domains, based on the sensitivity of the data that is managed in a particular domain or the trustworthiness of endpoints in a particular domain. In this approach, boundary filtering systems—or firewall systems—control all network interactions between adjacent domains to reduce risk.

Security domains are separated by physical or logical separation methods, which ensure that traffic can flow between domains only through a single designated transit point (a chokepoint).

With physical separation of domains, a firewall system connects to two physically distinct network infrastructures using two physical network interfaces, and all traffic between the two domains pass through the system. From a security perspective, physical separation is always the best method of domain separation, because it can only be circumvented by physical means or other compromise of the firewall system. However, this separation can be costly, especially when multiple security domains need to be created in large access networks, in which a common switched infrastructure interconnects systems of different roles, such as IP phones, clients, and servers.

Logical separation provides separate communication channels for different groups of users over the same physical infrastructure. Such logical separation methods include VLANs, virtual storage area networks (VSANs), or Multiprotocol Label Switching (MPLS) VPNs, in which tagging of LAN or WAN frames or packets provides separation between domains. Logical separation introduces additional risks inside the separation mechanism itself, which could fail and enable a bypass of the firewall system. For that reason, logical separation is less trusted than physical separation. However, the cost benefits of logical separation may offset its potential security shortcomings, and with the current trends in IT virtualization, such separation may become standard in many environments, after a proper risk assessment has been conducted.

https://t.me/learningnets

certcollection.net

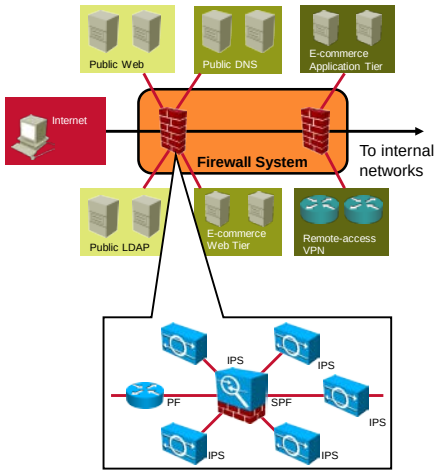
Firewall Characteristics

This topic describes the characteristics of firewalls.

Firewall Definition

A firewall is a system that enforces access control between security domains:

- It should be resistant to attacks itself and the only transit point between security zones.
- A “system” means it can be made of many devices (such as stateful packet filters, proxies, and network IPS devices).



LDAP = Lightweight Directory Access Protocol
SPF = stateful packet filter
PF = packet filter

© 2012 Cisco and/or its affiliates. All rights reserved.

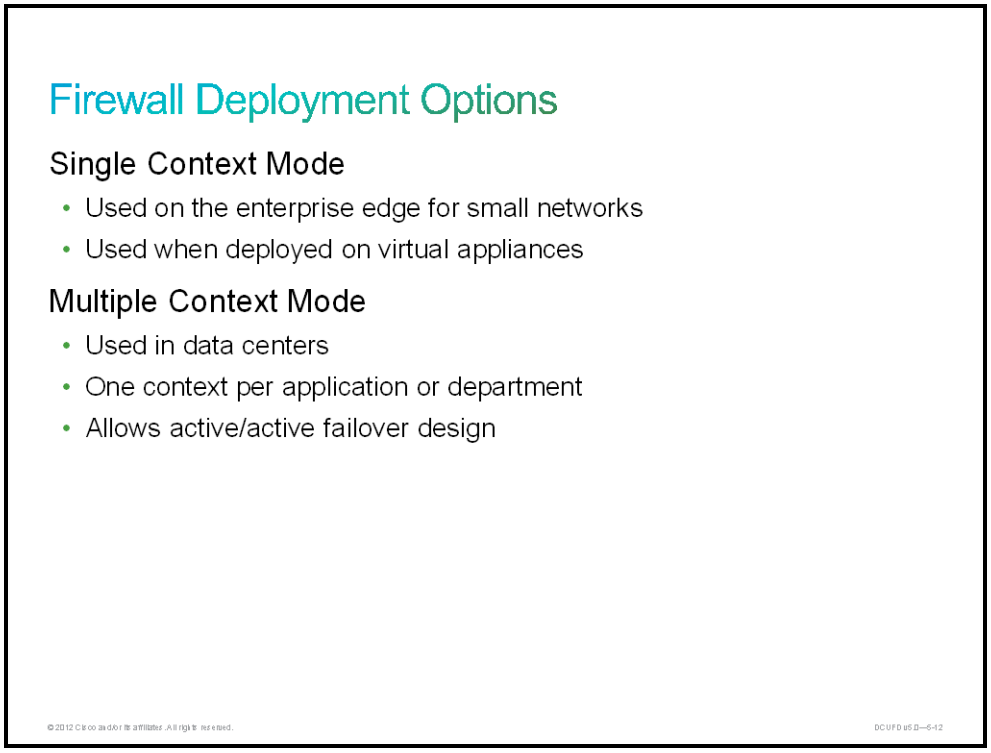
A firewall is a system that enforces an access control policy between two or more security domains. All firewalls share two common properties:

- The firewall itself must be resistant to attack, otherwise it would enable an attacker to disable the firewall or change its access rules to bypass its controls.
- All traffic between security domains must flow through the firewall. This prevents a backdoor connection that could be used to bypass the firewall, violating the desired access control policy.

A firewall system can be a single device or a set of devices, each device providing a specific traffic filtering role to achieve the desired set of controls (countermeasures). For example, a firewall designer may choose to include stateful filtering devices, advanced application inspections, proxy-based devices, network intrusion prevention systems (IPSs), and similar components to build a firewall system.

https://t.me/learningnets

certcollection.net



The Cisco ASA adaptive security appliance and Adaptive Security Appliance Security Module (ASA-SM) firewalls can be deployed in single or multiple context modes.

Single Context Mode

In the single context mode, the firewall forwards, filters, and inspects the traffic in the main system context. This mode simplifies administration, because all administration is done in the same, main system context.

Examples of such firewall deployments are at the enterprise edge, in small branches, in small office, home office (SOHO) environments, and when deploying virtualized appliances in the data center.

Fault-tolerant configuration is possible in this mode, and one firewall is always the standby unit.

Multiple Context Mode

In multiple context mode, the firewall is divided into several security contexts, each one having its own interfaces, filtering and inspection rules, configuration, and dedicated resources.

This is a typical deployment mode for physical firewalls in a data center, where every application, department, or customer gets its context. The configuration is contained within that context, and the firewall is managed through a dedicated context, called the “admin context.”

Fault tolerance is achieved by two firewalls running in multiple context mode, where some contexts are active on the first unit and some contexts are active on the second unit. This is how active/active mode is accomplished.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

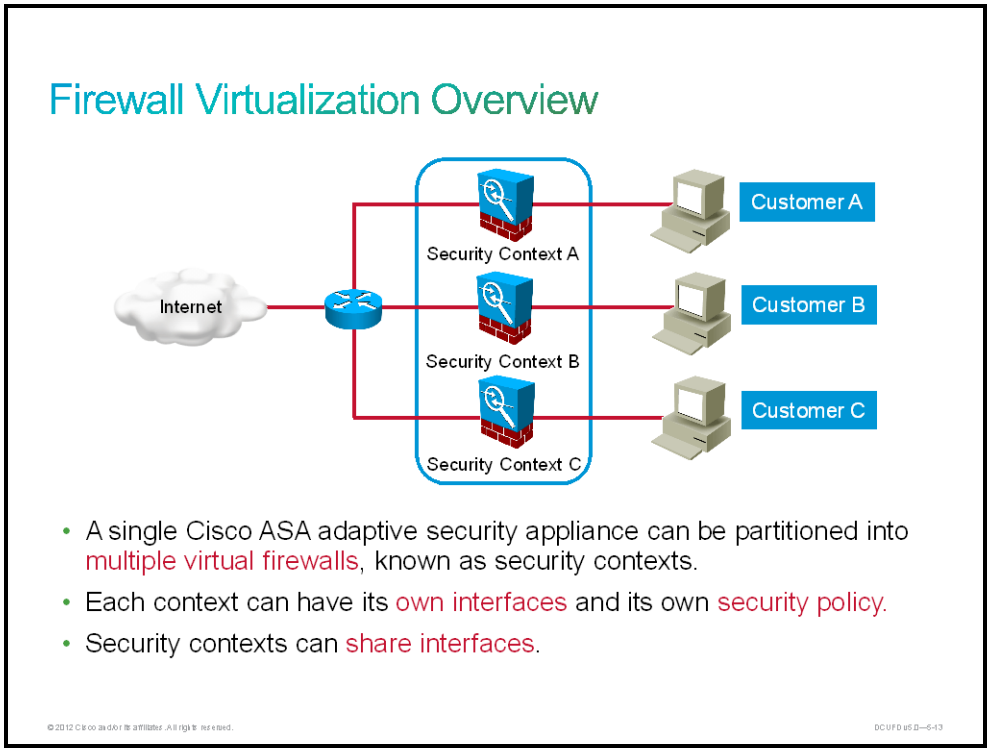
Mixing Firewall Modes on Multiple Contexts

With later software releases, the Cisco ASA adaptive security appliance firewalls can operate several contexts in different modes—some contexts can be in transparent mode and some contexts can be in routed mode.

Note Previously, the only device that was capable of such a configuration was the Cisco Firewall Services Module (FWSM). Mixed modes are now supported on the Cisco ASA adaptive security appliance and ASA-SM firewalls.

https://t.me/learningnets

certcollection.net



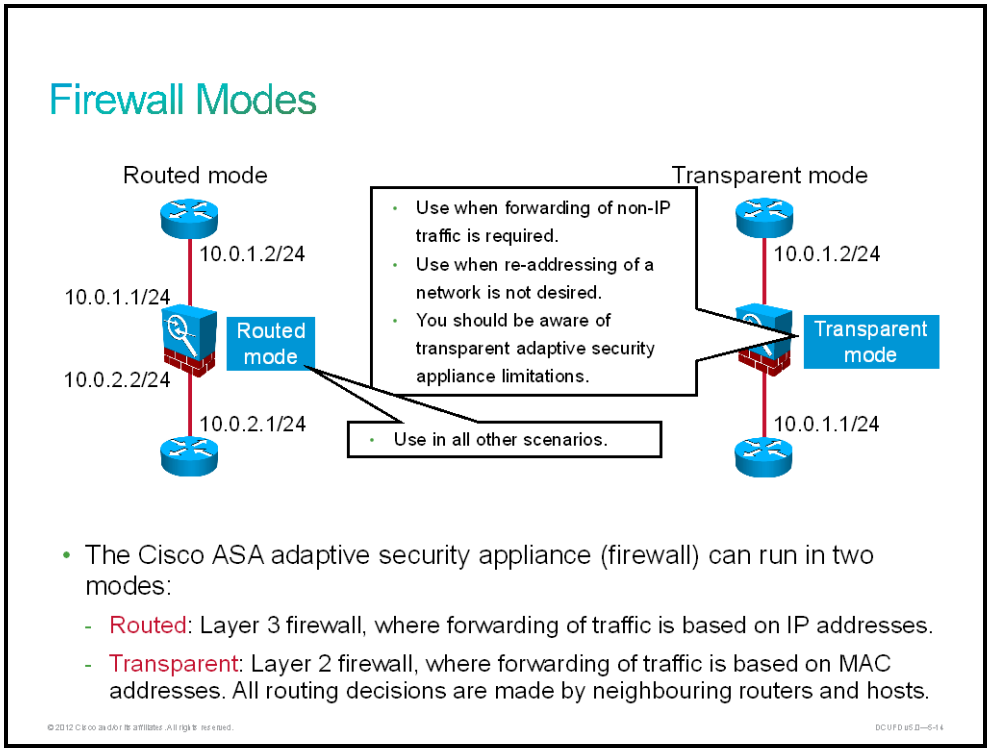
You can partition a single security appliance into multiple virtual firewalls, known as security contexts. Each context is an independent firewall, with its own security policy, interfaces, and administrators. Having multiple contexts is similar to having multiple standalone firewalls. The security appliance that hosts multiple security contexts must be first converted into “multiple mode,” which supports virtualization. Most of the single mode Cisco ASA adaptive security appliance features are also supported in multiple context mode, including static routing, access control features, security modules, and management features. Some features are not supported, including both IP Security (IPsec) and Secure Sockets Layer (SSL) VPNs, and dynamic routing protocols.

Each security context on a multiple mode Cisco ASA adaptive security appliance has its own configuration that identifies the security policy, interfaces, and almost all the options you can configure on a single mode firewall. Administrators can configure each context separately with having access to their own context only. In cases where different security contexts connect to the same network (for example, the Internet), you can use one physical interface that is shared across all security contexts.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net



Adaptive security appliances can run in two basic traffic-forwarding and network integration modes—routed mode and transparent mode.

Routed Firewall

In routed mode, the adaptive security appliance acts as a routed (Open Systems Interconnection [OSI] Layer 3) hop and presents itself as a router for hosts or routers that connect to one of its networks. The adaptive security appliance can cooperate in routing protocols or use static routes. Traffic forwarding across the adaptive security appliance is based on destination IP addresses.

Transparent Firewall

In transparent mode, the adaptive security appliance is a Layer 2 device that acts like a “bump in the wire” and is not seen as a routed hop to connected devices. The adaptive security appliance connects the same IP subnet on its inside and outside interfaces, and performs secured transparent bridging between the two interfaces. Traffic forwarding is based on destination MAC addresses. Access controls, such as access lists, authentication, authorization and accounting (AAA), stateful packet inspection, and application inspection control are supported for unicast IP version 4 (IPv4) and IP version 6 (IPv6) traffic. Other traffic, such as multicast and non-IP traffic, can pass the adaptive security appliance if you explicitly allow it with an access list.

Because a transparent firewall is not a routed hop, you can easily introduce a transparent firewall into an existing network. IP readdressing is unnecessary and network manageability can be simplified because there are no complex routing functions to troubleshoot.

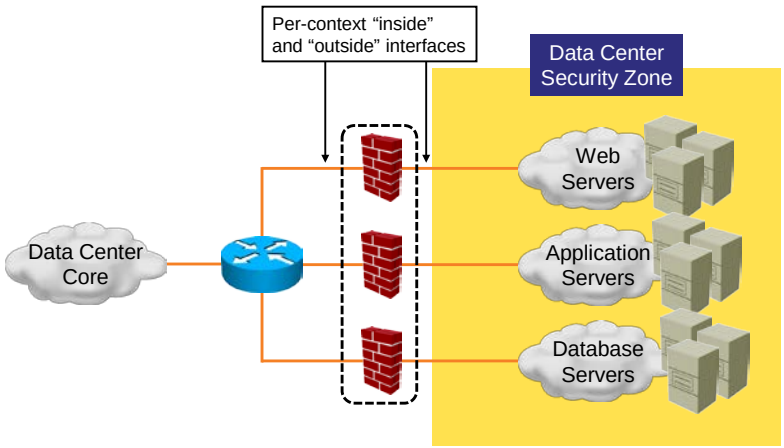
In the example on the left side of the figure, the adaptive security appliance in routed mode connects to two different IP networks, and each adaptive security appliance interface has an IP address from the appropriate subnet. The adaptive security appliance on the right side operates in transparent mode and connects to the same IP network on both interfaces. The interfaces do not have any IP addresses assigned. The only IP address that is needed on a transparent adaptive security appliance is a management IP address.

https://t.me/learningnets

certcollection.net

Dedicated Firewall Interfaces

- Multiple contexts that do not share interfaces
- Required for transparent mode



© 2012 Cisco and/or its affiliates. All rights reserved.

When using the firewall in multiple context mode, the contexts can either use interfaces exclusively or share them with another context.

Typically, a context in a data center firewall has one inside and one outside interface. More complex implementations use more interfaces with specific rules that allow data to pass from one security zone into another.

When using the firewall in transparent mode, inside and outside interfaces cannot be shared with other contexts. The reason is that the firewall bridges between the two interfaces, and sharing is not possible.

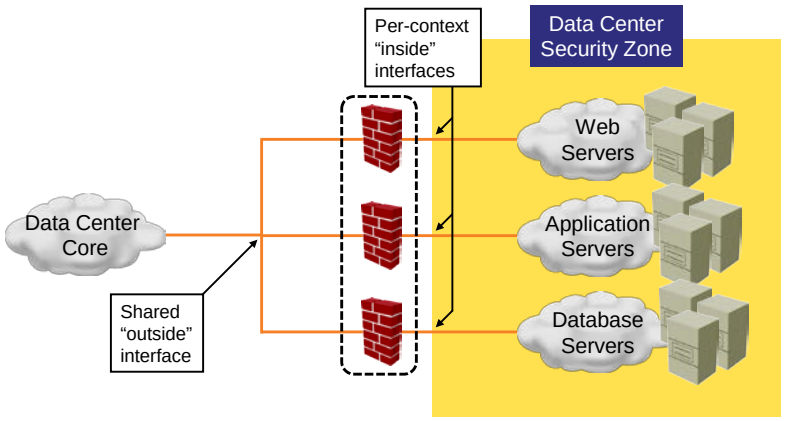
© 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

Shared Firewall Interfaces

- A single interface is shared among contexts.
- Cascading of the contexts on a single physical service module is not supported.
- Only routed mode is supported.



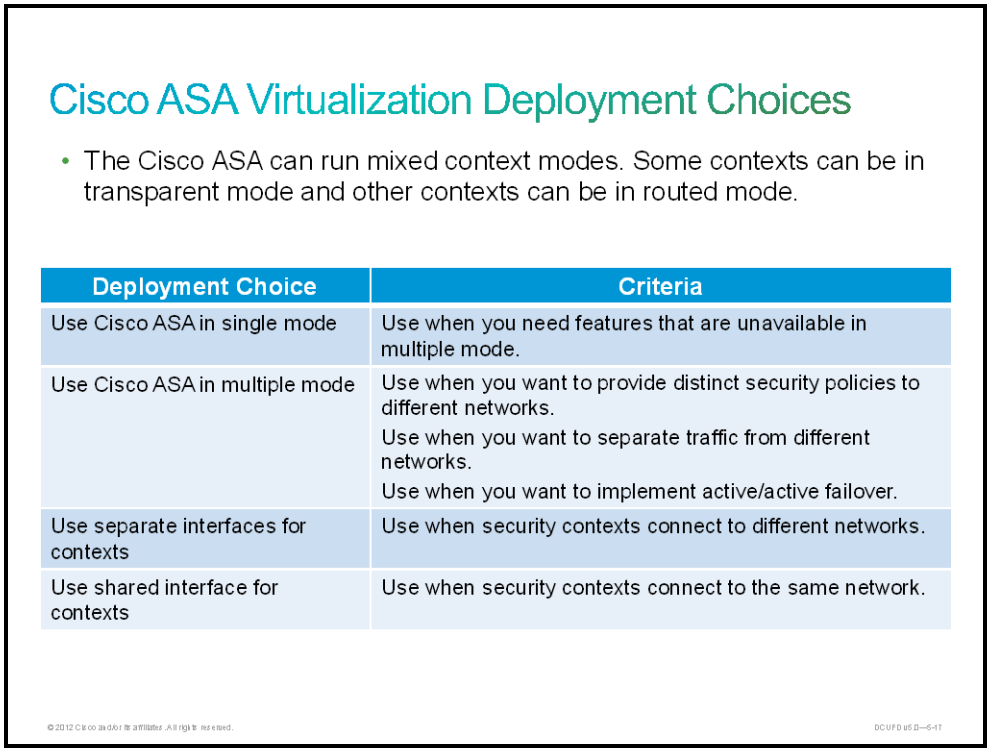
© 2012 Cisco and/or its affiliates. All rights reserved.

When using the firewall in routed mode, interfaces can be shared between the contexts because the firewall is routing traffic.

On the shared interface, the firewall runs a context classifier that places the packet in the correct context, where the packet is processed.

https://t.me/learningnets

certcollection.net



You should use the adaptive security appliance in multiple mode in these situations:

- You need to provide distinct security policies to different departments or users.
- You are a service provider and need to offer a different security context to each customer to separate traffic.
- You want to use the active/active failover feature. Active/active failover uses two contexts on the security appliance.

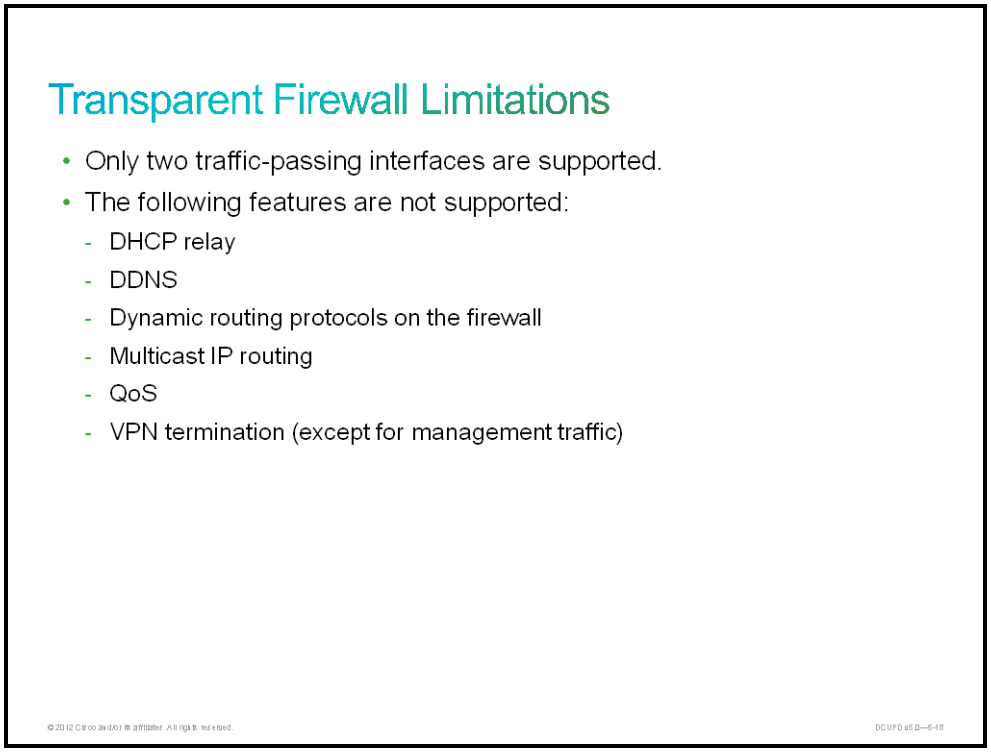
You should use the adaptive security appliance in single mode when you have to use features that are not available in multiple mode. These features are IPsec and SSL VPNs or dynamic routing protocols.

Another deployment option is the use of shared interfaces. You can use shared interfaces when the security appliance is in routed mode and the security contexts connect to the same network. When the security contexts connect to different networks, you should use separate interfaces. When you use the security appliance in transparent mode, you cannot use shared interfaces.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net



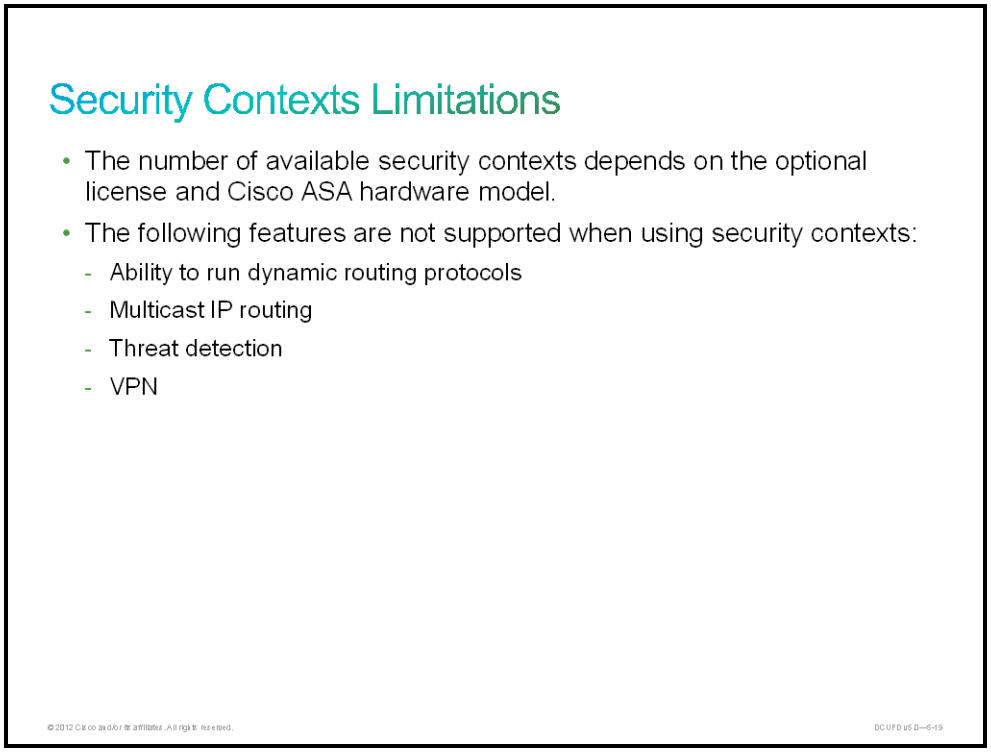
The transparent adaptive security appliance supports only two traffic-passing interfaces. If the adaptive security appliance platform supports a dedicated management interface, you can also enable the management interface for management traffic only.

The following features are not supported in transparent mode:

- **DHCP Relay:** The transparent firewall can act as a DHCP server, but it does not support the DHCP relay commands. DHCP relay is not required because you can allow DHCP traffic to pass through by using an extended ACL.
- **Dynamic Domain Name System (DDNS):** The firewall cannot act as a Layer 3 device.
- **Dynamic Routing Protocols:** The adaptive security appliance in transparent firewall mode will not be able to run any dynamic routing protocols. You can, however, add static routes for traffic originating on the adaptive security appliance. You can also allow dynamic routing protocols through the adaptive security appliance by using an extended ACL so that routers on each side of the transparent firewall can establish routing adjacency.
- **Multicast IP Routing:** You can allow multicast traffic through the adaptive security appliance by allowing it in an extended ACL.
- **QoS:** There is limited QoS support when runing in transparent mode.
- **VPN Termination:** The transparent firewall supports site-to-site VPN tunnels for management connections only. It does not terminate VPN connections for traffic through the adaptive security appliance. You can pass VPN traffic through the adaptive security appliance by using an extended ACL, but it does not terminate non-management VPN connections. SSL VPN is also not supported.

https://t.me/learningnets

certcollection.net



The maximum number of security contexts depends on the adaptive security appliance hardware model and on the optional security context license.

The following features are not supported in security contexts:

- Dynamic routing protocols
- Multicast IP routing
- Threat detection
- IPsec and SSL VPN

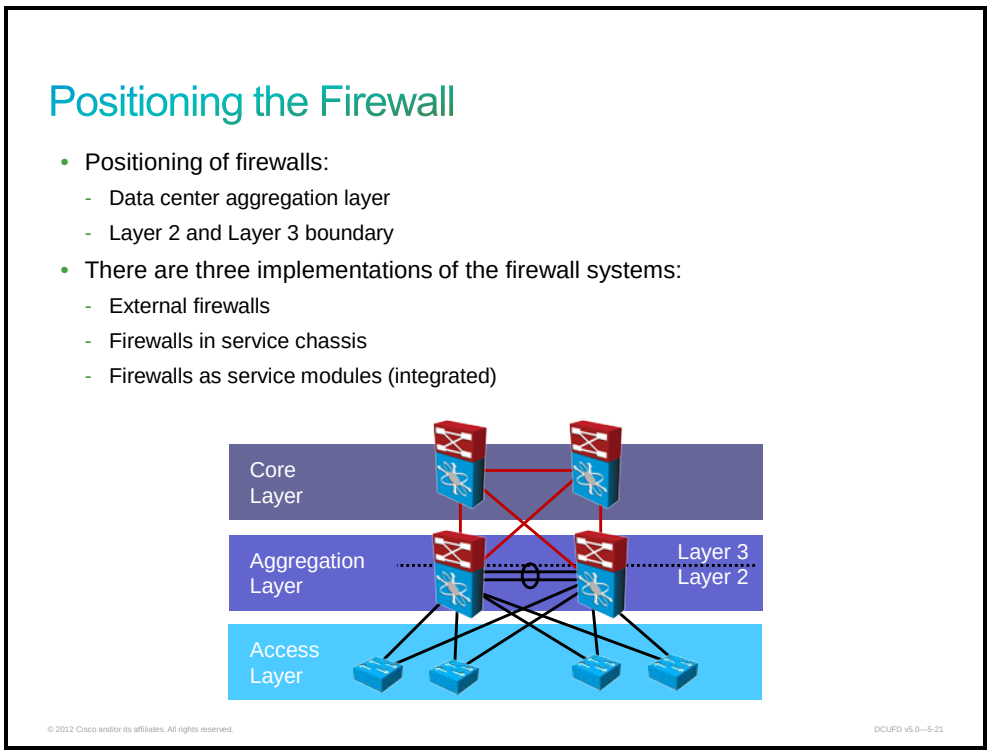
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

Positioning Firewalls Within Data Center Networks

This topic describes how to position security appliances within data center networks.



A firewall can be positioned at the Layer 2 or Layer 3 boundary, depending on the traffic that needs to pass the firewall. In routed mode, the adaptive security appliance acts as a routed (OSI Layer 3) hop and presents itself as a router for hosts or routers that connect to one of its networks. A routed mode firewall can be placed at the Layer 3 boundary. In transparent mode, the adaptive security appliance is a Layer 2 device and can be placed at the Layer 2 boundary.

A firewall is positioned at the Layer 2 boundary, in transparent mode, when forwarding of non-IP traffic (Internetwork Packet Exchange [IPX], for example) is required, or when readdressing of a network is not desired.

Cisco security appliances are available in different implementations:

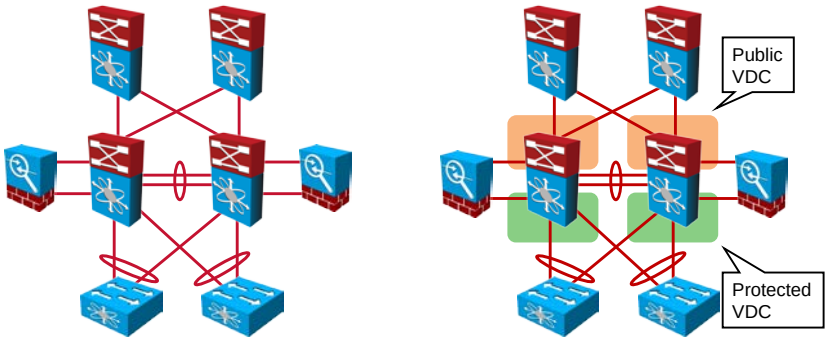
- As a standalone device or as external firewalls like Cisco ASA 5585-X Series Adaptive Security Appliances. They are purpose-built solutions that integrate firewall, unified communications security, VPN, IPSs, and content security services in a unified platform.
- As service modules (integrated) in the service chassis. The most typical deployment is using a Catalyst 6500 chassis and with an ASA-SM installed in the chassis. The ASA-SM supports most features, as found on the adaptive security appliance standalone firewalls, and allows additional flexibility for configuration of the topology because it does not use physical cables.

https://t.me/learningnets

certcollection.net

Designs with Standalone Firewall Appliances

- Typically positioned in the data center aggregation layer.
- Traffic flows from the aggregation switch to the appliance to be inspected.
- Allowed traffic flows toward the aggregation switch, to be further switched southbound to the access layer.
- Need to provision sufficient bandwidth to and from the firewall.



Regular deployment

"VDC sandwich" design

© 2012 Cisco and/or its affiliates. All rights reserved.

One of the most popular designs is placing the firewalls as standalone devices. High-performance adaptive security appliance devices, such as the Cisco ASA 5580-40 and 5585-X can be deployed. These security appliances feature traffic filtering and inspection for 10-Gb bandwidth and more.

The traffic flow is from the aggregation switch to the security appliance where traffic is filtered and inspected, and then back to the aggregation switch. Traffic can use a single link or multiple links (one per direction), depending on the amount of traffic and how congested these links are. Segment the traffic using VLANs—an outside VLAN and an inside VLAN.

Note On the trunk links between the adaptive security appliance and the switch, you also need to carry VLANs that are used for firewall failover.

The example on the left side of the figure does not use virtual device contexts (VDCs) on the Cisco Nexus 7000 aggregation switch, while the example on the right side does. Using VDCs is recommended, so as to additionally separate trusted and untrusted zones on the aggregation switches.

When using the design with VDCs, you need to provision multiple links to the firewall appliance—one from the unsecured VDC and one to the secured VDC.

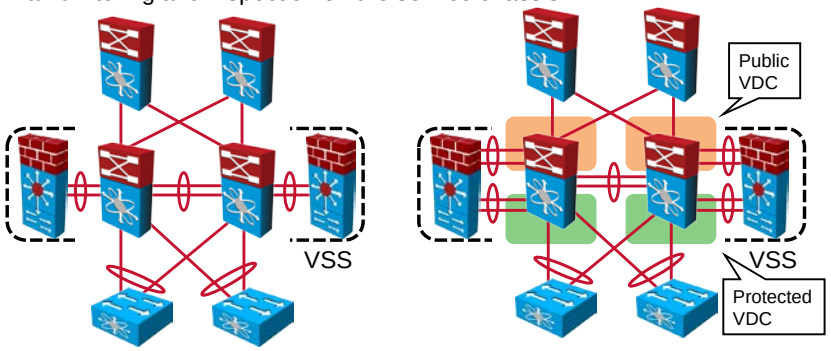
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

Service Chassis Design

- Traffic from the aggregation switch is sent to the service chassis.
- Aggregation switch uses a vPC to both service chassis.
- Service chassis running Cisco Catalyst 6500 VSS and MEC to aggregation switch.
- The aggregation switch can use VDCs.
- Traffic filtering and inspection on the service chassis.



© 2012 Cisco and/or its affiliates. All rights reserved.

A design with a service chassis is used when you need multiple types of IP services in the data center—not only firewalling, but also server load balancing and so on. In the service chassis, you can deploy the ASA-SM firewall and the Cisco ACE30 Module to deliver application services. This way, you have a common platform—the service chassis—to deliver all services.

When using the option with VDCs on the Cisco Nexus 7000 aggregation switches, you must provision links from the public VDC to the service chassis, and from the service chassis to the private VDC. You can use a virtual port channel (vPC) on the Cisco Nexus aggregation switch, and Multichassis EtherChannel (MEC) on the Virtual Switching System (VSS) for the service chassis.

The links to the private VDC can be used for transport of fault-tolerant VLANs for service modules.

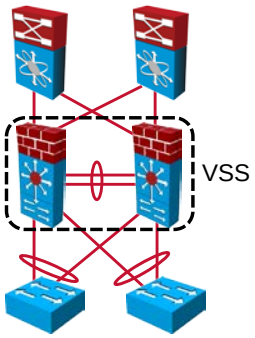
The service chassis can operate as a Layer 3 device or as a Layer 2 switch.

https://t.me/learningnets

certcollection.net

Fully Integrated Services Design

- Services running within the aggregation layer.
- Formed by a Cisco Catalyst 6500 VSS.
- The switch has a single control plane. Firewall modules have two control planes and operate in active/standby or active/active using multiple contexts.
- A VSS interswitch link is used to forward traffic to the service module if the context is active on the switch that is not local.



© 2012 Cisco and/or its affiliates. All rights reserved.

The design in this figure features the Cisco Catalyst 6500 VSS as the aggregation switch, with service modules installed in both chassis.

The switch has a single control plane, but the service modules still operate as standalone devices and use their own failover mechanisms. Service modules are deployed in active/active, multiple context mode, and some security contexts operate in active/standby mode.

It is possible that traffic going from the server to the service module is received on the switch that does not have the service module active for that context. In this case, traffic is forwarded through the link between the physical switches to reach the correct service module.

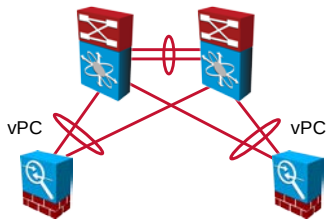
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

Cisco ASA with vPC Attachment

- Better redundancy, faster convergence, and better load distribution
- vPC from the switch to the ASA appliance
- Similar topology as a vPC to the service chassis



© 2012 Cisco and/or its affiliates. All rights reserved.

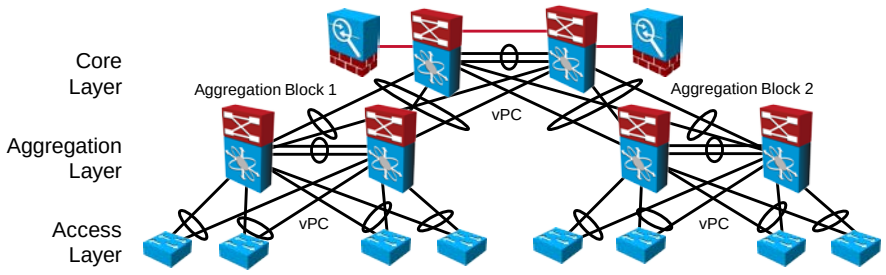
The Cisco ASA adaptive security appliance can take advantage of the vPC connectivity to the aggregation switches. The vPC offers better redundancy, faster convergence, and better load distribution, providing benefits similar to that of the firewall in a service chassis, connected to the aggregation switches using a combination of a vPC and MEC.

https://t.me/learningnets

certcollection.net

Firewall at the Boundary of Layer 2 and Layer 3

- In large, “flat” data center networks with Layer 2 domains stretching to the core layer, firewalls are connected to core switches.
- The core switch provides the Layer 2 or Layer 3 boundary, switch virtual interfaces for default gateway, First Hop Redundancy Protocol (FHRP), and routing upstream to the core layer.
- This is not a very common design.



© 2012 Cisco and/or its affiliates. All rights reserved.

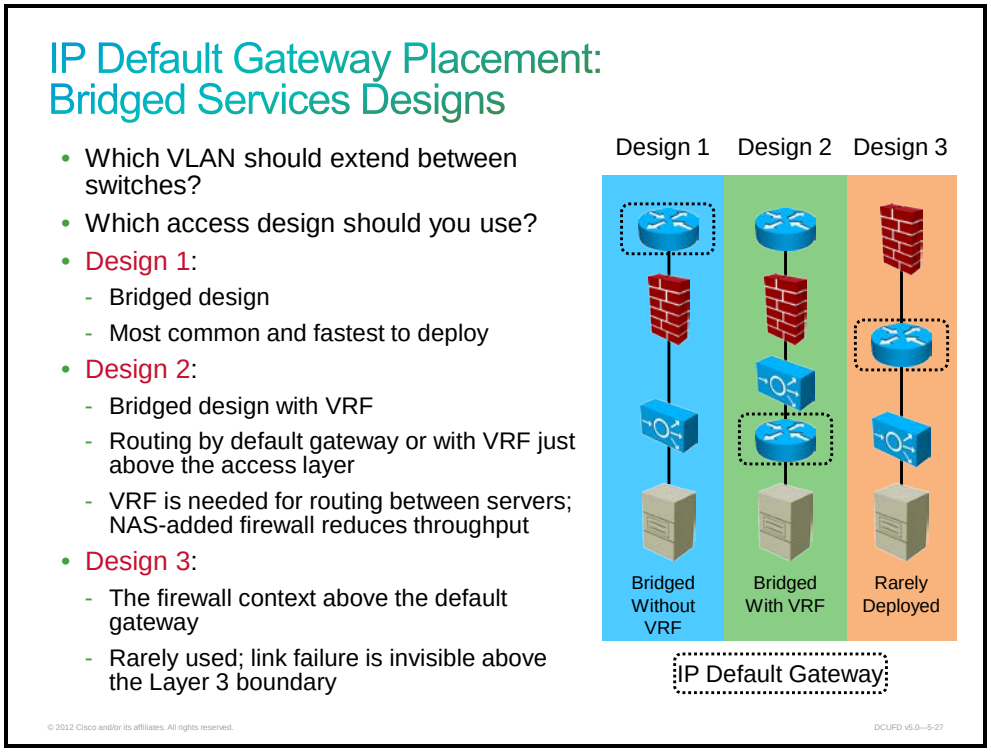
When you have a data center with Layer 2 domains stretched to the core layer, you need to position the firewalling devices at the boundary of Layer 2 and Layer 3, at the data center core. Positioning the firewalls at the aggregation layer would make sense only if they would operate as Layer 2 devices only (in transparent mode) and would not guarantee consistency in firewalling. The core switches aggregate all traffic that is received in and sent out from the data center, and provide a strategic point to place inspection devices.

This way, the firewalls are not burdened by traffic that is switched within the Layer 2 domains, such as intracluster traffic, VMotion, and so on.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net



Cisco ACE and Firewall Service Modules Design Choices in Cisco Catalyst 6500 Series Switches

When deploying these services, there are various choices to make:

- Bridging versus routing toward the access with virtual routing and forwarding (VRF)
- Routed versus transparent mode
- Inline versus one-arm design
- Which VLANs to extend between aggregation switches
- How traffic will flow
- Which access design to use
- How to accomplish server-to-server load balancing

The bridged design leaves all of the routing to the global Multilayer Switch Feature Card (MSFC). The VRF-routed design adds a VRF south of the service modules to perform the access layer routing. This is useful when there is a requirement to route between subnets behind the firewall. For example, network-attached storage (NAS) resources might be located on a different subnet than the web server, but within the same security domain. Forcing these flows through the firewall reduces the overall capacity, without providing any additional security.

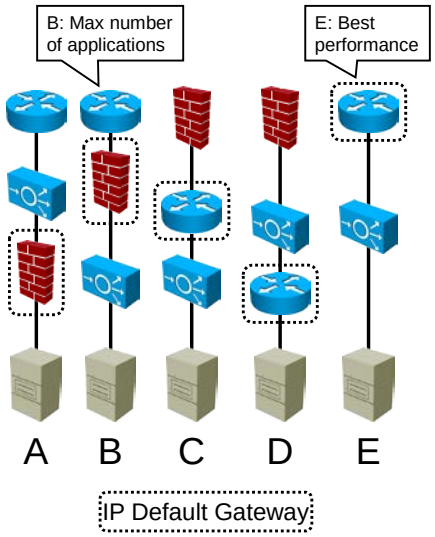
Another option is to place the firewall context above the global MSFC, between the aggregation and core tiers. This approach, however, is undesirable for a number of reasons. STP processes are introduced into the core of the network, the MSFC loses direct visibility to link failures, and the regular changes to Cisco FWSM contexts are potentially disruptive to the entire network. Alternatively, when dedicated VRFs are used to provide routing functionality, the integrity of the core is maintained, while maximum flexibility is provided for access connections. VRFs can also provide a way to manage overlapping address space without the need for Network Address Translation (NAT).

https://t.me/learningnets

certcollection.net

IP Default Gateway Placement: Routed Services Designs

- The default gateway must be selected based on the data center requirements: Cisco ACE Module, firewall, or switch FHRP IP address.
- Different choices offer different benefits:
  - Per-farm independent security
  - Firewall in routed mode
  - Cisco ACE Module in routed or bridged mode
  - Different sizes of Layer 2 and Layer 3 failure domains
- Avoid placing the default gateway on the firewall for high loads of traffic.



© 2012 Cisco and/or its affiliates. All rights reserved.

Service Module Default Gateway Redundancy

The default gateway for a server in a data center can be a router, a firewall, or a load balancer. In the same data center, some server farms use only Layer 2 and Layer 3 functions, some use only load balancing or firewalling, and some use both firewall and load-balancing services. Firewalling and load balancing can be provided by placing these devices in front of the server farm in either transparent mode or bridge mode, or by making the firewall or the load balancer the default gateway for the server farm. The default gateway must be selected based on the data center requirements. Different choices offer different disadvantages and advantages.

Note	For high-bandwidth traffic flows or demanding applications, avoid placing the IP default gateway device on the firewall. This is not a scalable solution.

Example A

When the load balancer is operating in routed mode or bridged mode, the firewall configuration must enable server health-management traffic from the content switching module to the server farm. This adds management and configuration tasks to the design. In this design, the firewall provides the default gateway to the server farm.

Example B

This example shows a server with the load balancer default gateway if it is deployed in routed mode. If a load balancer is deployed in bridged mode, then the default gateway is the firewall. This configuration facilitates the creation of multiple instances of the firewall and Cisco Application Control Engine (ACE) combination for the segregation and load balancing of each of the server farms independently. Placing the load balancer in bridged mode between the server farm and the firewall and configuring the firewall as the default gateway provides the maximum number of application and security services.

https://t.me/learningnets

certcollection.net

Example C

The firewall that faces the core IP network in this example must have routing capabilities for easy integration with the routed network. This makes the capability to independently secure each server farm more complicated, because the packets can be routed from the server farm and back to the server farm by Cisco ACE Module and the firewall without passing the firewall. If Cisco ACE Module is deployed in routed mode, then Cisco ACE Module or the router can be the default gateway. When Cisco ACE Module is deployed in bridged mode, the default gateway is the router.

Example D

The advantage of this design is that the router is the default gateway for the server farm; therefore, the servers can take advantage of Hot Standby Router Protocol (HSRP) tracking, QoS, or DHCP relay functions, which are available only on routers.

Example E

Servers send traffic directly to the router. The default gateway that is configured on the servers is the IP address of the router.

Note Configuring the router as the default gateway provides the best performance.

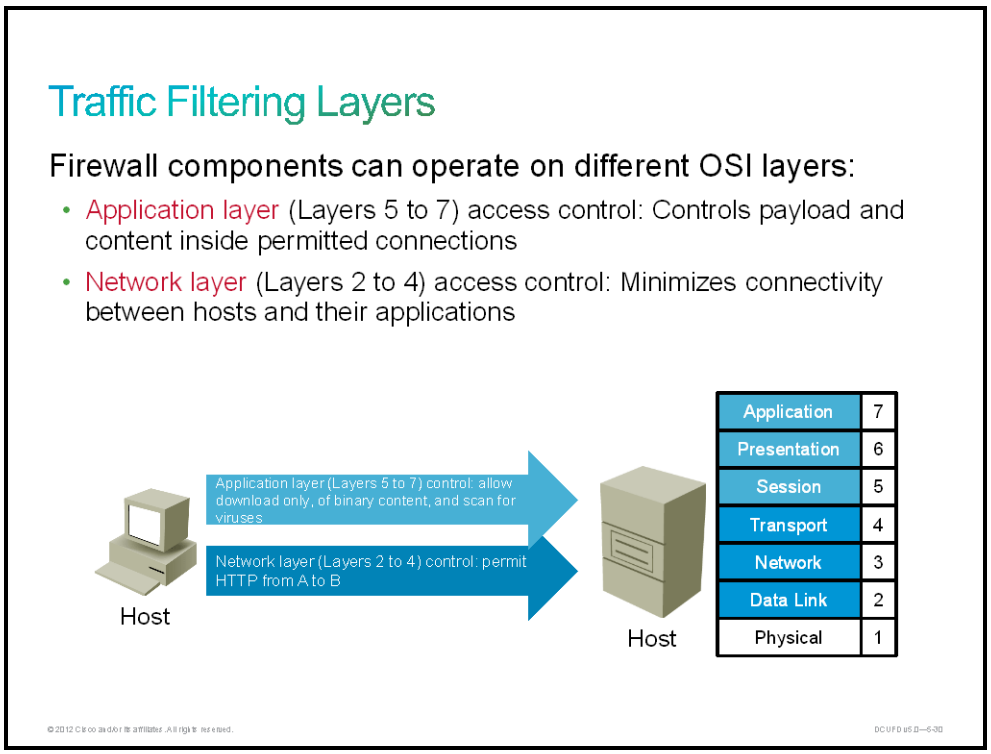
When Cisco FWSM is in transparent mode, the traffic that flows from a server farm on one VLAN to a server on a different VLAN traverses the device twice, because routing occurs on the switch.

https://t.me/learningnets

certcollection.net

Secure Communication on Multiple Layers

This topic describes how to design secure communication on multiple layers.



A network access policy defines which network connectivity is allowed according to the security policy of an organization. Firewall systems enforce network access control on two basic (coarse) layers:

- Network layer access control (OSI Layers 2 to 4) determines which application hosts can intercommunicate using which protocols and applications. An example of network layer access control is a firewall that permits all inside users to open HTTP connections to all servers on the Internet.
- Application layer access control (OSI Layers 5 to 7) determines what a user can do within an application. An example of such access control is a firewall that can verify the session adherence to the standard application layer protocol, allows users to view web pages but prohibits them from posting data to untrusted servers, blocks viruses in email messages by examining application layer content, or permits only well-formed XML messages inside a web services application.

 © 2012 Cisco Systems, Inc.

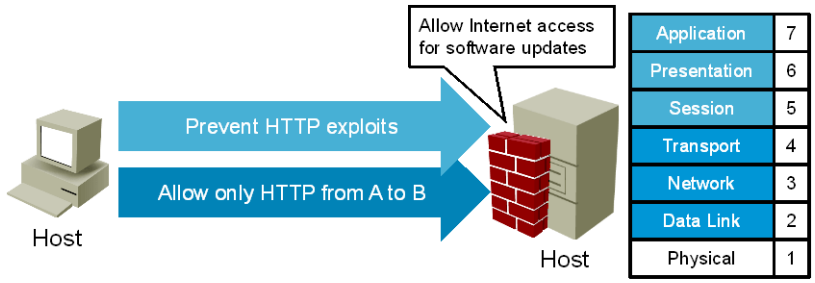
https://t.me/learningnets

certcollection.net

Filtering Approaches

Firewall components can operate in different access control modes:

- Restrictive access control: Everything not explicitly allowed is prohibited.
- Permissive access control: Everything not explicitly prohibited is allowed.



A firewall system can implement access control using one or both of two approaches:

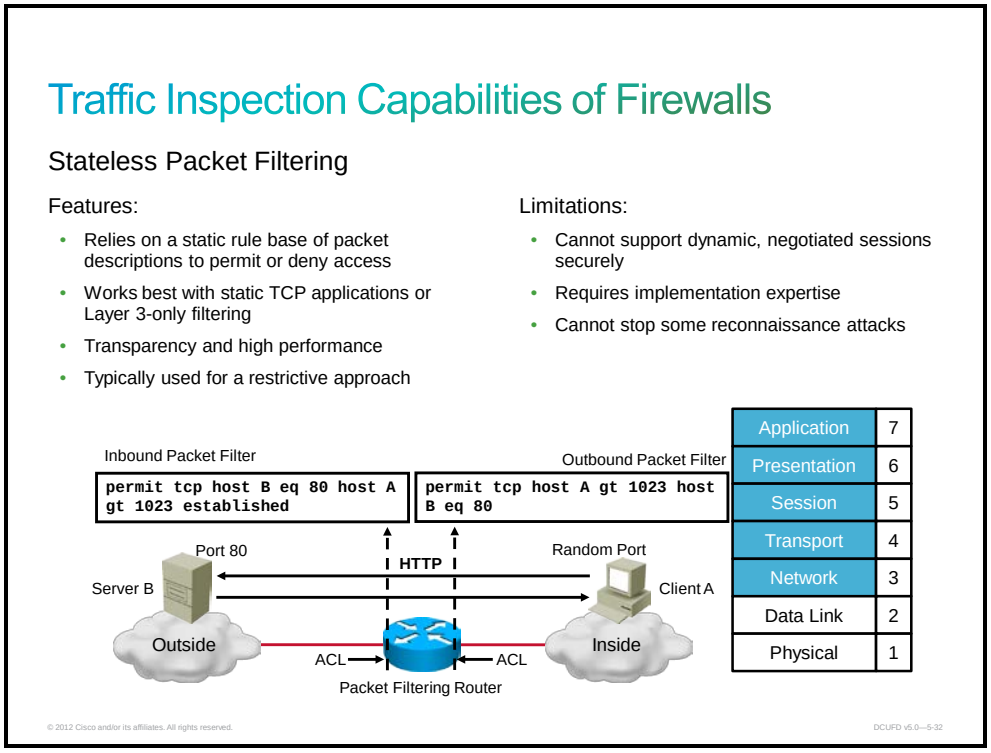
- The restrictive (or proactive) approach, in which the firewall, by default, denies all communication and only allows the aspects of communication that are explicitly permitted. Examples of this approach are stateful packet filtering devices that only allow specific hosts and applications to pass, or a mail proxy that would only allow text-based file attachments.
- The permissive (or reactive) approach, in which the firewall, by default, permits all communication and only blocks the aspects of communication that it considers malicious based on its attack signature database. Examples of this approach are network intrusion prevention systems and network antivirus.

Restrictive and permissive controls often work together. For example, only HTTP traffic can be allowed through a firewall. However, inside HTTP, all known HTTP exploits are prohibited.

On the firewall system, you should allow Internet access for data center servers for software updates.

https://t.me/learningnets

certcollection.net



There are several mainstream traffic filtering technologies that are used in modern firewall systems. The most basic is stateless packet filtering.

Stateless packet filtering is one of the oldest and most widely used network access control technologies and is usually employed by an OSI Layer 3 device, such as a network router. Stateless packet filters use a statically defined set of rules that independently (statelessly—without regard to previous or future packets) examine each packet header or payload to permit or deny its forwarding across the device. Stateless packet filtering usually examines protocol headers of network and transport OSI layers but can be extended to the application layer by examining packet payloads, and even parsing packets to decode their application layer protocols for simpler access rule configuration.

Stateless packet filters have the following features:

- They work best with simple TCP-based applications (which do not negotiate dynamic ports) or when filtering is performed strictly on Layer 3 of the OSI model (for example, in manual ingress or egress antispoofing filters).
- They are cost effective to deploy because they are generally present in existing network software and do not require any software changes.
- Stateless packet filters are generally efficient and high performing, and are often accelerated in hardware.

Stateless packet filters also have the following limitations:

- Stateless packet filters cannot permit applications with dynamically negotiated transport layer sessions (that is, dynamic ports) without the administrator creating suboptimal access rules, which permit unwanted traffic as well.
- The correctness of rules relies on the ability of the designer to set up the rules according to his knowledge of applications and protocols.
- Usually, an attacker can still send some reconnaissance traffic through a stateless packet filter due to its stateless nature.

 © 2012 Cisco Systems, Inc.

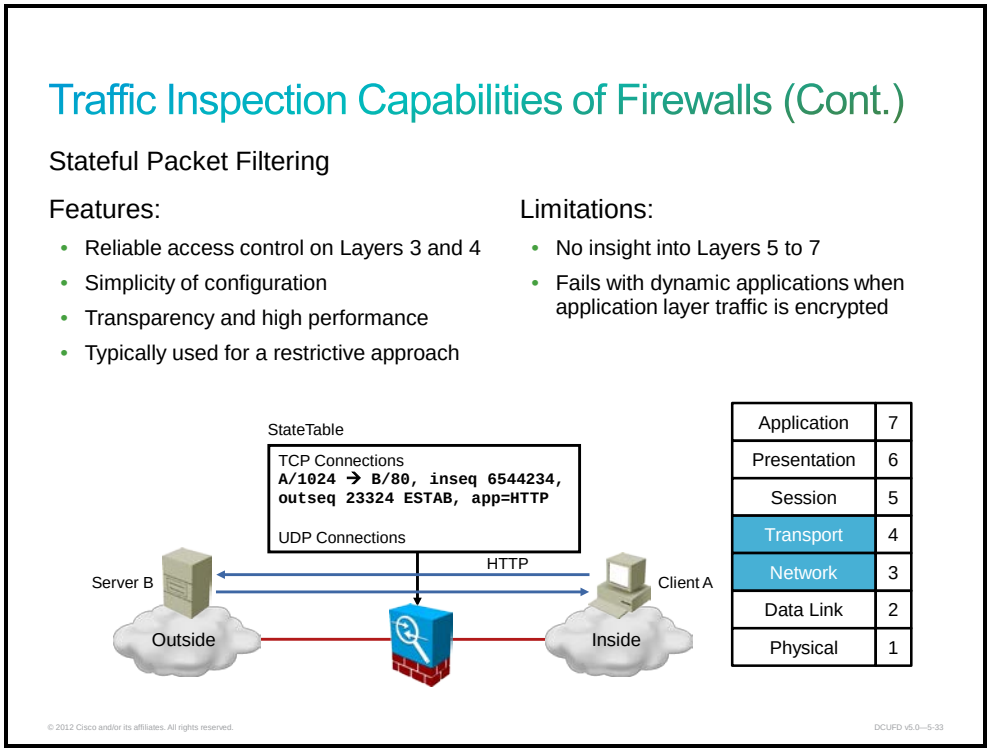
https://t.me/learningnets

certcollection.net

Examples

Examples of OSI Layer 3 and Layer 4 stateless packet filters are interface ACLs and Cisco Catalyst VLAN ACLs (VACL). Such ACLs can filter on network addresses, protocols, ports, and specific per-protocol flags, such as TCP flags, IP options, or Internet Control Message Protocol (ICMP) types and codes.

An example of OSI Layers 3 to 7 stateless packet filters is the Cisco IOS Flexible Packet Matching (FPM) functionality, which is a superset of classic ACLs. FPM allows for decoding of OSI Layers 3 to 7 protocols and matching based on packet payload.



Stateful packet filtering is an application-aware method of packet filtering that works on the connection—or flow—level, with occasional peeks into the application layer of an application.

Stateful packet filters maintain a state table to keep track of all active sessions that cross the firewall. A state table, which is an internal data structure of a stateful packet filter, tracks all OSI Layer 4 sessions and inspects all packets that pass through the device. Based on its memory of previous packets in a session, a stateful packet filter can expect what kind of traffic should arrive soon from communicating hosts. If the packets have the expected properties that were predicted by the state table, they are forwarded. The state table changes dynamically as a result of traffic flow.

Stateful packet filters are also application-aware through additional, deeper inspection of transit traffic, which is required to process dynamic applications. Dynamic applications typically open an initial connection on a well-known port, and then negotiate additional OSI Layer 4 connections through the initial session. Stateful packet filters support these dynamic applications by analyzing the contents of the initial session and parsing the application protocol just enough to learn about the additional negotiated channels. A stateful packet filter typically assumes that if the initial connection was permitted, any additional transport layer connections of that application should be permitted as well.

https://t.me/learningnets

certcollection.net

Stateful packet filters have the following features:

- They provide a reliable method to filter network traffic on OSI Layers 3 and 4 between security domains.
- They are simple to configure, because the firewall operator does not need to be aware of how the application is using the network. The stateful intelligence processes any exceptional behavior of dynamic applications.
- They are transparent to hosts and have high performance. Some stateful packet filters even include QoS features, such as interface queuing and policing.

However, pure stateful packet filters engines do not provide reliable and extensive application layer filtering or protocol verification mechanisms, and they fail to pass legitimate traffic of dynamic applications, if their application layer traffic is encrypted, because they cannot observe protocol negotiations.

Examples

Examples of devices that can employ stateful packet filtering include the Cisco ASA adaptive security appliance, the Cisco Firewall Services Module, and the Cisco IOS Software zone-based policy firewall. In the figure, there is one TCP connection that is established over the adaptive security appliance. A connection is established from host A to host B with source port 1024 and destination port 80. The connection is used by the HTTP application.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

Traffic Inspection Capabilities of Firewalls (Cont.)

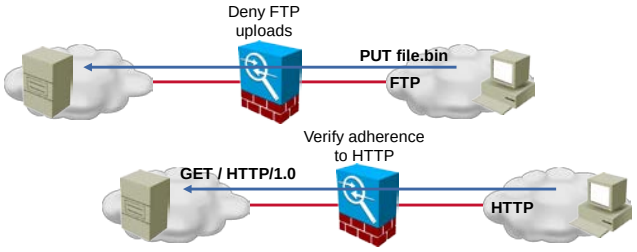
Stateful Packet Filtering with Application Inspection Control

Features:

- Reliable access control on Layers 3 to 7
- Simplicity of configuration
- Transparent, medium-performance operation
- Typically used for a restrictive approach

Limitations:

- Application inspection control impacts performance
- Limited buffering capability for deep content analysis



Application	7
Presentation	6
Session	5
Transport	4
Network	3
Data Link	2
Physical	1

© 2012 Cisco and/or its affiliates. All rights reserved.

Many users of stateful packet filtering technology have increasingly demanded higher application layer awareness in their stateful packet filtering-based firewalls. Most vendors responded by improving application layer analysis on their pure stateful packet filtering devices by enhancing the traffic analysis engine with the following services:

- In-memory reassembly of Layer 4 (TCP, UDP) sessions to obtain a sequential stream over which the application layer inspection engine can reliably parse the application layer protocol
- Application layer protocol decoding, to allow for (restrictive or permissive) filtering inside the protocol and its content
- Application layer protocol verification, in which the engine drops application layer protocol units that do not conform to the protocol standard

This additional functionality is called application inspection control or deep packet inspection and may considerably affect performance, if enabled. Application inspection control-enabled stateful packet filters have the same features and limitations as normal stateful packet filters, with these additional features:

- The ability to control access on OSI Layers 3 to 7
- Protocol verification on OSI Layers 3 to 7

Performance tends to be lower compared to stateful packet filters, depending on the amount of application layer inspection enabled inside the application inspection control engine. Additionally, because stateful packet filters with application inspection control usually do not have a hard disk or extreme amounts of RAM to perform the buffering that is required for deep content analysis (such as file-based antivirus), their application filtering is usually limited to application protocol headers, without detailed data inspection.

https://t.me/learningnets

certcollection.net

Examples

Examples of application inspection control-enabled stateful packet filters are the Cisco ASA adaptive security appliance (from version 7.0 on), the Cisco Firewall Services Module (from version 3.2 on), and the Cisco IOS zone-based policy firewall (the Application Firewall feature, available from Cisco IOS Software Release 12.3(14)T on).

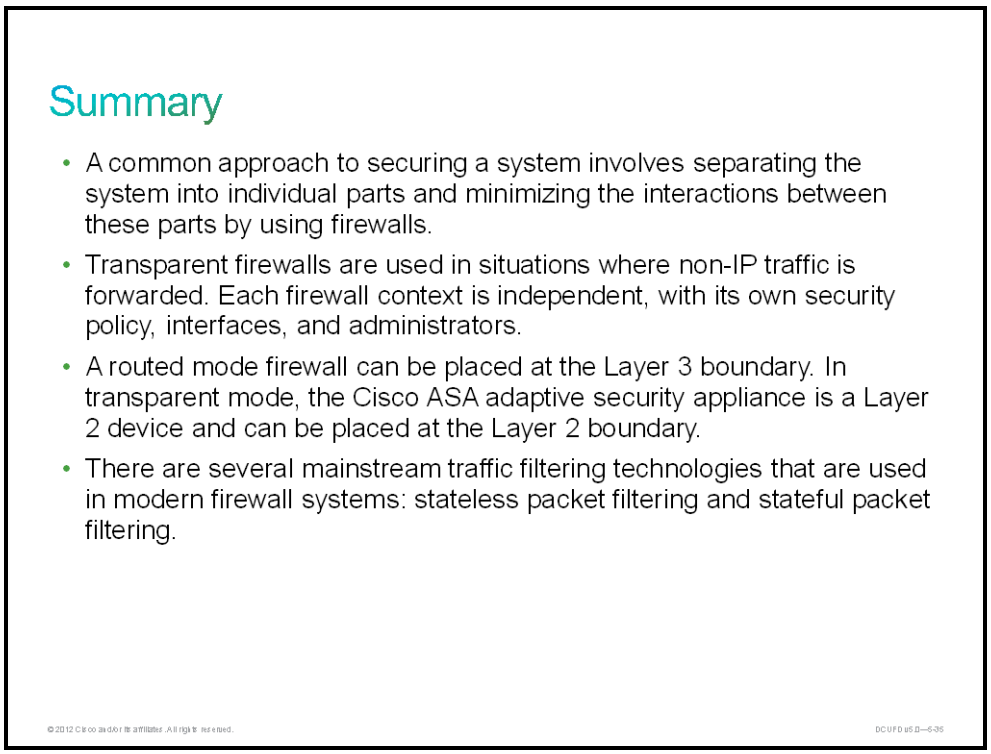
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

Summary

This topic summarizes the primary points that were discussed in this lesson.



https://t.me/learningnets

certcollection.net

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

Lesson 2

Designing Link Security Technologies and Device Hardening

Overview

The network infrastructure is one of the foundation elements of enterprise IT infrastructures and is a critical business asset of telecommunications service providers. A compromise of the network infrastructure or its individual components can have far-reaching consequences that can significantly increase risk to enterprise and service provider business processes. This lesson describes the threats against the network infrastructure, describes basic classes of security controls that can reduce the risk of these threats, and provides an overview of the Cisco advanced security features that you can use to reduce these risks in your data center environment.

Objectives

Upon completing this lesson, you will be able to design secure data center networks on the network and device level. This ability includes being able to meet these objectives:

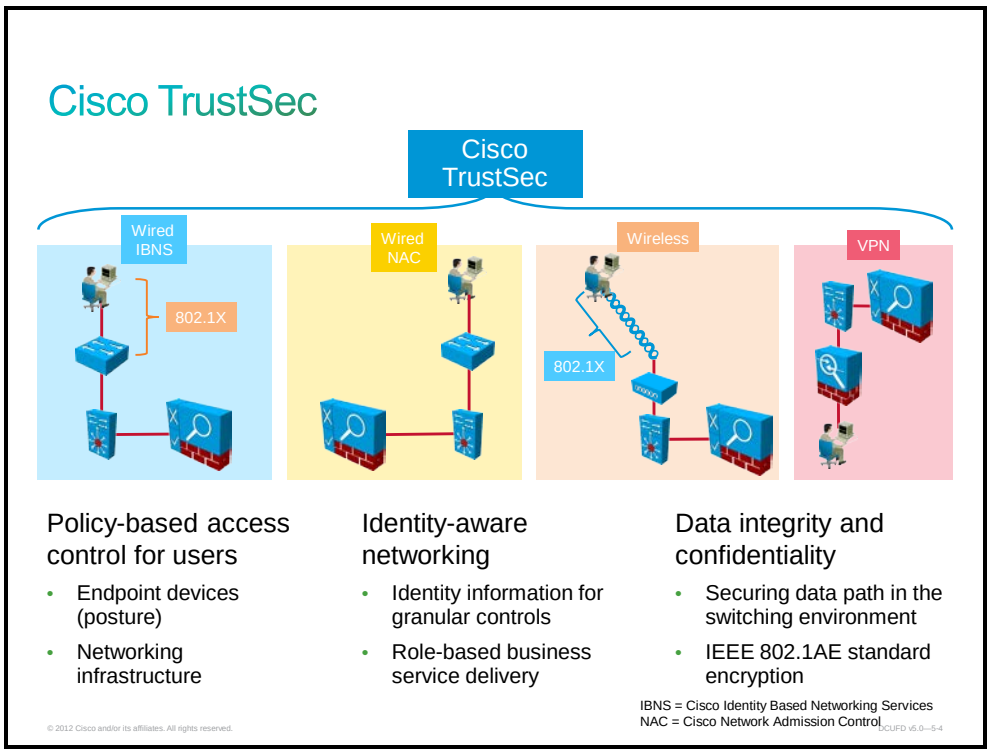
- Identify design requirements for Cisco TrustSec
- Describe device-hardening technologies
- Design secure management networks

https://t.me/learningnets

certcollection.net

Link Security

This topic describes how to identify design requirements for Cisco TrustSec.



The traditional desktop is no longer relevant. Customer networks must support all kinds of devices, such as personal mobile devices, or existing devices with no users connected to them. With so many devices connecting to the enterprise network, customers need a solution that helps them to ensure that they are meeting their security policies when these devices use the network.

From a data center standpoint, applications are progressing. Customers used to think about securing their applications using access control lists (ACLs). In a virtualized data center, however, applications move between data centers via virtual machines (VMs). Customers must think differently about how to secure their networks. As their applications are moving through the data center, they need an infrastructure that is as dynamic as the applications.

Cisco TrustSec is an intelligent access control solution. With minimal effort, Cisco TrustSec mitigates security risks by providing comprehensive visibility into who and what is connecting across the entire network infrastructure, as well as exceptional control over what and where they can go.

Whether you need to support employees who are bringing personal devices to work or you want to secure access to your data center resources, Cisco TrustSec provides a policy-based platform that offers integrated posture, profiling, and guest services to make context-aware access control decisions. Cisco TrustSec builds on an existing identity-aware infrastructure by enforcing these policies in a scalable manner. Additionally, Cisco TrustSec helps to ensure complete data confidentiality by providing ubiquitous encryption between network devices. A unique, single-policy platform that uses your existing infrastructure helps ensure highly effective management.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

Cisco TrustSec offers numerous features:

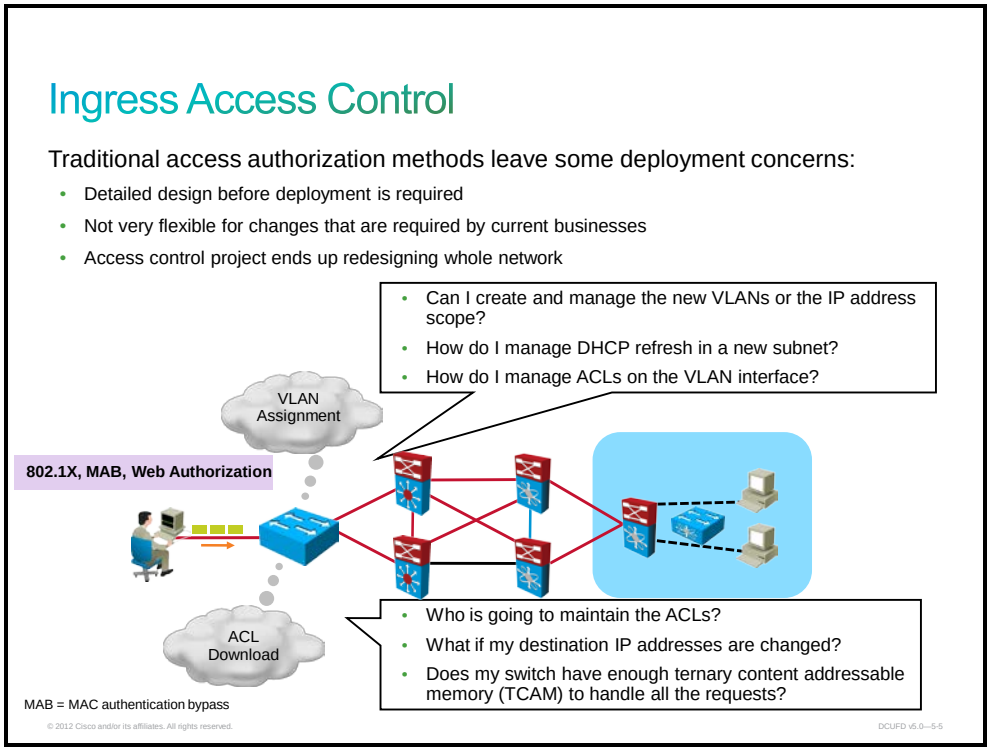
- Identify who is accessing to your network
- Determine how this access is attempted
- Locate where this person is trying to access
- Evaluate what privilege this person has

Cisco TrustSec provides numerous results:

- Admission to the network
- Scope of resources this person can access
- Level of services this person can access
- Record of network usage

https://t.me/learningnets

certcollection.net



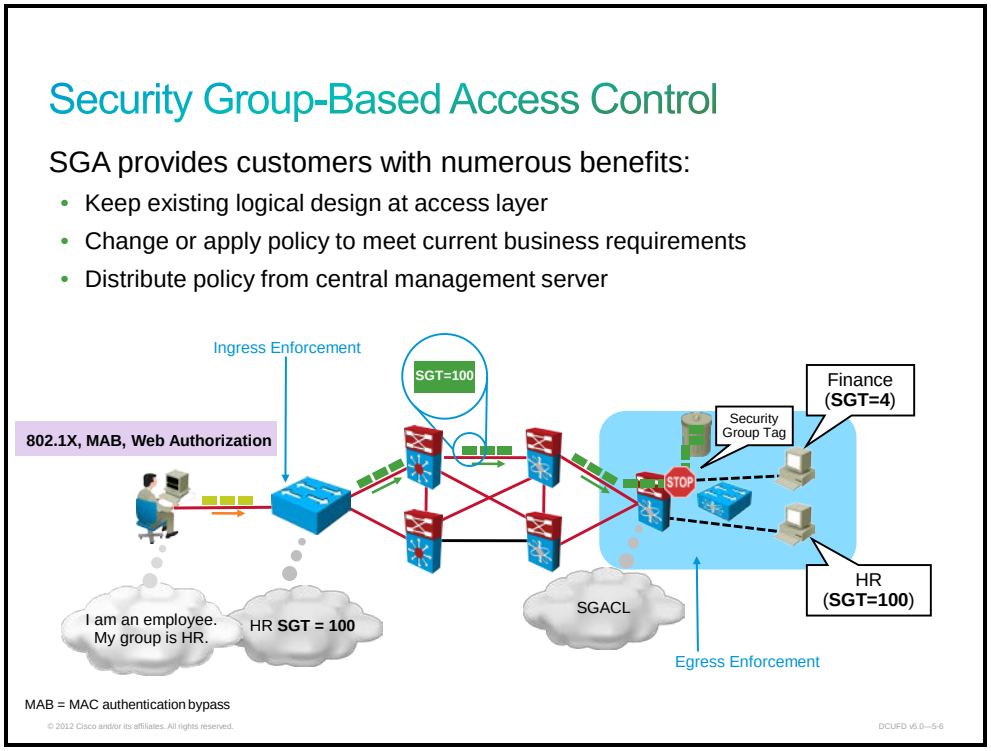
Three important functional areas of Cisco TrustSec are visibility, control, and management.

- **Comprehensive Visibility:** The differentiated identity features, next-generation network-based device sensors, and active endpoint scanning in Cisco TrustSec provide contextualized visibility of the "who, how, what, and when" for users and devices that are accessing the network, whether through wired, wireless, or remote connections. Because Cisco TrustSec provides comprehensive visibility into the broadest range of devices (whether smart phones, tablets, PCs, or even gaming devices), it lays a strong foundation for a Bring Your Own Device (BYOD) solution.
- **Exceptional Control:** A centralized policy and enforcement platform enables coordinated policy creation and consistent context-based policy enforcement across the entire corporate infrastructure. Noncompliant devices can be quarantined, remediated, or given restricted access with scalable and flexible next-generation enforcement mechanisms using existing identity-aware infrastructure. Cisco TrustSec helps to ensure secure access for devices via automated endpoint security configuration for the most common PC and mobile platforms.
- **Effective Management:** Cisco TrustSec combines authentication, authorization, and accounting (AAA), posture, profiler, and guest management functions in a single, unified appliance, which leads to simplified deployments and a single point of management. These features provide lower total cost of ownership (TCO).

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net



After network users and devices are authenticated and confirmed to comply with the security policy of an organization, they are allowed network access. Their subsequent resource and service entitlement is accomplished by the authorization process. Cisco TrustSec supports multiple authorization methods, including ACLs, VLANs, and Security Group Access (SGA).

These choices help organizations design their security architecture and services offerings with maximum flexibility and effectiveness. Downloadable, per-session ACLs and dynamic VLAN assignments can be implemented at the ingress point where users and devices gain their initial entry to the network. In addition, SGA allows user identity information to be captured and tagged with each data packet. A Security Group Access Control List (SGACL) can be implemented at an egress point where a network resource (such as a file server) is located. SGA-based access control allows organizations to keep the existing logical design at the access layer and, with flexible policies and services, to meet different business requirements without having to redeploy the security controls.

This figure shows how the role-based tag works:

Step 1 A user (or device) logs into network via IEEE 802.1X.

Step 2 The Cisco Identity Service Engine (ISE) server is configured to send a tag in the authorization result, based on the role of the user or device.

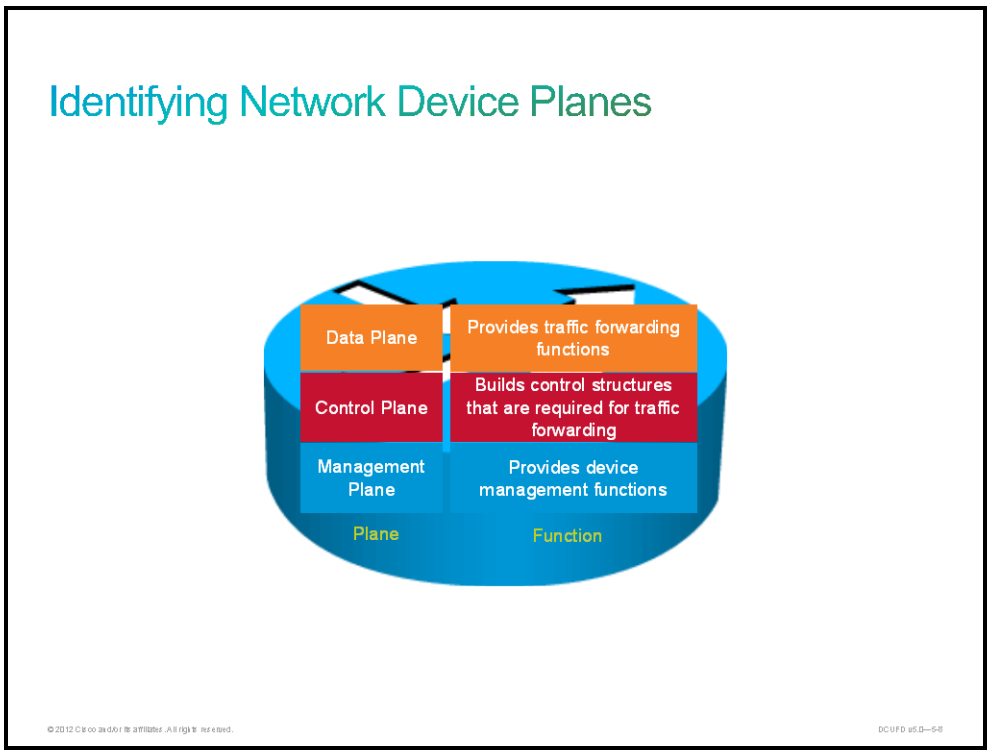
Step 3 The switch applies this tag to the user traffic.

https://t.me/learningnets

certcollection.net

Device Hardening

This topic describes device-hardening technologies.



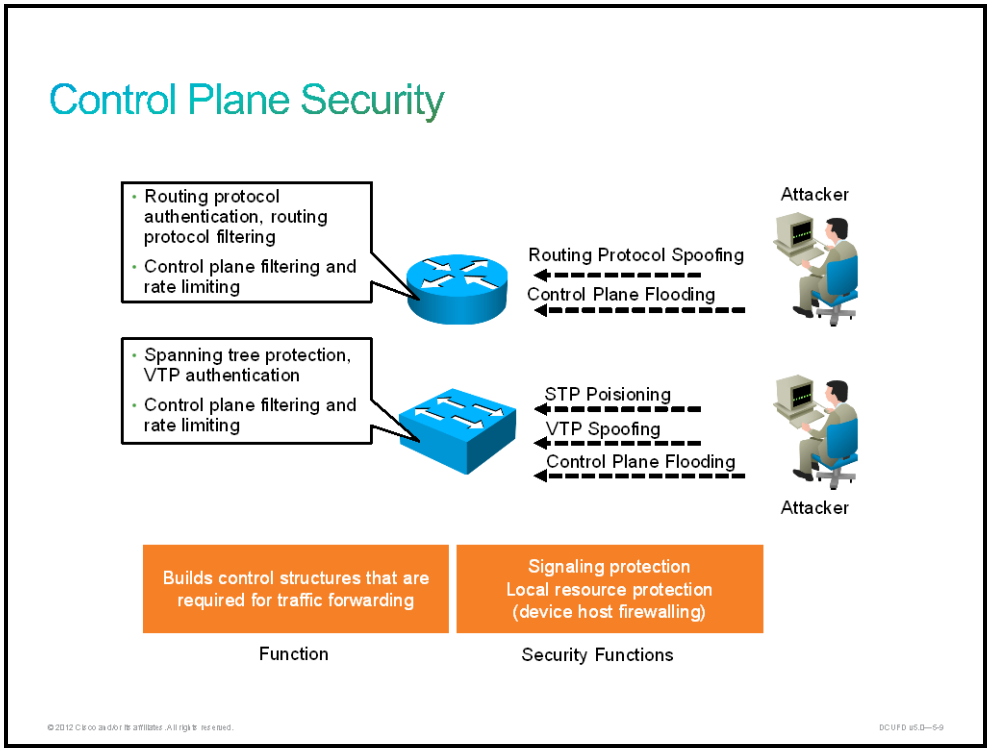
It is often beneficial to think of network devices in three separate contexts, as identified by their functionality planes. The functionality of a network device is therefore typically segmented into three planes of operation, each with a clearly identified objective:

- **Management plane:** The management plane provides the device with all functions that administrators need to provision the configuration and monitor the operation of the device.
- **Control plane:** The control plane allows the device to build all of the required control structures (such as routing table, forwarding table, and MAC address table) that will allow the data plane to operate correctly.
- **Data plane:** The data plane allows the device to forward network traffic and apply services (such as security, quality of service [QoS], accounting, and optimization) to it as it is forwarded.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

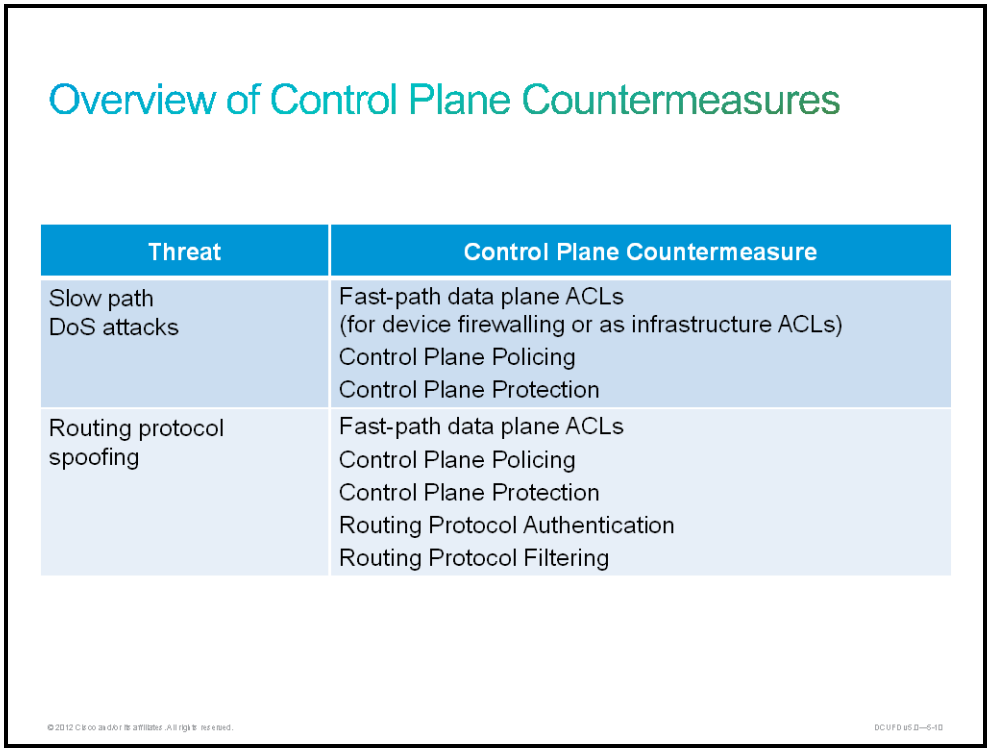


The control plane of a network device can provide the following security-related features to protect the network device against compromise:

- Signaling protection features, which prevent unauthorized entities from influencing traffic-forwarding control structures and, therefore, the traffic-forwarding process itself. The control plane of a device should authenticate signaling-protocol (Spanning Tree Protocol [STP], VLAN Trunking Protocol [VTP], or a routing protocol) information and possibly filter it before passing it on to other signaling partners.
- Methods for protecting the control plane processes of a device against access and flooding from untrusted entities. The control plane of a device should be able to filter or rate-limit packets that are destined for the control plane processes of the device. This filtering or packet rate-limiting is done to both minimize device exposure and to impose a strict limit on CPU and memory resources that can be consumed by control plane traffic.

https://t.me/learningnets

certcollection.net



Because the CPU is shared among the three functions (control plane, management plane, slow data path), excessive traffic to one of these three functions can, by default, overwhelm the entire CPU and influence the behavior of the other two functions. This can lead to flooding attacks, in which the attacker can disable these three functions by sending a high rate of packets to the CPU. There are multiple possible countermeasures that guard against this threat:

- The use of device fast-path data plane ACLs (usually these are interface ACLs of routers and switches) to deny most traffic before it is dispatched into the slow path to the CPU of the router. Because these ACLs are very efficient and are often implemented in hardware in Open Systems Interconnection (OSI) Layer 3 switches and high-end routers, they can drop most malicious traffic without any effect to the CPU. Instead of implementing these ACLs on every device, you can deploy the ACLs at the edge of your network (that is, infrastructure ACLs) to prevent endpoints from injecting traffic that would be forwarded to device CPUs. Be aware, however, that ACLs must be configured with appropriate destination addresses, which may not be scalable if the devices use many IP addresses (interfaces). If these addresses are not contiguous, they can make ACLs difficult to manage.
- The use of specific, on-device protection methods that can filter or rate-limit traffic to the CPU, while leaving fast-path transit traffic untouched. Control Plane Policing (CoPP) and Control Plane Protection are two such features.

Another significant threat to routed control planes is the injection of malicious routing information. Attackers can use malicious routing information to redirect or black-hole sensitive traffic and therefore violate its confidentiality, integrity, or ability to perform a denial-of-service (DoS) attack. These risks can be mitigated by the following:

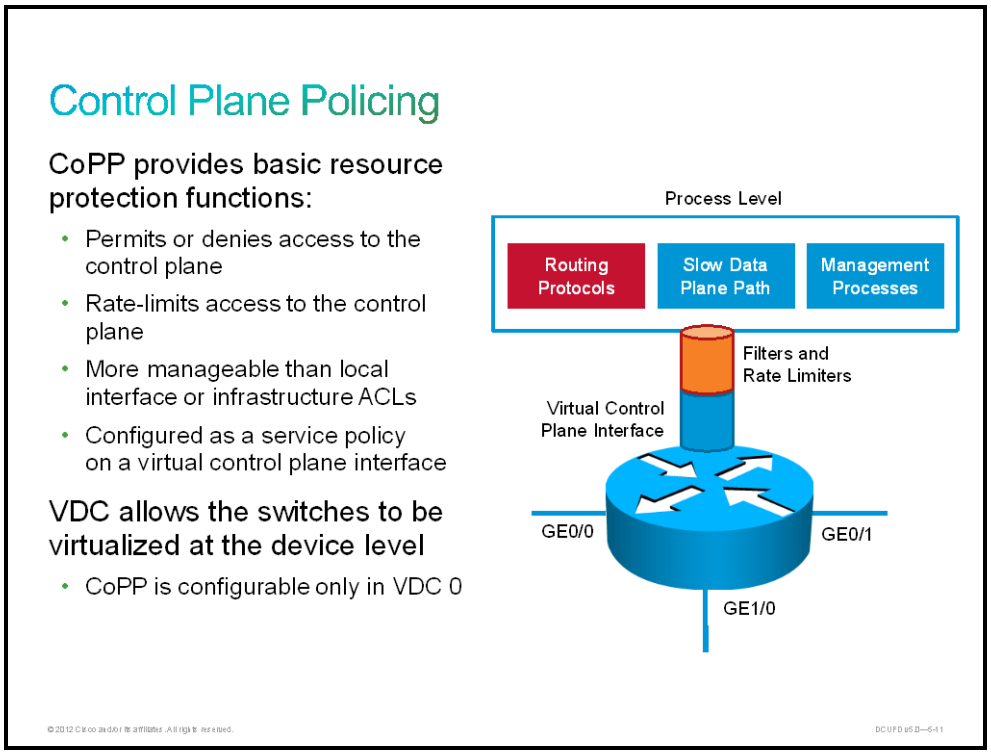
- Use fast-path data plane ACLs to limit who can send routing protocol information to network devices. This solution might not be scalable because of many interfaces and discontiguous IP addressing of network links.
- Use the CoPP and Control Plane Protection features to locally limit the authorized routing protocol peers by their IP address.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

- Use routing protocol authentication, in which a cryptographic integrity and authenticity proof is embedded within each routing protocol message that prevents routing adjacency and routing update spoofing.
- Use routing protocol filtering, which prevents injection of malicious routing information from known, authenticated peers.



CoPP uses early rate limiting and drops traffic that is destined for the CPU of the network device by applying QoS policies to a virtual aggregate CPU-bound queue, called the “control plane interface.” This queue receives all aggregated traffic that is destined for the control plane (which includes the routing protocols), the management plane (management processes), and the slow data plane path traffic of the network device.

CoPP can granularly permit, drop, or rate-limit traffic to the CPU using a Modular QoS CLI (MQC) interface. Because CoPP aggregates all traffic that is forwarded to the CPU of the network device, it is independent of interfaces. This independence allows a central configuration mechanism to protect the network device CPU resources of the process layer of a device.

Cisco Nexus Operating System (NX-OS) supports virtual device contexts (VDCs), which allows the switches to be virtualized at the device level. Each configured VDC presents itself as a unique device to connected users under that physical switch. The VDC runs as a separate logical entity within the switch, maintaining its own unique set of running software processes, having its own configuration, and being managed by a separate administrator.

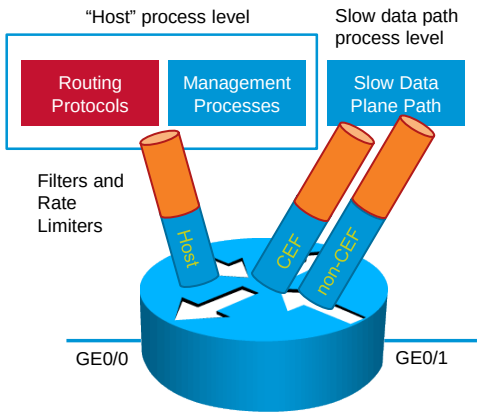
https://t.me/learningnets

certcollection.net

Control Plane Protection

Control Plane Protection provides flexible resource protection and device firewall functions:

- Creates multiple queues to the process level
- Automatically separates host-terminated and transit slow data plane traffic
- Configured as a service policy on a virtual control plane subinterface



© 2012 Cisco and/or its affiliates. All rights reserved.

Control Plane Protection extends the CoPP functionality by automatically classifying all CPU-bound traffic into three queues (subinterfaces) under the aggregate "control plane" interface. Each subinterface receives and processes a specific type of CPU-bound traffic, and each subinterface has a separate traffic policy that is attached to it, making limit configuration much easier.

Control Plane Protection is preferred over CoPP, if it is available on a device, because of its automatic preclassification of traffic into the three subinterfaces, with the ability to separate locally terminated control plane and management plane traffic from transit slow-path data plane traffic.

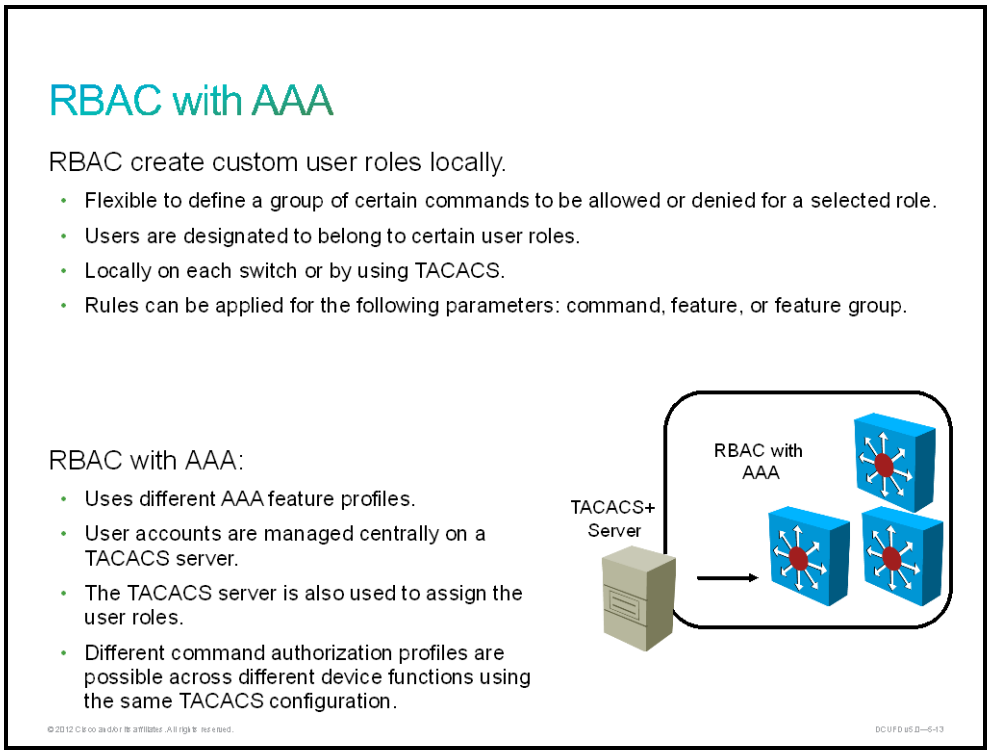
These are the three control plane subinterfaces that are automatically created by Control Plane Protection:

- **The host subinterface:** This interface receives all IP traffic that is directly destined for one of the router interfaces, aggregating the control plane and management plane traffic. Host IP traffic examples include management traffic or routing protocol traffic, such as Secure Shell (SSH), Simple Network Management Protocol (SNMP), Border Gateway Protocol (BGP), Open Shortest Path First (OSPF), and Enhanced Interior Gateway Routing Protocol (EIGRP).
- **The Cisco Express Forwarding exception subinterface:** This subinterface receives all CPU-bound traffic that is redirected to the slow-path data plane. This is a result of the inability to use the fast-path (interrupt-level or hardware-assisted) Cisco Express Forwarding routines to forward the packet. Traffic requires more detailed processing on the process level.
- **The transit (not Cisco Express Forwarding) subinterface:** This subinterface receives all CPU-bound traffic that is redirected to the slow-path data plane. This is a result of the inability to use the fast-path express forwarding routines that are not from Cisco to forward the packet (that is, when Cisco Express Forwarding is not configured on the input interface). On Cisco devices running recent software releases, this typically only includes process-switching traffic.

https://t.me/learningnets

certcollection.net

The transit and Cisco Express Forwarding exception subinterfaces process all slow-path data plane forwarding traffic together.



Cisco Nexus switches can use user roles, which is a local command authorization method. There are various default system user roles.

Role-based access control (RBAC) refers to the ability to create custom user roles locally on a Cisco Nexus switch. This gives the administrator the flexibility to define a group of commands to be allowed or denied for a selected role. Users can then be designated to belong to user roles. This designation can either be done locally on each switch or by using TACACS.

The AAA function and the user roles are mutually exclusive, because the AAA feature overrides the permissions that are allowed with user roles. Using RBAC with the AAA feature (not authorization) offers some interesting options, depending on the network design and requirements.

Custom user roles are defined by giving the role a name and by creating rules within the role. Each rule has a number, to decide the order in which the rules are applied. Rules are applied in descending order. Rule 3 is applied before Rule 2, which is applied before Rule 1. This means that a rule with a higher number overrides a rule with a lower number. Each role can have up to 256 rules configured. All the rules combined within a role determine what operations the role allows the associated user to perform.

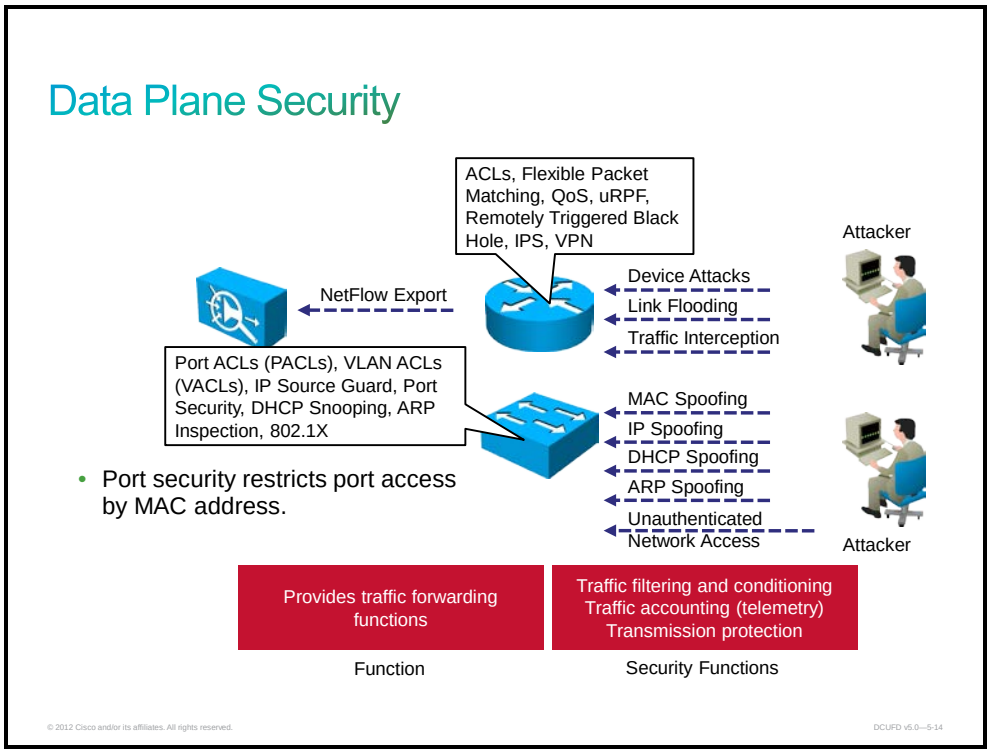
Rules can be applied for the following parameters:

- **Command:** A command or group of commands defined in a regular expression.
- **Feature:** Commands that apply to a function provided by the Cisco Nexus switch.
- **Feature group:** Default or user-defined group of features.

https://t.me/learningnets

certcollection.net

Using RBAC with AAA instead of relying on local usernames, or using different AAA profiles, makes way for favorable designs is certain networks. All user accounts are managed centrally on a TACACS server. The TACACS server is also used to assign the Cisco Nexus user roles. If the TACACS assigned user roles match the local user roles, different command authorization profiles are possible across different device functions using the same TACACS configuration. If the TACACS assigned user role does not match any local user role, the default “Network-Operator” role is applied.



The data plane of a network device provides various security-related features to protect the network device and network endpoints against compromise:

- Traffic filtering features, which can prevent identity theft (for example, features that can address MAC, IP, DHCP, and Address Resolution Protocol [ARP] spoofing attacks), limit access to network devices, or prevent attacks against network-connected endpoints
- Traffic conditioning QoS features, which can control and enforce proper network link use
- Traffic accounting features to enable local incident analysis and the export of network telemetry to centralized analysis systems, to detect malicious activity and provide an audit trail of network activity for incident investigation
- Transmission protection features that cryptographically encapsulate network traffic for transport over untrusted networks

Port security is a Layer 2 traffic control feature on Cisco Catalyst switches. It enables an administrator to configure individual switch ports to allow only a specified number of source MAC addresses to enter the port. Its primary use is to deter the addition, by users, of "dumb" switches to illegally extend the reach of the network so that two or three users can share a single access port.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

Overview of Switched Infrastructure Attack Countermeasures

Attack Type	Countermeasure
VLAN hopping	Static access ports Disable DTP Avoid trunk native VLAN on access ports
STP spoofing	BPDU guard, root guard
MAC spoofing	Port security **or** static MAC entries
CAM flooding	Port security with MAC limit
DHCP server spoofing	DHCP snooping
DHCP starvation	Limiting the maximum number of MAC addresses per port or a DHCP rate limiting
ARP spoofing	ARP inspection **or** static ARP on endpoints and Layer 3 network devices
IP spoofing	IP Source Guard **or** static port ACLs

© 2012 Cisco and/or its affiliates. All rights reserved.

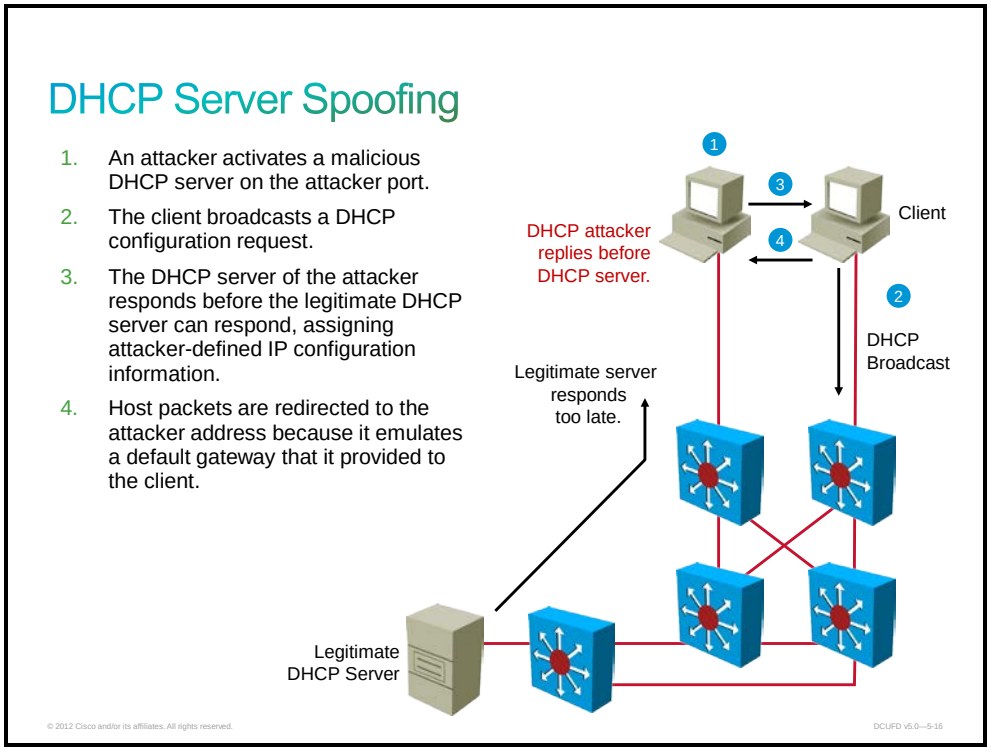
The table in this figure presents common attacks that are mounted inside switched infrastructures, and the security countermeasures that are offered by switched infrastructure devices:

- **VLAN hopping attacks:** The attacker attempts to inject frames that allow the attacker to access VLANs that should not be accessible on a particular port. Proper configuration of static access ports, disabling of the Dynamic Trunking Protocol (DTP), and avoidance of trunk native VLANs on access ports mitigate this common vulnerability.
- **STP spoofing:** The attacker attempts to influence STP operation and divert traffic, or black-hole traffic from the access layer. Effective countermeasures include the use of the bridge protocol data unit (BPDU) guard and root guard features on network switches.
- **MAC address spoofing:** The attacker attempts to steal the identity of endpoints by diverting or black-holing traffic that is destined to them by using a spoofed MAC address and poisoning switch forwarding tables with it. Effective network defenses include the port security feature or static content-addressable memory (CAM) table entries.
- **CAM table flooding:** The attacker attempts to overflow the switch CAM table and cause flooding of sensitive traffic to all switch ports. An effective countermeasure is the limitation of the number of allowed MAC addresses for each access port.
- **DHCP server spoofing:** The attacker acts as a legitimate DHCP server in the network and maliciously configures clients to use it as the DNS server or default gateway, allowing the attacker to intercept their traffic. An effective defense against this threat is to deploy the Cisco IOS Software DHCP snooping feature.
- **DHCP starvation:** The attacker attempts to use all available DHCP-assigned addresses in the network to deny DHCP service to legitimate users. Limiting the maximum number of MAC addresses per port or a DHCP rate-limiting function provides two layers of defense against these threats.
- **ARP spoofing:** The attacker attempts to redirect traffic inside a subnet in order to intercept sensitive flows or spoof endpoint identities. The Cisco IOS Software ARP inspection feature provides an effective control to thwart this threat.

https://t.me/learningnets

certcollection.net

- **IP spoofing attacks:** The attacker uses a spoofed IP address to either mask their identity or to steal identities of legitimate systems. Inside the switched infrastructure, the Cisco IOS Software IP Source Guard feature or port ACLs both provide a defense against it.



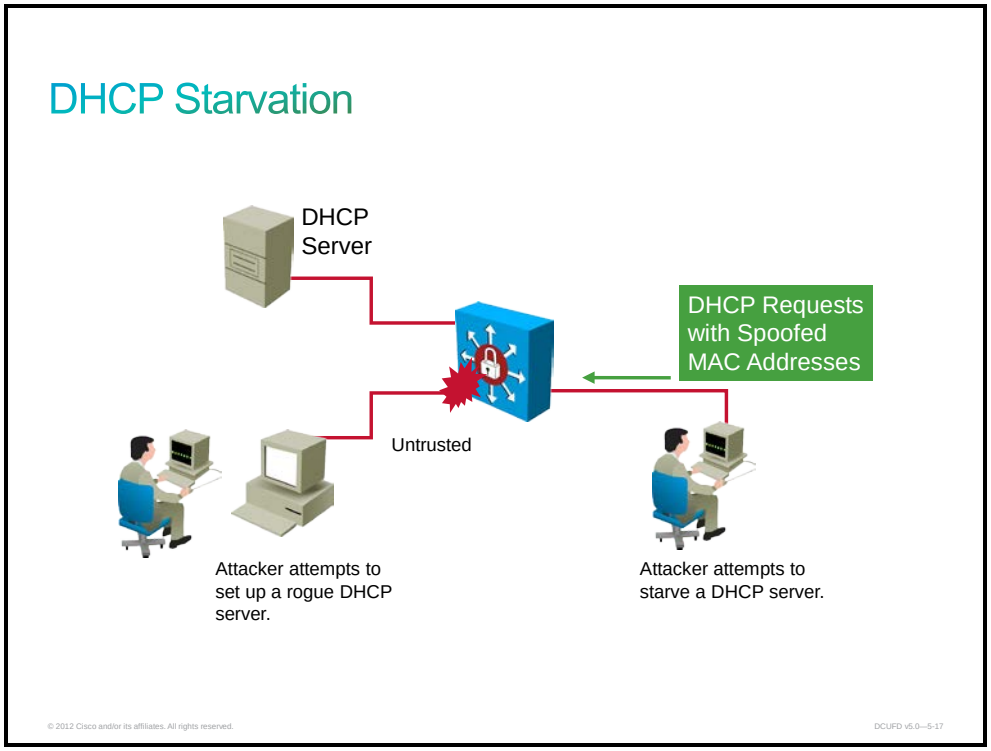
DHCP includes no authentication and is therefore easily vulnerable to spoofing attacks. The simplest attack is DHCP server spoofing, where the attacker pretends to be the DHCP server and replies to DHCP requests from legitimate clients, causing either DoS (by providing incorrect information), or confidentiality or integrity breaches via a man-in-the-middle attack. The attacker can assign himself as the default gateway or DNS server in all DHCP replies and then intercept all IP communication from the configured hosts to the rest of the network.

To mitigate this threat, you can use static IP addresses (this is obviously not scalable in large environments) or let the infrastructure control DHCP traffic by using DHCP snooping.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net



A DHCP starvation attack works by broadcasting DHCP requests with spoofed MAC addresses. If enough requests are sent, the network attacker can exhaust the address space that is available to the DHCP servers for a period of time. The network attacker can then set up a rogue DHCP server on their system and respond to new DHCP requests from clients on the network.

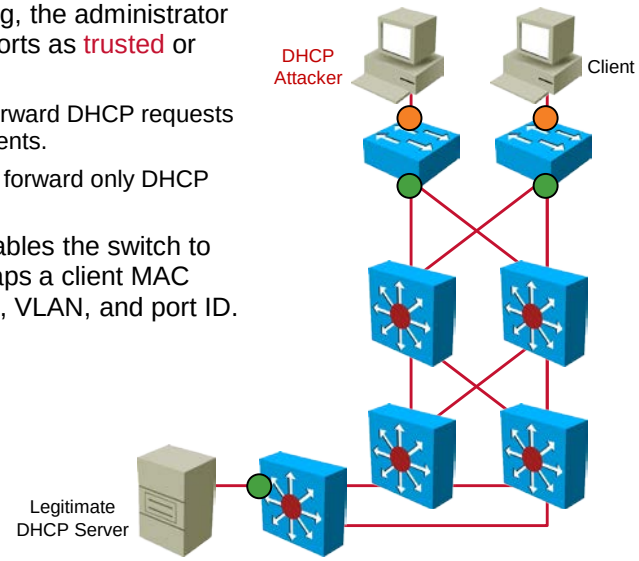
To mitigate DHCP address exhaustion attacks, you should deploy port security address limits, which set a higher limit of MAC addresses than can be accepted into the CAM table from any single port. Because each DHCP request must be sourced from a separate MAC address, this effectively limits the number of IP addresses that can be requested from a switch port-connected attacker. Set this to a value that is never legitimately exceeded in your environment.

https://t.me/learningnets

certcollection.net

DHCP Snooping

- With DHCP snooping, the administrator designates switch ports as trusted or untrusted.
  - Trusted ports can forward DHCP requests and acknowledgements.
  - Untrusted ports can forward only DHCP requests.
- DHCP snooping enables the switch to build a table that maps a client MAC address, IP address, VLAN, and port ID.



© 2012 Cisco and/or its affiliates. All rights reserved.

DHCP snooping is a Layer 2 security feature that prevents DHCP server spoofing attacks and mitigates DHCP starvation to a degree. DHCP snooping provides DHCP control by filtering untrusted DHCP messages and by building and maintaining a DHCP snooping binding database, which is also referred to as a DHCP snooping binding table.

For DHCP snooping, each switch port must be labeled as trusted or untrusted. Trusted ports are the ports over which the DHCP server is reachable and that will accept DHCP server replies. All other ports should be labeled as untrusted ports and can only source DHCP requests. Typically, this means the following:

- All access ports should be labeled as untrusted, except the port to which the DHCP server is directly connected.
- All interswitch ports should be labeled as trusted.
- All ports that point toward the DHCP server (that is, the ports over which the reply from the DHCP server is expected) should be labeled as trusted.

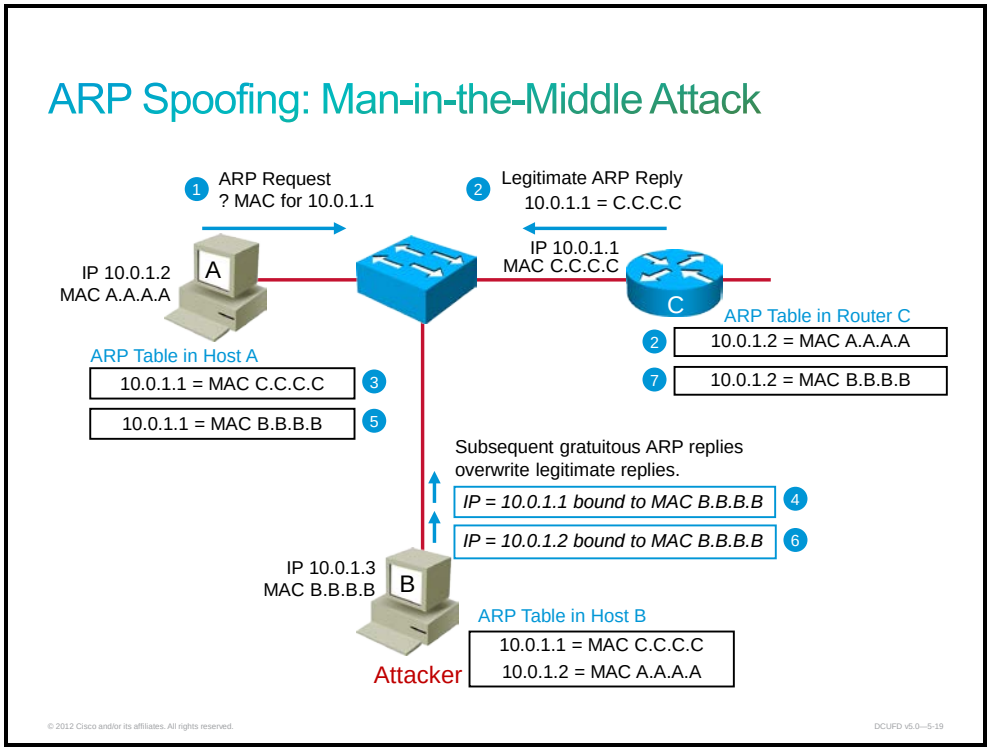
Untrusted ports are those ports that are not explicitly configured as trusted. A DHCP binding table is automatically built by analyzing normal DHCP transactions on all untrusted ports. Each entry contains the client MAC address, IP address, lease time, binding type, VLAN number, and port ID recorded as clients make DHCP requests. The table is then used to filter subsequent DHCP traffic. From a DHCP snooping perspective, untrusted access ports should not send any DHCP server responses, such as DHCPOFFER, DHCPACK, or DHCPNAK, and the switch will drop all such DHCP packets.

This figure shows the deployment of DHCP protection mechanisms on the access layer of the network. User ports are designated as untrusted for DHCP snooping, while Inter-Switch Links are designated as trusted if the DHCP server is reachable through the network core. User ports also have a limit of MAC addresses to prevent DHCP address exhaustion.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net



In normal ARP operation, a host sends a broadcast to determine the MAC address of a destination host with a particular IP address. The device with the IP address replies with its MAC address. The originating host caches the ARP response, using it to populate the destination Layer 2 header of packets that was sent to that IP address. By spoofing an ARP reply from a legitimate device with a Gratuitous ARP (GARP), an attacking device appears to be the destination host that is sought by the sender. The ARP reply from the attacker causes the sender to store the MAC address of the attacking system in its ARP cache. All packets that are destined for that IP address are forwarded to the attacker system.

An ARP spoofing attack, also known as ARP cache poisoning, can therefore target hosts, switches, and routers that are connected to your Layer 2 network by poisoning the ARP caches of systems that are connected to the subnet, and by intercepting traffic that is intended for other hosts on the subnet. The figure shows an example of ARP cache poisoning.

Step 1 Host A sends an ARP request for the Router C MAC address.

Step 2 Router C replies with its MAC and IP address. Router C also updates its ARP cache.

Step 3 Host A binds the Router C MAC address to its IP address in the ARP cache.

Step 4 Host B (attacker) sends a GARP to Host A, binding the MAC address of Router B to the IP address of Router C.

Step 5 Host A updates its ARP cache with the MAC address of Host B that is bound to the IP address of Router C.

Step 6 Host B (attacker) sends a GARP to Router C, binding the MAC address of Router B to the IP address of Router A.

Step 7 Router C updates its ARP cache with the MAC address of Host B that is bound to the IP address of Router A.

Step 8 Packets are now diverted through the attacker (Host B).

https://t.me/learningnets

certcollection.net

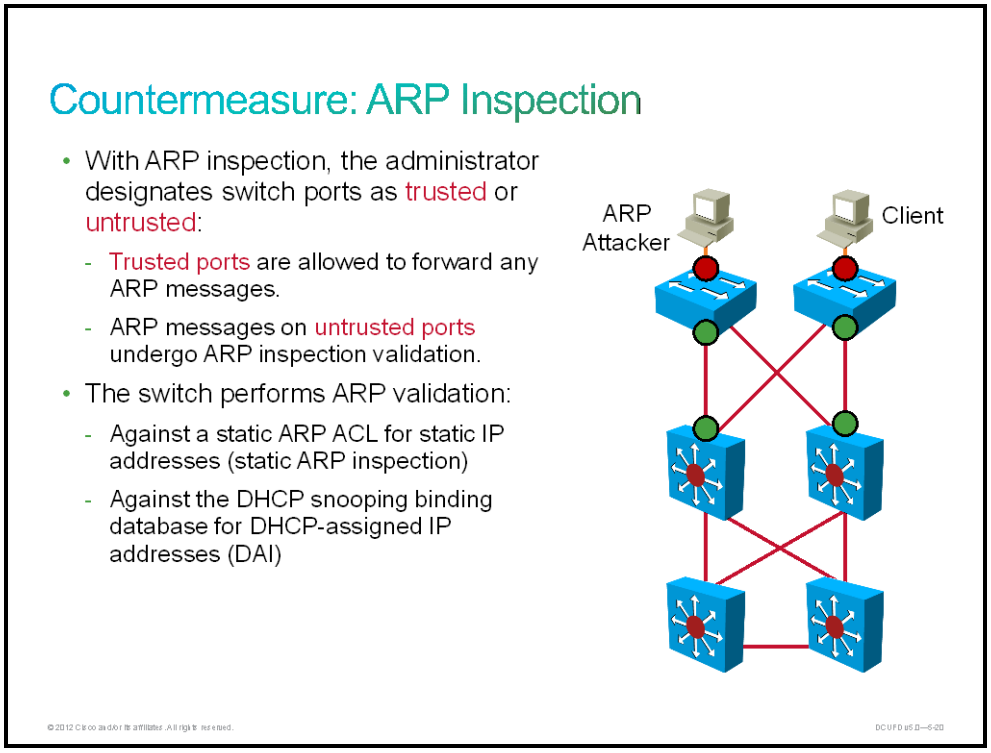
To address this ARP vulnerability in the infrastructure, you can use one of the following solutions:

- Static ARP or Dynamic ARP Inspection (DAI) in network switches
- Static ARP entries on infrastructure devices and, therefore, not use ARP on critical segments

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net



To prevent ARP spoofing, or “poisoning,” a switch can process transit ARP traffic to ensure that only valid ARP requests and responses are relayed. The ARP inspection feature of Cisco Catalyst switches prevents ARP spoofing attacks by intercepting and validating all ARP requests and responses. Each intercepted ARP reply is verified for valid MAC-to-IP address bindings before it is forwarded. ARP replies with invalid MAC-to-IP address bindings are dropped.

ARP inspection can determine the validity of an ARP reply based on bindings that are stored in a DHCP snooping database for DHCP-addressed hosts. For statically addressed hosts or network devices, ARP inspection can validate ARP packets against a user-configured ARP ACL that contains static MAC-to-IP-address mappings.

As with DHCP snooping, ARP inspection labels all switch ports as trusted or untrusted. The switch examines all ARP packets from untrusted ports and only forwards them if they contain an expected MAC-to-IP-address mapping.

In general, you should label the ports as follows:

- All ports that are connected to any host that is considered a possible source of attack should be labeled as untrusted. For hosts with static IP addresses, you must use static ARP ACL entries on the switch to permit their ARP traffic.
- All ports to other switches that are configured with ARP inspection should be labeled as trusted.
- All ports to other switches that do not support ARP inspection should be labeled as untrusted. In this case, make sure that DHCP traffic crosses the ARP inspection-enabled switch so that it can learn the legitimate IP-MAC mappings and permit associated ARP traffic.

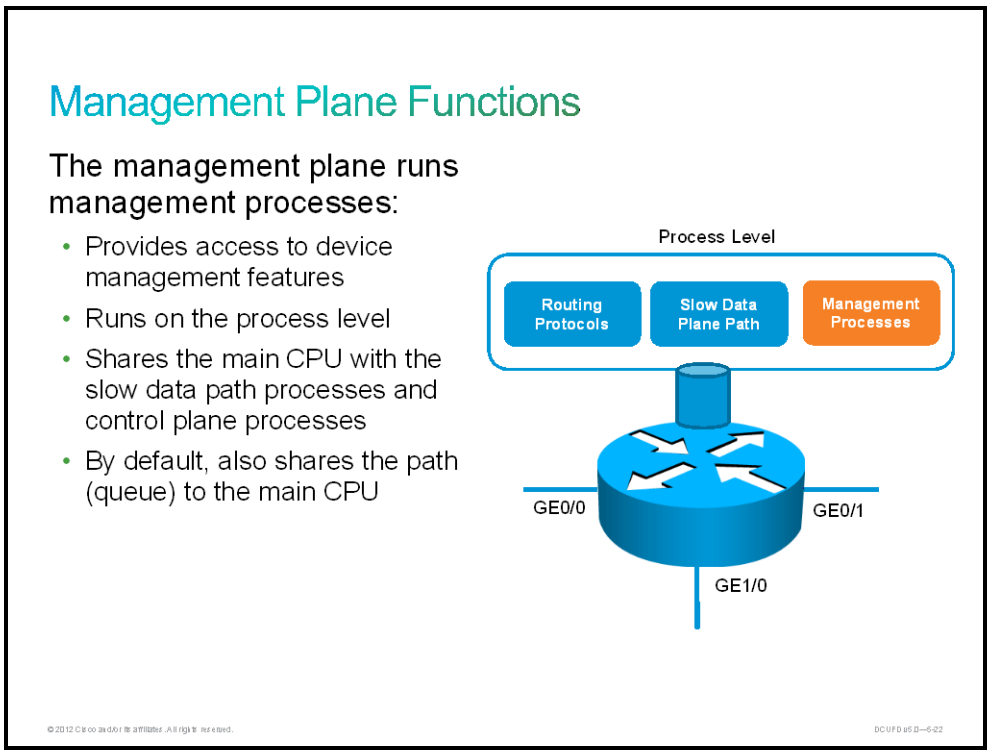
This figure shows the deployment of ARP protection mechanisms on the access (and partly distribution) layer of the network. User ports are designated as untrusted for ARP snooping, while Inter-Switch Links are designated as trusted.

https://t.me/learningnets

certcollection.net

Secure Management

This topic describes secure management networks.

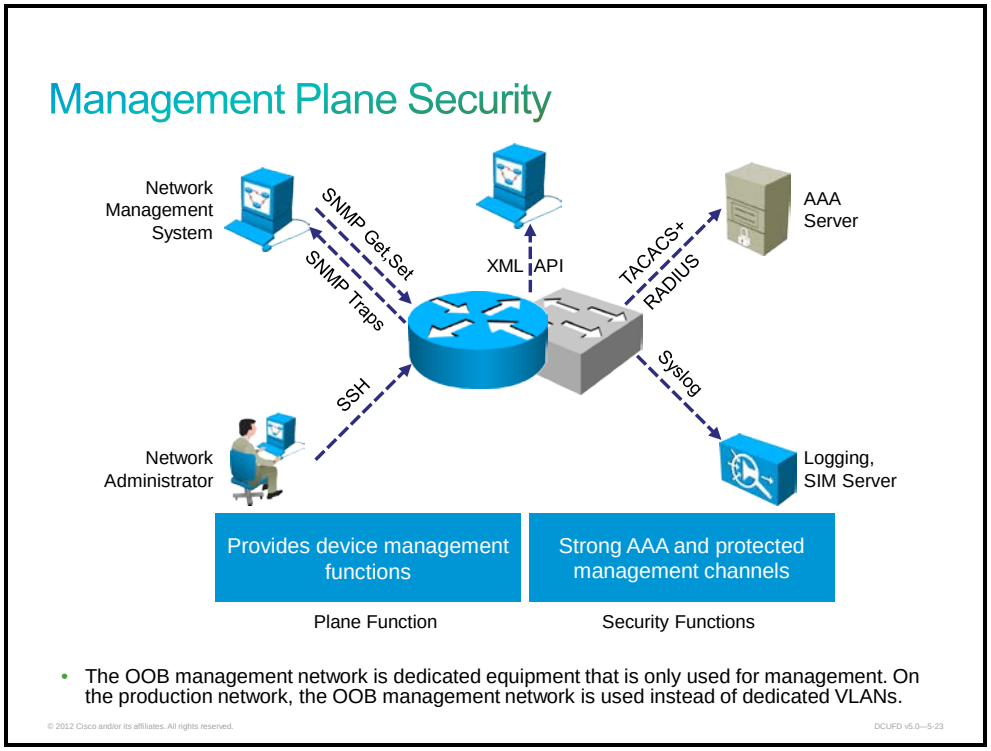


The management plane is a collection of processes that runs at the process level on the CPU of a network device and provides the device with management features and management access methods that administrators can use to locally or remotely access the device. The management plane functions share the CPU of the main router with control plane processes, such as routing protocols and data plane slow-path traffic processing functions. By default, these three aspects also share the packet path (queue) to the main CPU.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net



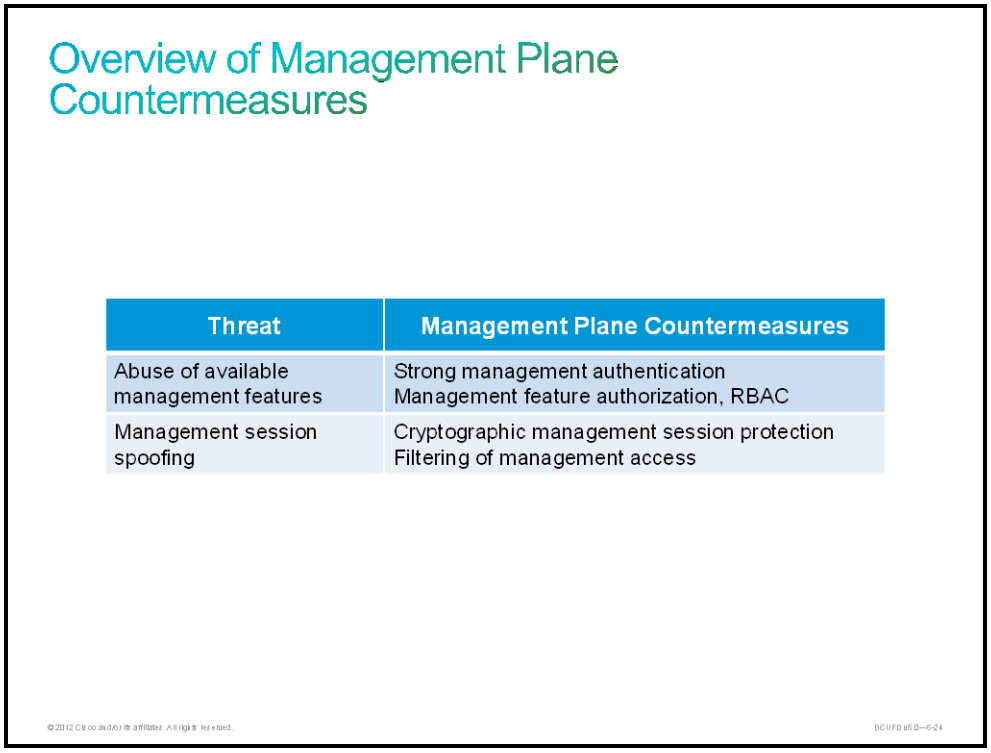
The management plane of a network device can provide security-related features that protect the device against compromise. These features include the following:

- A strong AAA feature set that validates administrator identity and suitably limits administrator access to device functions. Ideally, by using RBAC and minimal required privileges, the AAA feature audits all administrator actions and security-relevant device events. In most networks, centralizing AAA policy functions and audit trail collection—using a centralized AAA server function and a central logging or security information management system—is strongly recommended.
- Protected management channels over which administrators access devices. If the path between administrators and devices is not trustworthy, the management plane should provide cryptographic protection to management sessions or rely on data plane features of other network devices (for example, ARP inspection or VPN protection) to prevent interference with management sessions.

To protect management channels, one of the more secure options is an out-of-band (OOB) management network. The OOB management network is dedicated equipment that is only used for management. On the production network, the OOB management network is used instead of dedicated VLANs.

https://t.me/learningnets

certcollection.net



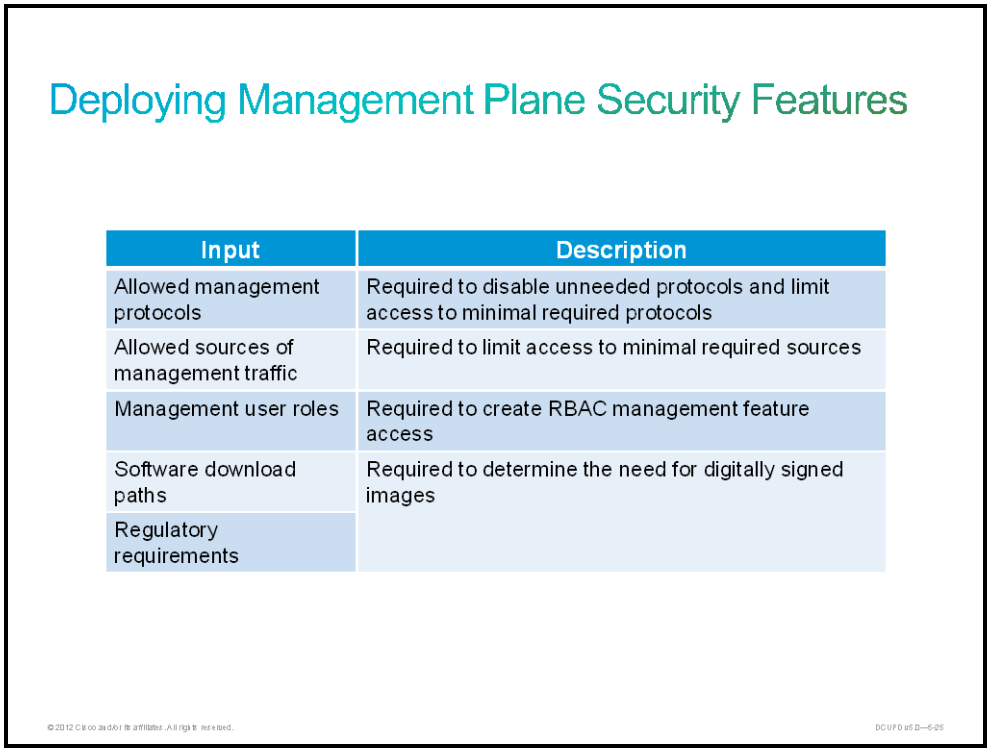
There are two major threat classes against the management plane:

- Abuse of management features by attackers and rogue administrators, where authenticated users act maliciously, or beyond their authorized profiles, to change the behavior of a network device in an undesirable manner. To reduce these risks, you can deploy strong authentication to prevent identity theft, and RBAC restrictions to limit access to specific management features.
- Spoofing of management sessions, where attackers hijack existing sessions, steal administrator credentials, or spoof legitimate administrator IP addresses in order to gain management access to a device. You can reduce these risks by using cryptographically protected management sessions and by tightly filtering access to devices to allow access from only specific networks if routing in your network is trusted and IP spoofing is unlikely.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

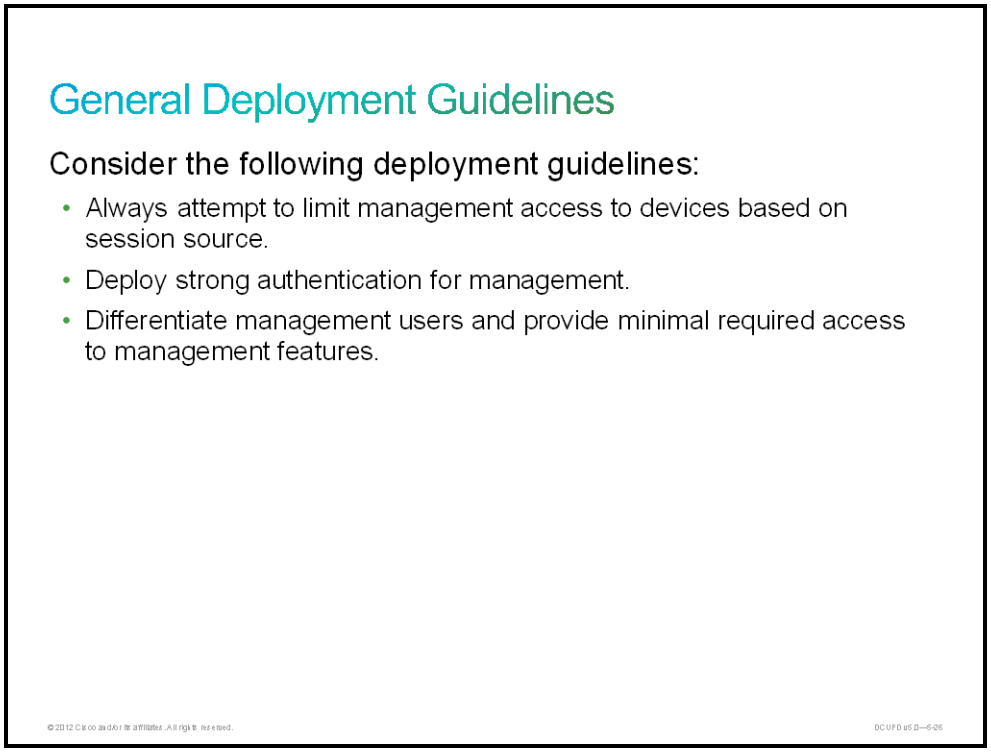


In order to deploy management plane security features, you first need to obtain some parameters from the environment in which these controls will be deployed. You should obtain the following information:

- A list of allowed management protocols in your environment, in order to disable all unneeded protocols and limit access to devices to a minimal set of protocols.
- A list of allowed sources of management traffic, in order to limit access to devices to a minimal set of management traffic sources.
- The various roles that are assigned to device administrators and the list of privileges for each role. This will allow you to implement a tight RBAC policy and only provide administrators with access to the management features that they need.
- The network paths that are used to download software to devices. If these paths use untrustworthy transport networks, you might consider migrating to platforms that support software image authenticity and integrity verification to reduce the risk of maliciously altered software being loaded to devices. Such verification may also be necessary for regulatory requirements .

https://t.me/learningnets

certcollection.net



When implementing management plane security features in your environment, consider the following general deployment guidelines:

- It is strongly recommended that you limit access to devices to the minimal needed sources of management traffic. This severely limits the attack surface of the device that an attacker can exploit.
- It is also strongly recommended that you use strong authentication methods for administrators in order to prevent attacks against administrators credentials. Consider using two-factor authentication (for example, one-time password generators together with PINs) in high-risk environments.
- To mitigate the threat of rogue administrators, differentiate management users and provide the minimal required management access to each management role. Also, deploy administrator auditing to generate a management audit trail.

 © 2012 Cisco Systems, Inc.

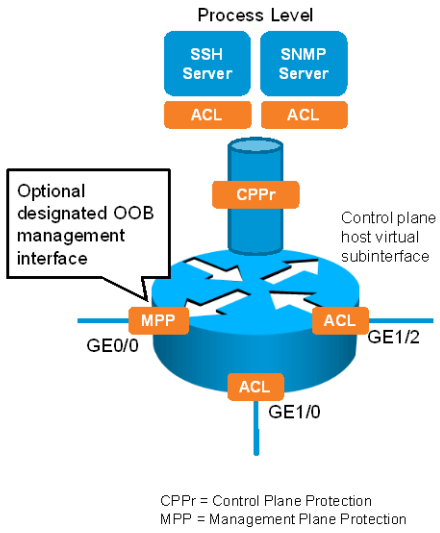
https://t.me/learningnets

certcollection.net

Management Plane Access Control

You can filter management access to devices using multiple independent mechanisms:

- Interface ACLs
- Service-specific ACLs
- Control Plane Protection
- Management Plane Protection



Limiting access to the management plane based on the source of management traffic can significantly reduce the risk of unauthorized management plane access. If your network is engineered to reduce the likelihood of IP spoofing—using, for example, Unicast Reverse Path Forwarding (uRPF) mechanisms, or ingress and egress antispoofing ACLs—you can effectively deploy IP-address-based filters to only allow access to device management planes from trusted hosts and networks.

For devices that are using Cisco IOS Software, you can employ various independent mechanisms to limit access to device management planes:

- You can deploy interface ACLs, which deny access to management IP addresses of the device on all device interfaces.
- You can deploy service-specific ACLs that limit access to a specific management process (using, for example, vty or SNMP server ACLs).
- Cisco IOS Software Control Plane Protection can simplify and enhance the management access control that is provided by interface ACLs by centralizing access control at a virtual control plane interface.
- The Cisco IOS Software Management Plane Protection (MPP) feature allows you to designate an interface on a device as the only interface over which management traffic is allowed to and from the device, allowing the device to connect to a dedicated OOB management network with minimal access control configuration.

The MPP feature in Cisco IOS Software enables you to restrict the interface (or interfaces) on which network management packets are allowed to enter a device. With the MPP feature, you can designate one or more router interfaces as management interfaces. Device management traffic is permitted to enter a device only through these management interfaces. After you enable MPP, no interfaces except designated management interfaces accept network management traffic that is destined to the device.

https://t.me/learningnets

certcollection.net

When you configure a management interface, all incoming packets through that interface are dropped except for those from the allowed management protocols. This configuration also drops packets on all interfaces from all of the remaining management protocols (supported in the MPP feature), including the interface that you configured. The allowed management protocols are dropped by all other interfaces unless the same protocol is enabled on those interfaces.

Designating management interfaces increases your control over the management of a device and provides more security for the device. There are various additional benefits of MPP:

- Improved performance for data packets on nonmanagement interfaces.
- Fewer ACLs are needed to restrict access to a device.
- Management packet floods on switching and routing interfaces are prevented from reaching the CPU.
- Network scalability.

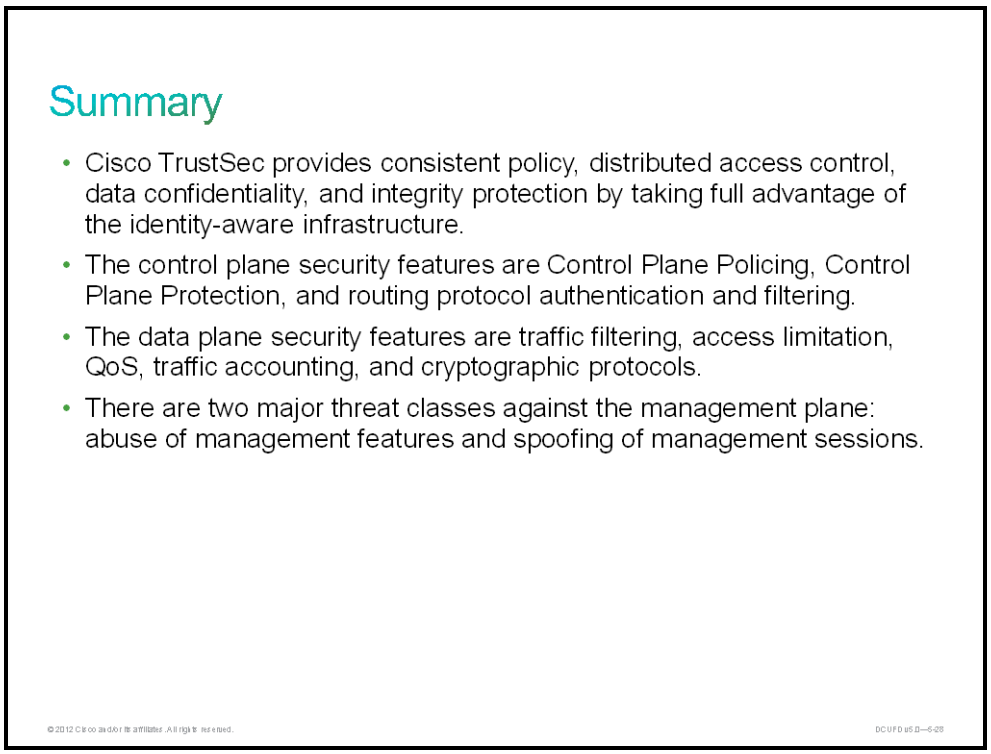
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

Summary

This topic summarizes the primary points that were discussed in this lesson.



https://t.me/learningnets

certcollection.net

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

Lesson 3

Designing Storage Security

Overview

All data centers store sensitive data. Securing the points where this sensitive data is stored is essential. Much focus is put on securing IP networks, but sensitive data also travels over SANs, where security measures also exist and can be implemented. In this lesson, you will learn how to design storage security on multiple levels—on the SAN level, while data is in transit, and when data is at rest.

Objectives

Upon completing this lesson, you will be able to design secure data center SANs. This ability includes being able to meet these objectives:

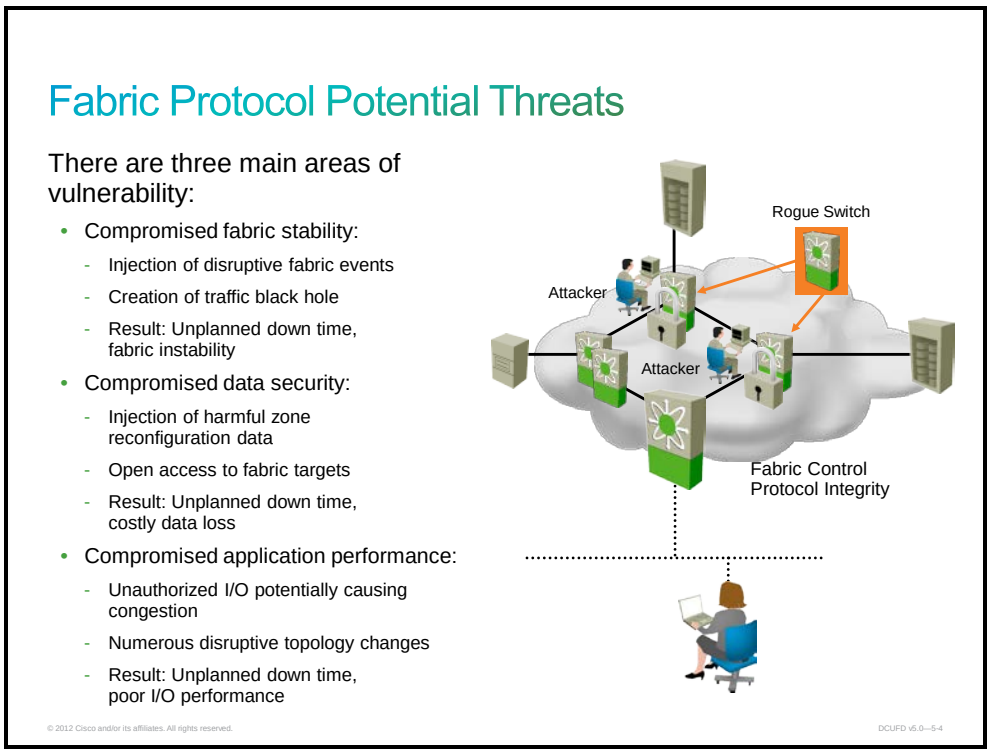
- Design secure SANs
- Explain security solutions for data encryption
- Outline security implications for IP-based storage

https://t.me/learningnets

certcollection.net

Design Secure SAN

This topic describes how to design a secure SAN.

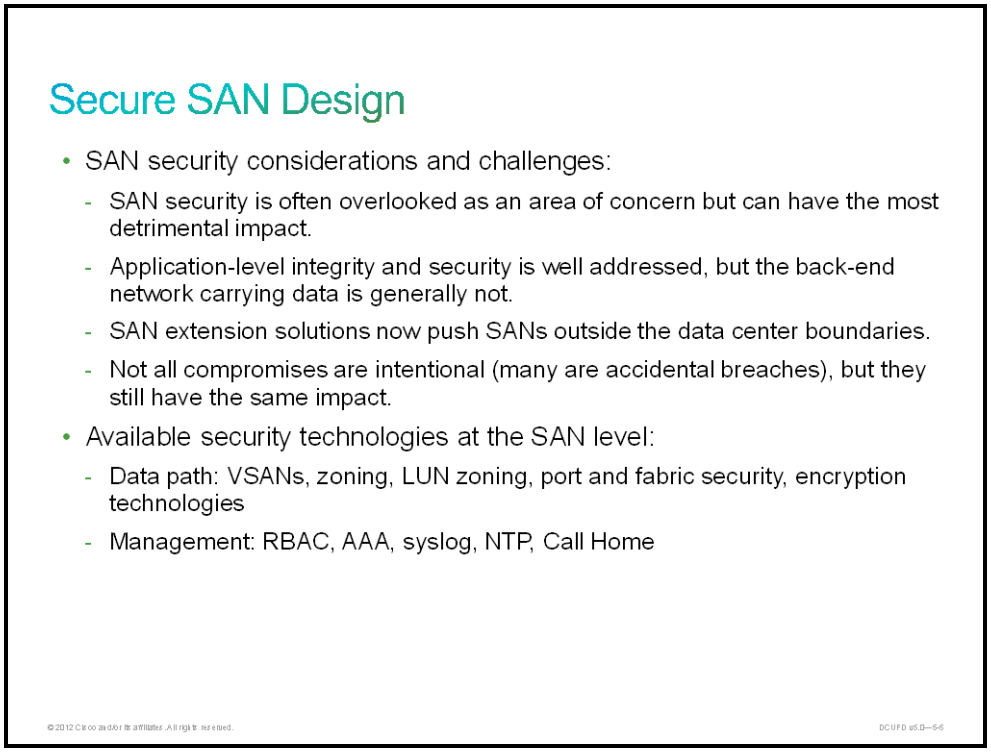


This figure summarizes potential threats to protocols that are running in the SAN fabric.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net



Generally, SAN environments are considered more secure than a typical LAN environment. To exploit a “security hole” in a SAN environment, you must have physical access to the data center or you must successfully break into a host that has unprotected access to the SAN fabric.

Because data centers have controlled physical access, this type of breach is not very common unless it is done by internal employees.

For someone to gain access to sensitive data by breaking into a host and exploiting its connection to the SAN, good knowledge of the SAN topology and of the storage array characteristics is required.

For designing secure SANs, there are technologies that restrict access from the hosts to the fabric. The most commonly used mechanism is called “zoning,” where the SAN administrator or the designer can define which host (initiator) can access which storage device (target) in the fabric.

Another mechanism that can be used to increase security is virtual storage area networks (VSANs). This mechanism is used to segment the SAN fabric into multiple virtual fabrics and, at the same time, provides complete isolation between VSANs.

Note Generally, no data can leak between VSANs, except if Inter-VSAN Routing (IVR) is configured. IVR is used to provide access to common devices (such as tape drives) from multiple fabrics.

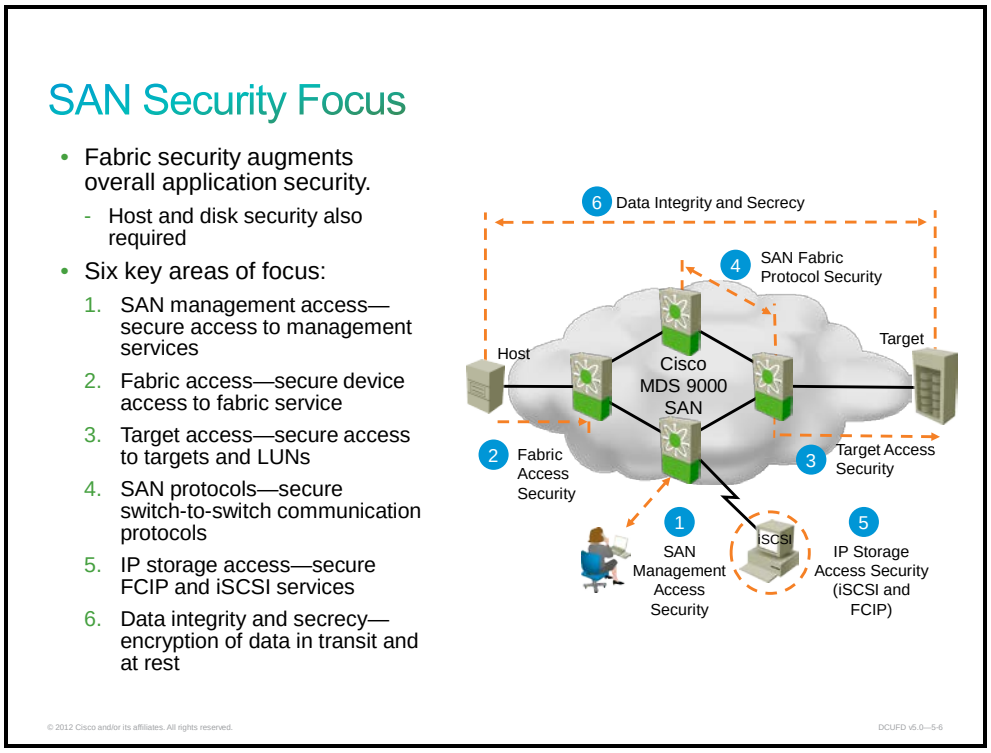
There are various mechanisms that provide additional security:

- Access control:
  - Port security
  - Device and switch authentication

https://t.me/learningnets

certcollection.net

- Encryption of data in transit:
  - Cisco TrustSec
  - IPsec for Fibre Channel over IP (FCIP) sessions
- Encryption of data at rest:
  - Cisco Storage Media Encryption (SME)
  - Storage array-based encryption technologies or tape drive-based encryption



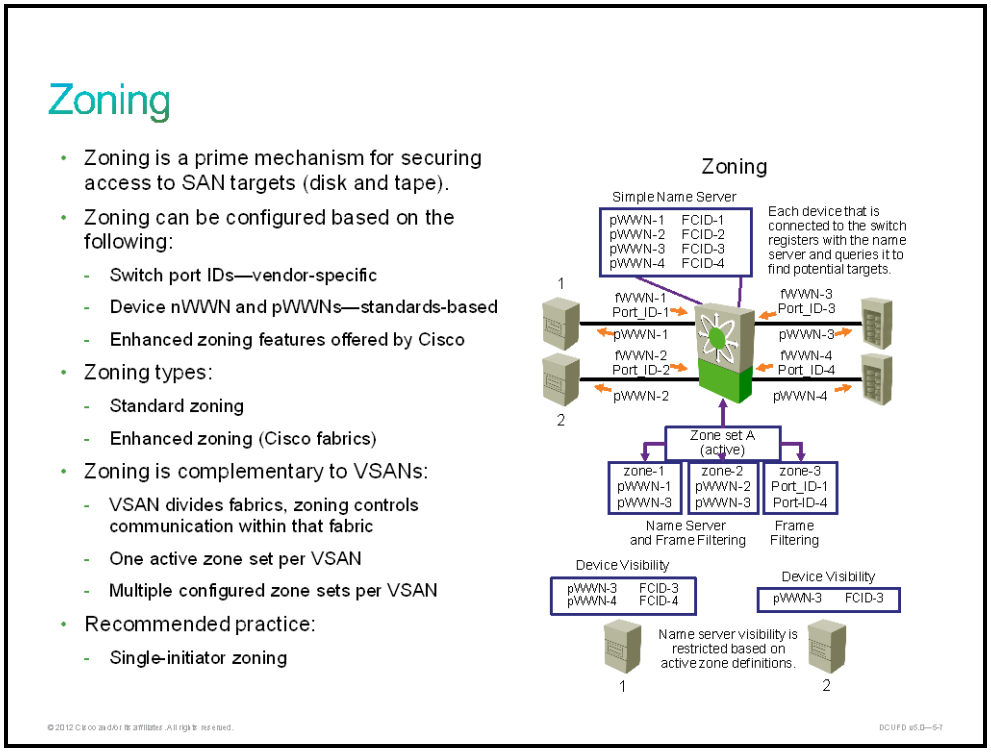
There are six major areas of focus in SAN security:

1. **SAN management access:** Secure access to management services. Securing the management plane is essential so that you do not have unwanted configuration changes.
2. **Fabric access:** Secure device access to the fabric service. Securing the fabric is important so that only devices that are authorized can access the SAN.
3. **Target access:** Secure access to targets and logical unit numbers (LUNs). You ensure that a host can read and write only on the device and portion of data that belongs to that host.
4. **SAN protocols:** Secure switch-to-switch communication protocols. You ensure consistency of the fabric so that no unauthorized devices can join the SAN and disrupt operation.
5. **IP storage access:** Secure FCIP and Internet Small Computer Systems Interface (iSCSI) services. Prepare the IP infrastructure to allow access only for flows that should be allowed.
6. **Data integrity and secrecy:** Encryption of data in transit and at rest. There are technologies that are available that ensure that your data is kept private while in transit and while at rest (written on the media).

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net



Zoning is a mechanism for fabric-based access control. Zoning is used to limit the devices that can communicate between themselves into zones. You can add a device into a zone, based on a unique device identifier—the port world wide name (pWWN), the Fibre Channel ID (FCID), or the alias (“device-alias” or “fc-alias”).

There are two major types of zoning, based on how zoning is implemented—soft zoning (which is switch-based name server filtering) and hard zoning (which is hardware-enforced frame filtering). Hard zoning is needed for true security.

Standard Zoning

All zoning information is stored fabric-wide in a zoning database. This database resides on a switch within the fabric that is responsible for distribution of this database to other switches in the SAN fabric. The switch that has the zoning database has both the full zone set and the active zone set (which is a subset of the full zone set). Zone sets are multiple zones that are grouped in a set. Other switches in the fabric have only active zone sets in their zoning database.

Be careful not to delete the full zone set on the switch. If you delete it, you lose the full zone set in the fabric, and only devices that are in active zone sets around the fabric are able to communicate.

Enhanced Zoning

Enhanced zoning takes away this limitation. Both active and full zone sets are distributed to switches in the fabric using the Cisco Fabric Services protocol.

Note	Enhanced zoning also supports adding devices into zones using their aliases. This can simplify the configuration.

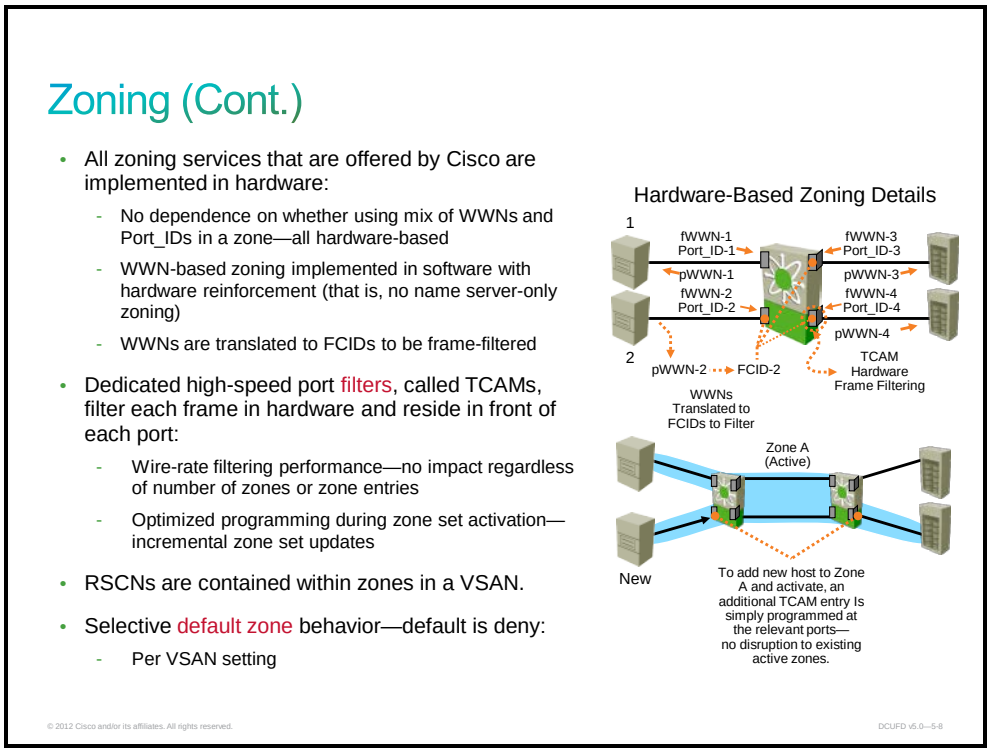
https://t.me/learningnets

certcollection.net

Single-Initiator Zoning

When adding devices into zones, you typically add one host (initiator) and one storage device (target). This is called single-initiator zoning.

When a host needs to have access to multiple volumes (also called LUNs), multiple targets are added to that zone. This is the case when a system needs to access its boot volume and a shared volume that contains virtual machines. This is a typical situation in a server virtualization scenario.



All zoning services that are offered by Cisco are implemented in hardware. There is no dependence on whether you are using a mix of world wide names (WWNs) and port IDs in a zone. The switch encodes the information in switching hardware.

WWN-based zoning is implemented in software with hardware reinforcement (that is, no name server-only zoning). WWNs are translated to FCIDs, and frame filtering on the interface is performed based on the FCIDs.

On the interface, there is a dedicated high-speed port filter that is called ternary content addressable memory (TCAM), which filters each frame in hardware and resides in front of each port, offering wire-rate filtering performance. The number of zones or zone entries has no effect on performance.

Changes to the zoning configuration are applied on the forwarding hardware when you activate the zone set. These changes are propagated using incremental updates, with no disruption to traffic flows.

Fibre Channel signals, such as Registered State Change Notifications (RSCNs), are contained within zones in a VSAN and do not disturb traffic or initiate changes in other zones.

The default setting for zone behavior is “deny” so that no traffic is allowed unless it is specifically permitted.

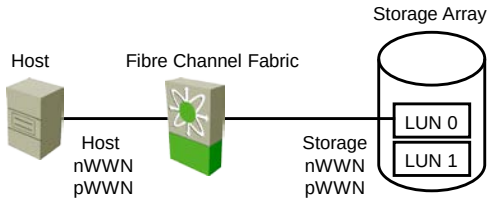
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

LUN Masking

- LUN masking is a mechanism that storage arrays use to present a single volume to a single host.
- Relation between a LUN ID and host nWWN and pWWN.
- A storage array presents only that volume to the host. The host operating system cannot access other volumes that belong to other hosts.
- LUN zoning achieves the same objective, but is confiugured in the fabric.



© 2012 Cisco and/or its affiliates. All rights reserved.

LUN masking is used on the storage array to provide an additional layer of access control and typically complements the zoning configuration.

When configuring zoning, you control which initiator can access which target, but you cannot control what an initiator can access on the target.

Storage arrays have multiple volumes and, typically, a volume is assigned to a single host. To bind the volume to a particular host (a server can have only one boot drive that cannot be shared by multiple servers), configure LUN masking.

Note If the storage array is not capable of LUN masking, or if the LUN masking license for your storage array is too expensive, you can use “LUN Zoning” on the Cisco MDS switch. This is a fabric-wide service and requires a license. The functionally does the same.

https://t.me/learningnets

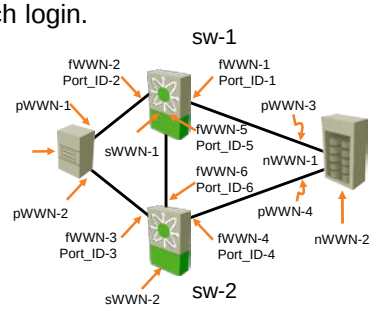
certcollection.net

Port Security

Port security is used to allow device-to-switch login.

Attributes to define binding configuration:

- pWWN—port WWN of attaching device
- nWWN—node WWN of attaching device
- fWWN—fabric WWN of switch port
- Port_ID—port identifier on switch (fc1/2)



Bind Host to sw-1 (Any Port)	**Security Group – sw-1 pWWN-1 or nWWN-1**
Bind Host, disk to sw-1 (Any Port)	**Security Group – sw-1 pWWN-1 or nWWN-1 pWWN-3 or nWWN-2**
Bind Host to sw-1, port 2	**Security Group – sw-1 pWWN-1 or nWWN-1 Port_ID-2 or fWWN-2**
Bind Host HBA-1 to sw-1, port 2	**Security Group – sw-1 pWWN-1 Port_ID-2 or fWWN-2**

© 2012 Cisco and/or its affiliates. All rights reserved.

Port security is a well-known mechanism that limits access to the fabric only to hosts with a particular WWN address. The functionality is very similar to LAN port security—if a host with an incorrect address attaches to the switch, the switch port is put in an error-disabled state. The administrator must manually review the situation, adjust the configuration, and re-enable the port.

Port security can be configured using the following parameters:

- **pWWN:** Port WWN of the attaching device
- **nWWN:** Node WWN of the attaching device
- **fWWN:** Fabric WWN of the switch port
- **Port_ID:** Port identifier on the switch (such as fc1/2)

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

Port-Mode Security

Port-mode security—allow edge ports to form F Ports or FL Ports only (that is, no ISL or EISL):

- Cisco MDS supports an Fx Port mode, which allows F Port or FL Port only
- Limit users who can change port mode via RBAC assignments

VSAN-based security—only allow access to devices within attached VSAN:

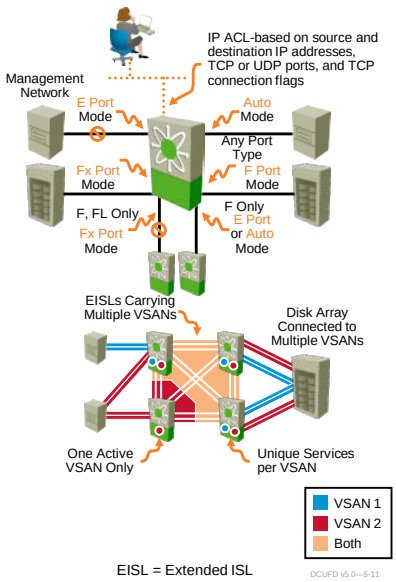
- Strict isolation based on fabric service partitioning and explicit frame tagging
- Independent name server table per VSAN
- Independent active zone set per VSAN
- Part of ANSI T11 fabric expansion study group

Management port access security:

- Provides IP ACLs for management traffic (such as SNMP, SSH, Telnet)

© 2012 Cisco and/or its affiliates. All rights reserved.

Port-Mode and VSAN-Based Security



Port-mode security allows ports to operate in various modes. For ports that are edge ports, allow Fx Port mode only. There is no establishment of interswitch links (ISLs) that use E or TE Port modes. As such, you cannot connect a rogue switch on the ports that are designated for access connections.

The Cisco MDS switch supports Fx Port mode, which allows F Port or FL Port types to be autodetected on the access port.

Note A general recommendation is to limit users who can change the port mode via role-based access control (RBAC) assignments.

In addition to port-mode security, you can configure VSAN-based security to only allow access to devices within an attached VSAN. VSANs offer strict isolation based on fabric service partitioning and explicit frame tagging when traffic traverses the links between switches.

In addition, the VSAN separation offers the following:

- Independent name server table per VSAN
- Independent active zone set per VSAN
- Part of ANSI T11 fabric expansion study group

To provide for management port access security, you can use IP access control lists (ACLs) for management traffic (such as Simple Network Management Protocol [SNMP], Secure Shell [SSH], and Telnet).

https://t.me/learningnets

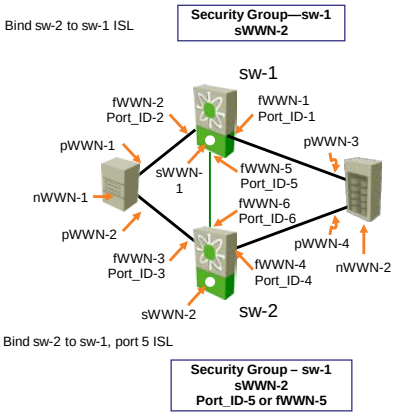
certcollection.net

Fabric Binding

Fabric binding is used to allow ISL establishment.

Attributes to define binding configuration:

- fWWN—fabric WWN of switch port
- sWWN—switch WWN
- Port_ID—port identifier on switch (fc1/2)



© 2012 Cisco and/or its affiliates. All rights reserved.

Fabric binding is a mechanism that allows you to specify which Fibre Channel switch can join the fabric. Fabric binding is used to prevent rogue SAN switches from attaching to the fabric and changing the topology and active fabric databases. Fabric binding also prevents access of rogue hosts to legitimate targets.

To facilitate the initial configuration, you can use automatic learning. After a switch is connected to the fabric, only that switch is allowed in the future. Typically, the switch WWN is used to restrict which switch can be a member of the fabric.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

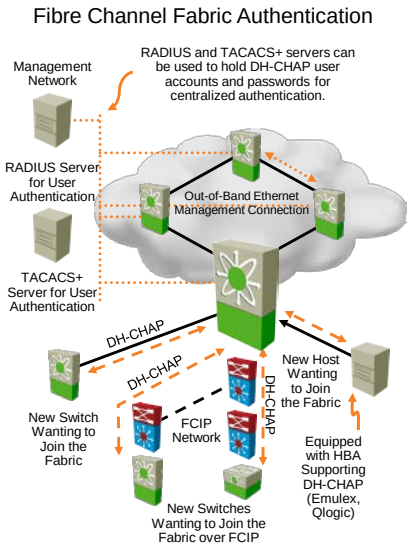
Authentication

Device authentication provides a stronger means of ensuring device identity.

- WWNs can be spoofed by simple means.

DH-CHAP provides authentication mechanisms:

- Switch-to-switch authentication
- Device-to-switch authentication (when adopting HBA that supports DH-CHAP)



© 2012 Cisco and/or its affiliates. All rights reserved.

Authentication can be used for SAN management access. The Fibre Channel switches and the SAN fabric are protected from unauthorized access and configuration changes.

Device Authentication

Device authentication provides a stronger means of ensuring device identity, rather than just using port security mechanisms. (WWNs can be spoofed easily. They are even customizable.)

The Diffie-Hellman Challenge Handshake Authentication Protocol (DH-CHAP) provides an authentication mechanism that allows switch-to-switch authentication and device-to-switch authentication if the host bus adapter (HBA) supports it.

Note The ANSI T11 FC-SP security protocols working group is responsible for standards for device authentication. Cisco was the prime contributor to the working group.

User Authentication for Switch Access

If an intruder has access to the switch configuration, there is a lot of damage that can be done, from alterations of LUN visibility to corruption of data on the disks to data theft. An intruder can change the zoning configuration, can provide rogue host access to a data LUN, and can copy sensitive data.

SAN management security protects all aspects of the SAN switch, including console sessions, GUI management, file transfer, and Network Time Protocol (NTP).

RBAC is typically combined with authentication, authorization, and accounting (AAA), giving administrators the required privileges over the whole fabric or over only parts of the fabric (VSANs) for which the administrator is responsible.

There is a general recommendation that, together with management authentication, you should enable accounting as well. Accounting enables you to track what commands were issued by which user on the fabric.

To facilitate correlation of events on multiple devices, you should enable NTP on devices in the fabric.

https://t.me/learningnets

certcollection.net

Best Practice

Use zoning services to isolate where required:

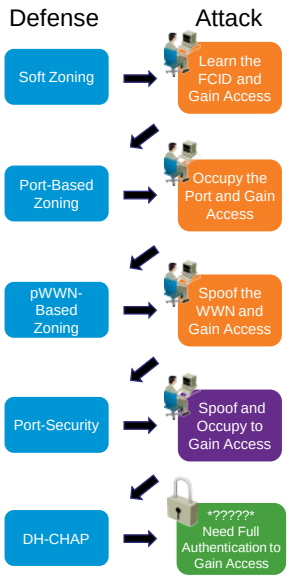
- Port or WWN-based, all hardware enforced
- Set default zone policies to deny

It is suggested to only allow zoning configuration from one or two switches to minimize access:

- Use RBAC to create two roles, only one allowing zoning configuration
- Alternatively, use RADIUS or TACACS+ to assign roles based on particular switch; more flexible

Use WWN-based zoning for convenience and use port-security features to harden switch access:

- Works well for interoperability with switches that are not from Cisco



© 2012 Cisco and/or its affiliates. All rights reserved.

This figure summarizes best practices to prevent attacks to the SAN fabric.

It is important to secure the fabric control protocols to ensure fabric stability:

- The first step is to secure access to the control protocol configuration via RBAC.
- Enable port security for switch binding.
- Use Fibre Channel Security Protocol (FC-SP) for switch-to-switch authentication to block rogue ISLs.

The “plug and play” (meaning that it will work as soon as it is active on the network) fabric protocol configuration is convenient. However, static configuration is more secure.

- Configure the static principle switch.
- Enable static domain IDs.
- Enable static FCIDs (optional but recommended):
  - Benefit for HP/UX and AIX environments
- Enable Reconfigure Fabric (RCF)-reject, especially on long-haul links.
- Enable RSCN suppression where necessary.

Use VSANs to divide the fabric and to manage each part individually. This approach also improves resiliency.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

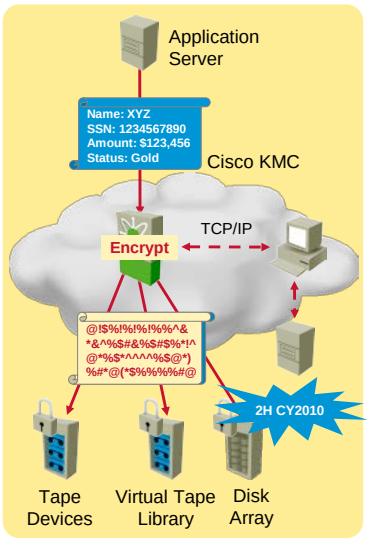
Data Security Solutions

This topic describes security solutions for data encryption.

Encryption of Data at Rest

Cisco Storage Media Encryption

- Encrypts media for SAN-attached tapes, virtual tape libraries, and disk arrays
  - Uses IEEE AES-256 encryption
  - Disk—XTS; Tape—GCM
- Solution includes Cisco KMC for provisioning and key management
  - Integration with RSA Key Manager
- Handles traffic from any VSAN in fabric



© 2012 Cisco and/or its affiliates. All rights reserved.

There are two possibilities regarding where and how to encrypt data—encryption of data at rest and encryption of data in transit.

Cisco Storage Media Encryption (SME) encrypts data that is being written to a disk or to a tape.

The necessary keys are managed by Cisco Key Management Center (KMC).

Cisco SME can process any traffic in the fabric, as long as that traffic is redirected to the service modules that perform the actual encryption, such as the Cisco MSM-18/4, the SSN-16 modules for Cisco MDS 9500 director, or a Cisco 9222i switch.

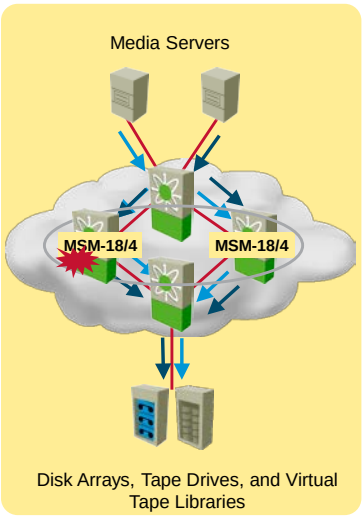
Note You can recover the data offline by using a Linux-based tool. However, you need to have appropriate keys to decrypt the data.

https://t.me/learningnets

certcollection.net

SME Scalablity and High Availability

- Allows rapid deployment:
  - No SAN reconfiguration or rewiring
  - Provision as a simple, logical process of selecting what to encrypt
  - Provision at the data center level and not at the module level
  - Integrates transparently in Cisco MDS fabrics using Fibre Channel redirect
- Modular, clustered solution offers highly scalable and reliable performance
  - Up to 4 switches and 32 encryption units
  - Support dual-fabric configurations
- Automatically load-balances
- Redirects traffic if a failure occurs



© 2012 Cisco and/or its affiliates. All rights reserved.

To overcome failure scenarios, the Cisco SME software can run in clustered pairs. The primary functionality on which the Cisco SME system is based is Fibre Channel redirect, which is available on Cisco MDS systems. The Fibre Channel redirect function redirects the data flow that needs to be encrypted to the Cisco MSM-18/4 or SSN-16 modules that are running Cisco SME.

© 2012 Cisco Systems, Inc.

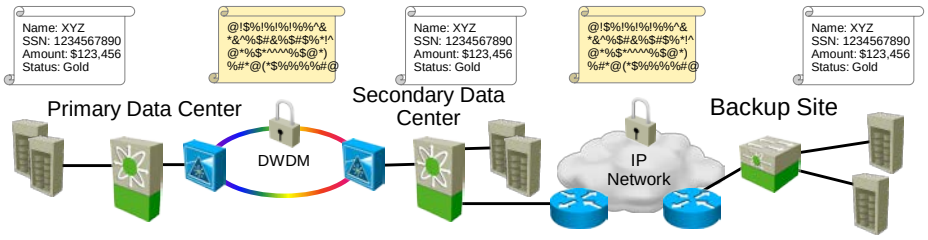
https://t.me/learningnets

certcollection.net

Encryption of Data in Transit

Link Layer Security

- Data confidentiality requirements are part of business today.
- Businesses need to ensure that data is not compromised while being transmitted between data centers.
- Cisco TrustSec (Fibre Channel) and IPsec (FCIP) are used to secure data over ISLs between switches.



DWDM = dense wavelength-division multiplexing

© 2012 Cisco and/or its affiliates. All rights reserved.

When using a SAN extension between data centers, the requirement to encrypt the data between them naturally emerges, especially if the network between data centers is public.

Cisco TrustSec is a technology that allows encryption of Fibre Channel data on the interface level, just before the frames leave the switch.

The encryption is done in hardware, using the Advanced Encryption Standard (AES) algorithm.

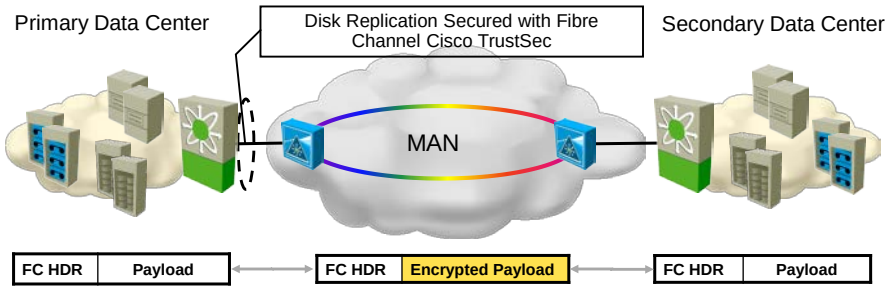
When using the IP protocol for transport of Fibre Channel frames, such as in the case of FCIP, you can use IP Security (IPsec) to encrypt the data stream as you would do with any other VPN traffic.

https://t.me/learningnets

certcollection.net

Encryption of Data in Transit (Cont.)

Cisco TrustSec Fibre Channel Link Encryption



- Extension to FC-SP to provide encryption of data:
  - DH-CHAP is used for peer authentication
  - Encryption: AES 128-bit key
- Integrity, confidentiality, authentication, no replay across dark fiber, MAN
- Hardware-based 8-Gb Fibre Channel wire rate on third-generation 8-Gb Fibre Channel blades
- No change to existing SAN, functionality provided by edge switches

© 2012 Cisco and/or its affiliates. All rights reserved.

FC HDR = Fibre Channel header

Cisco TrustSec Fibre Channel encryption is an extension to FC-SP. The DH-CHAP protocol is used for authentication of the peer device.

Integrity, confidentiality, authentication, and anti-replay protection are guaranteed across a dark fiber link, or a metropolitan-area network (MAN).

Encryption is hardware-based, wire rate, even for 8-Gb Fibre Channel connections on the third generation of 8-Gb Fibre Channel I/O modules for Cisco MDS 9500 switches.

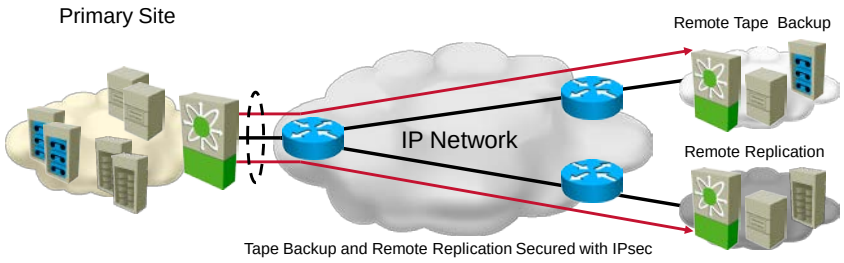
© 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

Encryption of Data in Transit (Cont.)

FCIP Security—IPsec Encryption



- Standards-based IPsec Encryption
  - IKE for protocol and algorithm negotiation, and key generation
  - Encryption: AES (128- or 256-bit key), DES (56-bit), 3DES (168-bit)
- Hardware-based Gigabit Ethernet wire-rate performance with latency ~ 10µs per packet
- Provides integrity, confidentiality, origin authentication, anti-replay across the IP network

© 2012 Cisco and/or its affiliates. All rights reserved.

When using IP transport for Fibre Channel traffic, standards-based IPsec encryption is used. The Internet Key Exchange (IKE) for the protocol algorithm is used for key negotiation and generation. The encryption used can be AES (128- or 256-bit key), Data Encryption Standard (DES) (56 bit), or Triple DES (3DES) (168 bit).

FCIP encryption using the IPsec protocol suite can be performed on Cisco MDS 9500 Series switches, with hardware-based Gigabit Ethernet wire-rate performance and with a latency of approximately 10 µs per packet.

IPsec provides integrity, confidentiality, origin authentication, and antireplay across any private or public IP network.

https://t.me/learningnets

certcollection.net

Secure IP-Based Storage Design

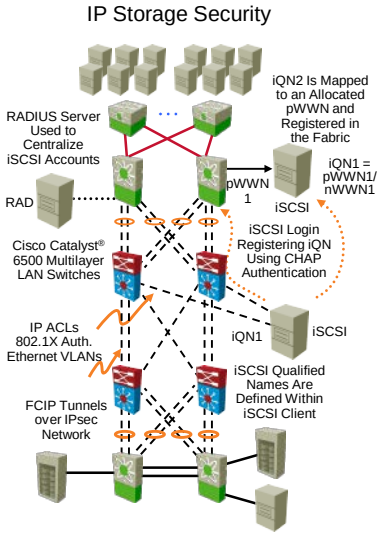
This topic describes security implications for IP-based storage.

Securing IP-Based Storage

iSCSI leverages many of the security features inherent in Ethernet and IP:

- Ethernet ACLs ↔ Fibre Channel zones
- Ethernet VLANs ↔ Fibre Channel VSANs
- Ethernet 802.1x port security ↔ Fibre Channel port security
- iSCSI authentication ↔ Fibre Channel DH-CHAP authentication

iSCSI offers a LUN masking and mapping capability as part of the gateway function.



© 2012 Cisco and/or its affiliates. All rights reserved.

IP-based storage using the iSCSI protocol is popular for customers that do not wish to use Fibre Channel storage. iSCSI is a block-based protocol that uses IP and TCP for transport of SCSI blocks.

The iSCSI protocol has many of the security features inherent in Ethernet and IP:

- Ethernet ACLs are functionally equivalent to Fibre Channel zones.
- Ethernet VLANs are functionally equivalent to Fibre Channel VSANs.
- Ethernet IEEE 802.1X port security is functionally equivalent to Fibre Channel port security.
- iSCSI authentication is functionally equivalent to Fibre Channel DH-CHAP authentication.

The iSCSI solution offers a LUN masking and mapping capability as part of its gateway function.

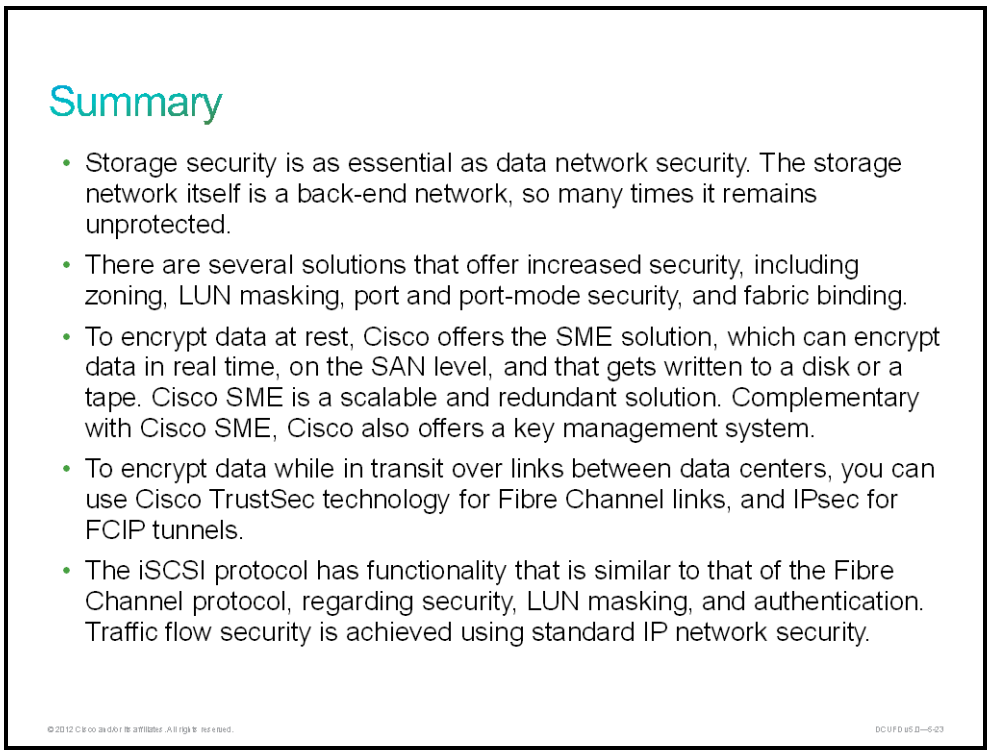
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

Summary

This topic summarizes the primary points that were discussed in this lesson.



https://t.me/learningnets

certcollection.net

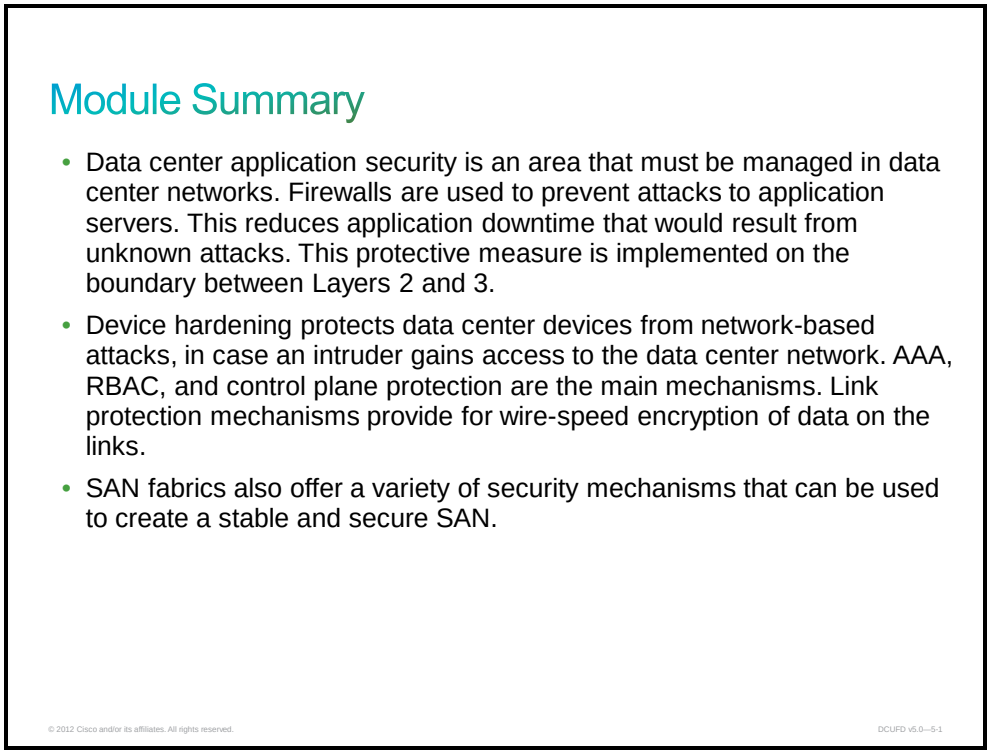
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

Module Summary

This topic summarizes the primary points that were discussed in this module.



Data center security is a broad topic, ranging from application security to device and link security. Application security is enforced with firewalls to protect servers from attacks from the Internet and from users. Device security offers protection from attacks that target network devices.

Mechanisms that help with security are authentication, authorization, and accounting (AAA) services and role-based access control (RBAC). Only users with dedicated roles can manage their part of device configuration.

Storage security is often overlooked because it is considered as back end, but severe service disruptions can occur if an intruder can gain access to the fabric or by misconfigurations.

https://t.me/learningnets

certcollection.net

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

Module Self-Check

Use these questions to review what you learned in this module. The correct answers and solutions are found in the Module Self-Check Answer Key.

Q1) What are three major threat classes? (Choose three.) (Source: Designing Data Center Application Security)

A) threats to information confidentiality
B) threats to information integrity
C) repudiation threats
D) threats to information availability
E) theft of service threats

Q2) One common approach to protecting network resources involves partitioning of the network into security domains. What is normally implemented on the domain boundaries? (Source: Designing Data Center Application Security)

A) filtering with firewalls
B) filtering with routers
C) routing with routers
D) routing with firewalls

Q3) Which firewall mode is used when forwarding non-IP traffic is required? (Source: Designing Data Center Application Security)

Q4) Which three items are independent for each context firewall? (Choose three.) (Source: Designing Data Center Application Security)

A) security policy
B) interfaces
C) administrators
D) power supply
E) hardware

Q5) On which two boundaries can firewalls be implemented? (Choose two.) (Source: Designing Data Center Application Security)

A) Layer 1 boundary
B) Layer 2 boundary
C) Layer 3 boundary
D) Layer 4 boundary

Q6) Match the firewall system approach to its correct description. (Source: Designing Data Center Application Security)

A) restrictive (or proactive) approach
B) permissive (or reactive) approach

_____ 1. The firewall, by default, permits all communication and only blocks the aspects of communication that it considers malicious, based on its attack signature database.

_____ 2. The firewall, by default, denies all communication and only allows the aspects of communication that are explicitly permitted.

https://t.me/learningnets

certcollection.net

Q7) Cisco TrustSec mitigates security risks by providing comprehensive visibility into who and what is connecting across the entire network infrastructure, as well as exceptional control over what and where they can go. (Source: Designing Link Security Technologies and Device Hardening)

A) true
B) false

Q8) Match the role-based tag steps to their correct descriptions. (Source: Designing Link Security Technologies and Device Hardening)

A) Step 1
B) Step 2
C) Step 3

_____ 1. The switch applies a tag to the user traffic.

_____ 2. A user or device logs into the network via IEEE 802.1X.

_____ 3. The Cisco ISE server sends a tag in the authorization result, based on the role of the user or device.

Q9) The functionality of a network device is segmented into three planes of operation. Match each plane to its correct description. (Source: Designing Link Security Technologies and Device Hardening)

A) management plane
B) control plane
C) data plane

_____ 1. This plane provides the device with all the functions that administrators need to provision the configuration and monitor the operation of the device.

_____ 2. This plane allows the device to forward network traffic and apply services (such as security, QoS, accounting, and optimization) to it as it is forwarded.

_____ 3. This plane allows the device to build all required control structures (such as routing table, forwarding table, and MAC address table) that will allow the data plane to operate correctly.

Q10) What are three control plane countermeasures for slow path denial-of-service attacks? (Choose three.) (Source: Designing Link Security Technologies and Device Hardening)

A) Fast-path data plane ACLs
B) Control Plane Policing
C) Control Plane Protection
D) Routing Protocol Authentication
E) Routing Protocol Filtering

Q11) What are the two management plane countermeasures for abuse of available management features? (Choose two.) (Source: Designing Link Security Technologies and Device Hardening)

A) strong management authentication
B) management feature authorization, RBAC
C) cryptographic management session protection
D) filtering of management access

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

Q12) What are the two management plane countermeasures for management session spoofing? (Choose two.) (Source: Designing Link Security Technologies and Device Hardening)

A) strong management authentication
B) management feature authorization, RBAC
C) cryptographic management session protection
D) filtering of management access

Q13) What are three major threats to SAN fabrics? (Choose three.) (Source: Designing Storage Security)

A) unplanned downtime
B) fabric instability
C) poor I/O performance
D) selective isolation
E) security threats

Q14) What are three SAN security mechanisms that prevent attachment of rogue devices? (Choose three.) (Source: Designing Storage Security)

A) port security
B) port mode security
C) fabric binding
D) user authentication
E) unavailability of remote access

Q15) Which solution prevents unauthorized reading of data that is stored on tapes? (Source: Designing Storage Security)

A) Cisco iSCSI Gateway
B) Cisco Storage Media Encryption
C) Cisco Data Mobility Manager
D) Cisco TrustSec

https://t.me/learningnets

certcollection.net

Module Self-Check Answer Key

Q1) A, B, D

Q2) A

Q3) transparent

Q4) A; B, C

Q5) B, C

Q6) 1-B
2-A

Q7) A

Q8) 1-C
2-A
3-B

Q9) A-1
B-3
C-2

Q10) A, B, C

Q11) A, B

Q12) C, D

Q13) A, B, C

Q14) A, B, C

Q15) B

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

Module 6

Data Center Application Services

Overview

In this module, you will learn about application services in the data center. The applications can be of many types, ranging from client-server to clustered or distributed. The network must accommodate the application design to provide optimum performance.

Cisco provides solutions for application services that use hardware acceleration and offer load balancing to simplify the application delivery to the end users.

Module Objectives

Upon completing this module, you will be able to design data center infrastructure that is required to implement network-based application services. This ability includes being able to meet these objectives:

- Present data center application architecture
- Design the network infrastructure that is needed to perform application services
- Design global load-balancing solutions

https://t.me/learningnets

certcollection.net

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

Lesson 1

Designing Data Center Application Architecture

Overview

This lesson explains the impact of data center application design on the network architecture and describes application design and application architectures. This lesson also explains application optimization technologies. You will learn about the application challenges of a remote network and about the features of Cisco Wide Area Application Services (WAAS) technology, which is designed to help overcome and enable more efficient application delivery.

Objectives

Upon completing this lesson, you will be able to design data center application architecture. This ability includes being able to meet these objectives:

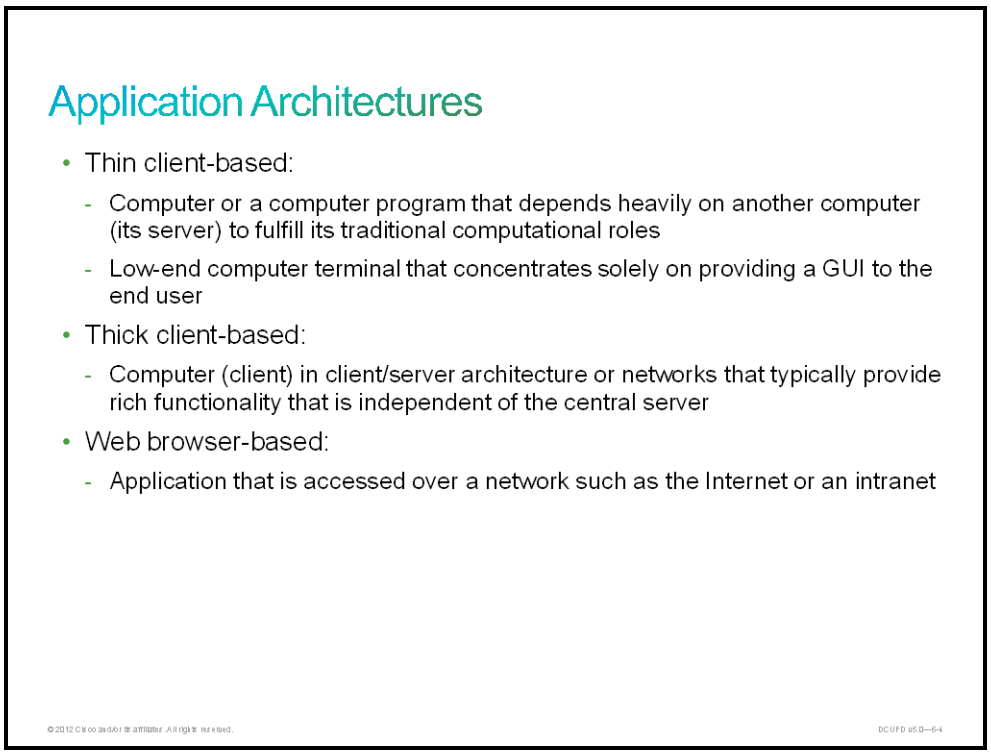
- Explain application architecture and design
- Explain application tiering
- Explain application optimization technologies

https://t.me/learningnets

certcollection.net

Application Architecture and Design

This topic describes application architecture and design.



A thin client (sometimes also called a lean or slim client) is a computer or a computer program that depends heavily on another computer (its server) to fulfill its traditional computational roles. This stands in contrast to the traditional fat client, a computer that is designed to take on these roles by itself. The exact roles that are assumed by the server may vary, from providing data persistence (for example, for diskless nodes) to actual information processing on the behalf of the client. Thin clients are components of a broader computer infrastructure, where many clients share their computations with the same server. As such, thin client infrastructures can be viewed as the amortization of computing services across several user interfaces. This is desirable in contexts where individual fat clients have much more functionality or power than the infrastructure either requires or uses. This can be contrasted, for example, with grid computing. The most common type of modern thin client is a low-end computer terminal that concentrates solely on providing a GUI to the end user. The remaining functionality—in particular, the operating system—is provided by the server.

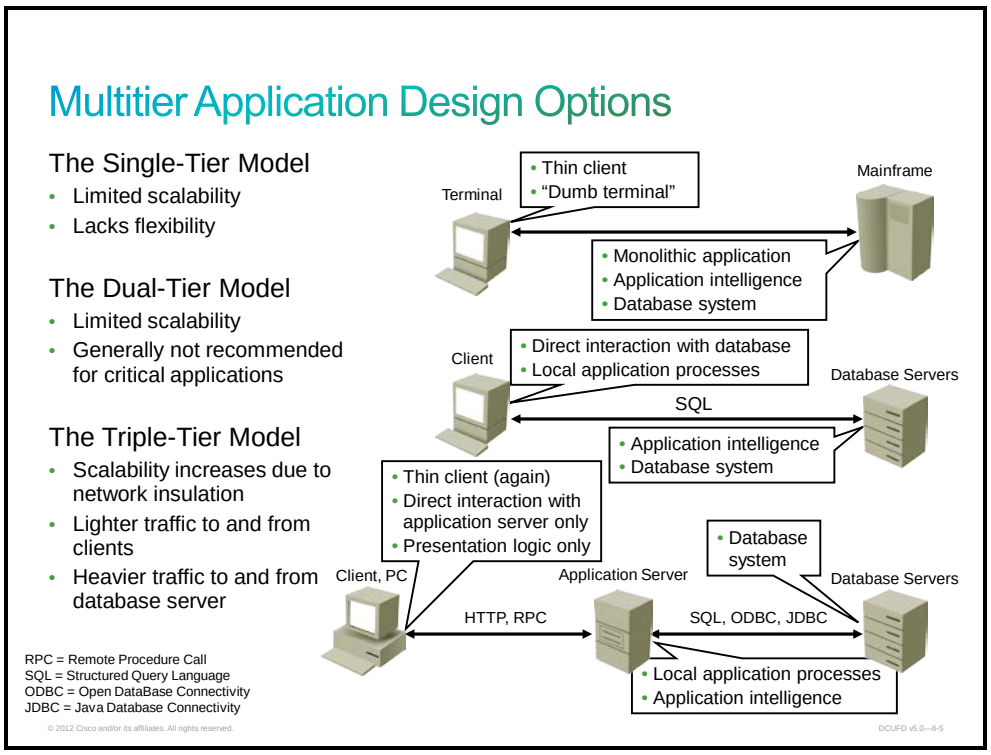
A thick client is a computer (client) in a client/server architecture or networks that typically provide rich functionality that is independent of the central server. Known as just a “client” or “fat client,” its name is contrasted to “thin client,” which describes a computer that is heavily dependent on the applications of a server. A fat client still requires at least a periodic connection to a network or central server, but is often characterized by the ability to perform many functions without that connection. In contrast, a thin client generally does as little processing as possible and relies on accessing the server each time input data needs to be processed or validated.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

A web application is an application that is accessed over a network such as the Internet or an intranet. The term can also mean a computer software application that is hosted in a browser-controlled environment or coded in a browser-supported language and reliant on a common web browser to render the application executable. Web applications are popular because of the ubiquity of web browsers and the convenience of using a web browser as a client, which is sometimes called a thin client. The ability to update and maintain web applications without distributing and installing software on, potentially, thousands of client computers is a major reason for their popularity, as is the inherent support for cross-platform compatibility. Common web applications include webmail, online retail sales, online auctions, wikis, and many other functions.



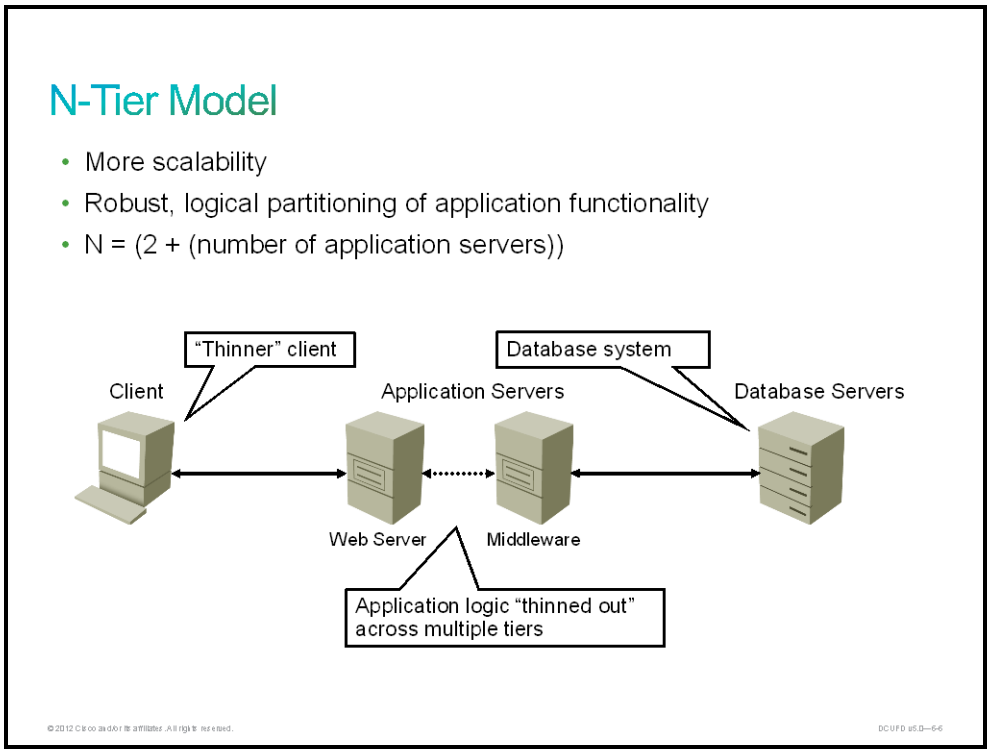
There are three main application design options:

- The single-tier model is the classic "dumb terminal" situation, where the client has very little intelligence and mostly only screen refreshes like characters, pixels, and so on. The old green screen is a classic example of this model. A thin client is a computer that depends heavily on some other computer (its server) to fulfill its traditional computational roles.
- The dual-tier model describes a client/server where the client has some processing power via an application-specific engine that resides on the PC. This model is distinctly different from the previous one-tier model because, from the perspective of the application, some logic has been distributed among two machines or tiers. The limitation of the dual-tier model is scalability. It is limited by the number of connections that the database server can manage. Also, there are no built-in limitations or business logic that limits how the client can query the database. A single client could dominate the resources of the database server with certain queries. Therefore, this model is generally not recommended for critical applications.

https://t.me/learningnets

certcollection.net

- The triple-tier model depicts the separation of the database server on the back end. The three-tier model is a distributed system that is characterized by clusters of autonomous functionality across multiple tiers. This model is more scalable and more resilient than the dual-tier model. Business logic can be applied in the application server tier to limit the extent and scope of queries from the clients. In this way, the application designers can better manage loads on the system.



The n-tier model describes the unbound set of potential application tiers. The number of tiers determines how scalable, portable, and manageable an application can be. As applications grow, different services that support that application can grow independently, as needed.

 © 2012 Cisco Systems, Inc.

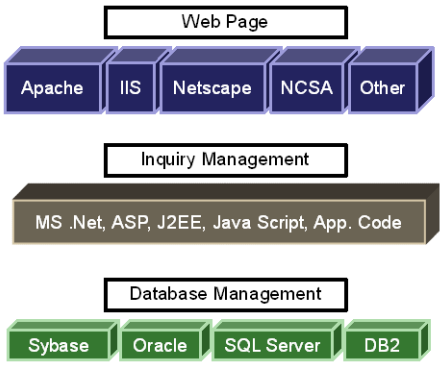
https://t.me/learningnets

certcollection.net

Web Services Model

Tasks performed by multiple hosts, each with a specific role:

- Web server
  - Serve pages
  - Static versus dynamic content
- Application server
  - Implements business logic
  - Manipulates data
- Database server
  - Accesses data store
  - Retrieves and writes efficiently and quickly



© 2012 Cisco and/or its affiliates. All rights reserved.

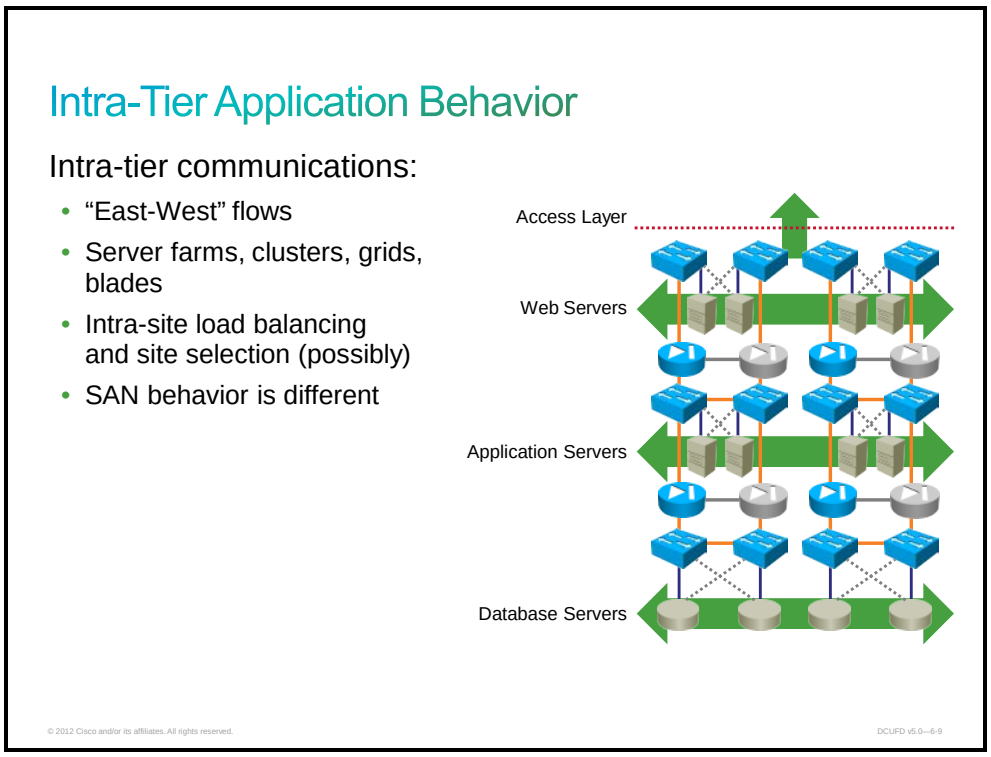
The web services model is an example of an n-tier system. Some examples of applications that might use this model include customer relationship management (CRM) or enterprise resource planning (ERP) applications. Web servers on the front end process individual user requests. These requests can be load-balanced to the web servers using a network-based solution. The web servers communicate directly with the application tier, and the application tier communicates with the database tier on the back end. Business logic can be applied against the user requests. For example, important users (such as large customers and executives) can be provided with priority service.

https://t.me/learningnets

certcollection.net

Application Tiering

This topic describes application tiering.



Server farms represent single tiers in the n-tier model, if not physically, then at least logically (but, usually, both). In any case, the types of clustering that occur at each tier must be understood so that the network can provide the proper resources and services to the application, such as common VLANs, private VLANs (PVLANs), and access control lists (ACLs).

Though most network-based functionality is currently aimed at the web server environment, many applications use a proprietary form of application server clustering technology that precludes Cisco from performing any hardware-based load-balancing or clustering assistance. This is also true in the database tier.

 © 2012 Cisco Systems, Inc.

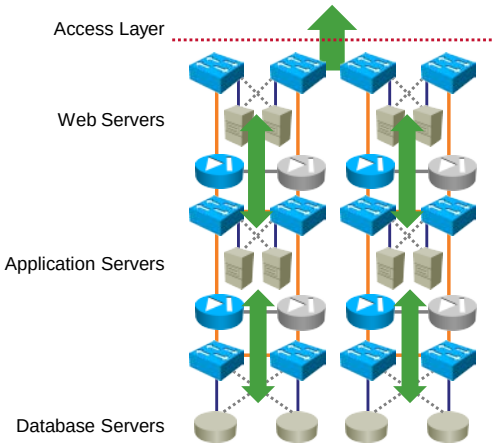
https://t.me/learningnets

certcollection.net

Inter-Tier Application Behavior

Inter-tier communications:

- "North-South" flows
  - Front-end to application
  - Application to back-end
  - Client to application
  - Client to back-end
- Security
- Protocol transparency



© 2012 Cisco and/or its affiliates. All rights reserved.

Clients communicate directly with web servers, and sometimes with application and database servers too. Web servers communicate with application servers. Application servers communicate with database servers. Database servers communicate with and across SANs.

This vertical communication mandates that the data center network must also be designed to accommodate inter-tier communications.

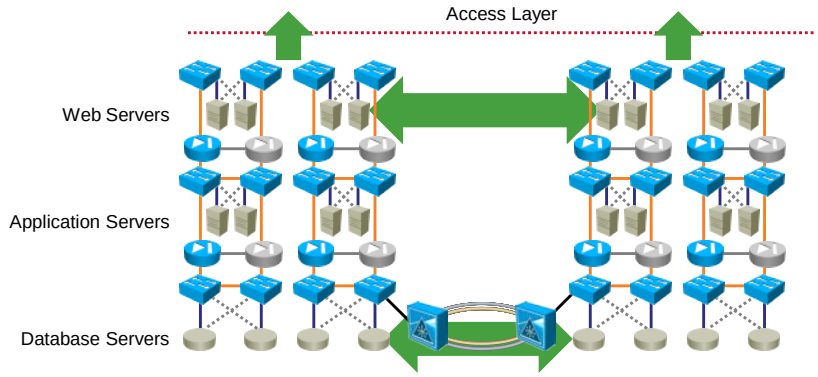
https://t.me/learningnets

certcollection.net

Multisite Application Behavior

Inter-site requirements:

- Site selection and load balancing
- Disaster recovery and business continuance
- Synchronous and asynchronous transactions



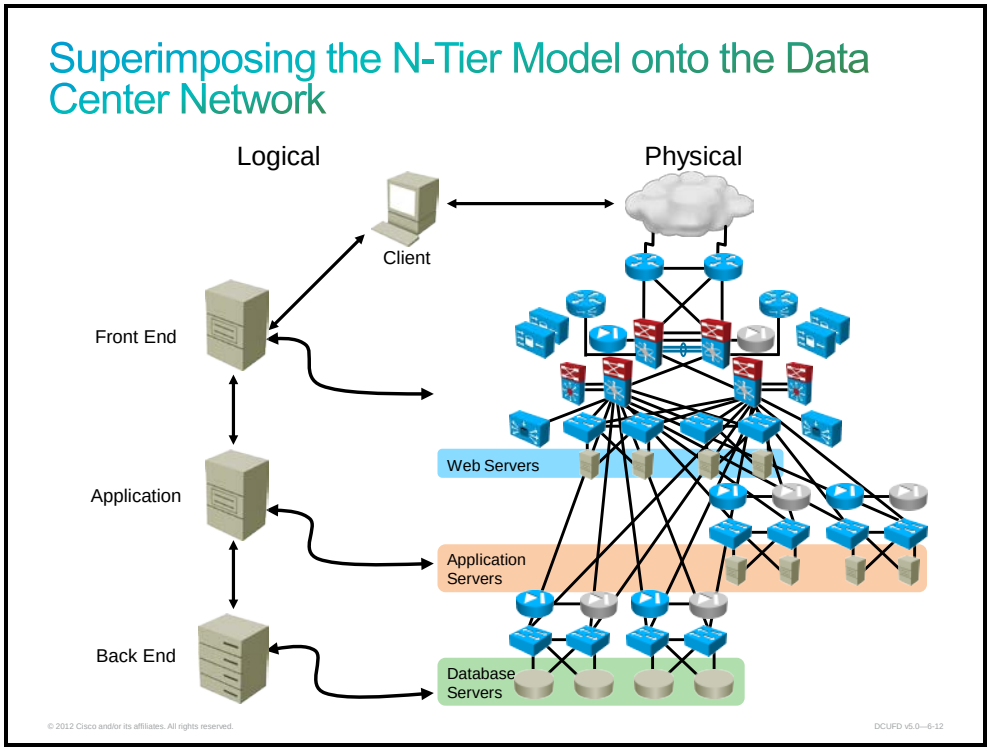
© 2012 Cisco and/or its affiliates. All rights reserved.

Distributed data centers offer globally distributed application-type services, which means that the communications that are required between these sites must be understood. Communication includes the server cluster heartbeats, server storage synchronous/asynchronous storage replication, and backup and failover functions between the primary and the secondary data center site.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net



Data centers are being built in such a way to satisfy the application requirements and traffic flow requirements.

Application performance dictates how much oversubscription you can afford between different access and aggregation layers.

For example, the type of flows, either inter-tier or intra-tier, determines if you will need an access switch (a lot of local traffic) or a fabric extender (FEX) (a lot of inter-tier traffic).

https://t.me/learningnets

certcollection.net

Wide-Area Application Optimization

This topic explains application optimization technologies.

Typical Distributed Enterprise

Expensive, distributed IT infrastructure:

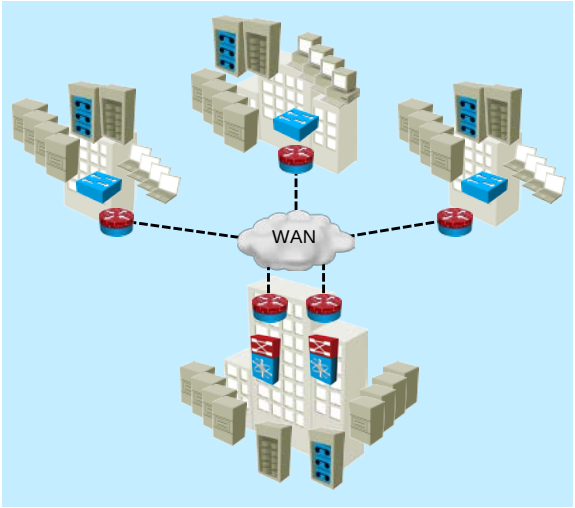
- File and print servers
- Email servers
- Tape backup

Application delivery problems:

- Congested WAN
- Bandwidth and latency
- Poor productivity

Data protection risks:

- Failing backups
- Costly offsite vaulting
- Compliance



© 2012 Cisco and/or its affiliates. All rights reserved.

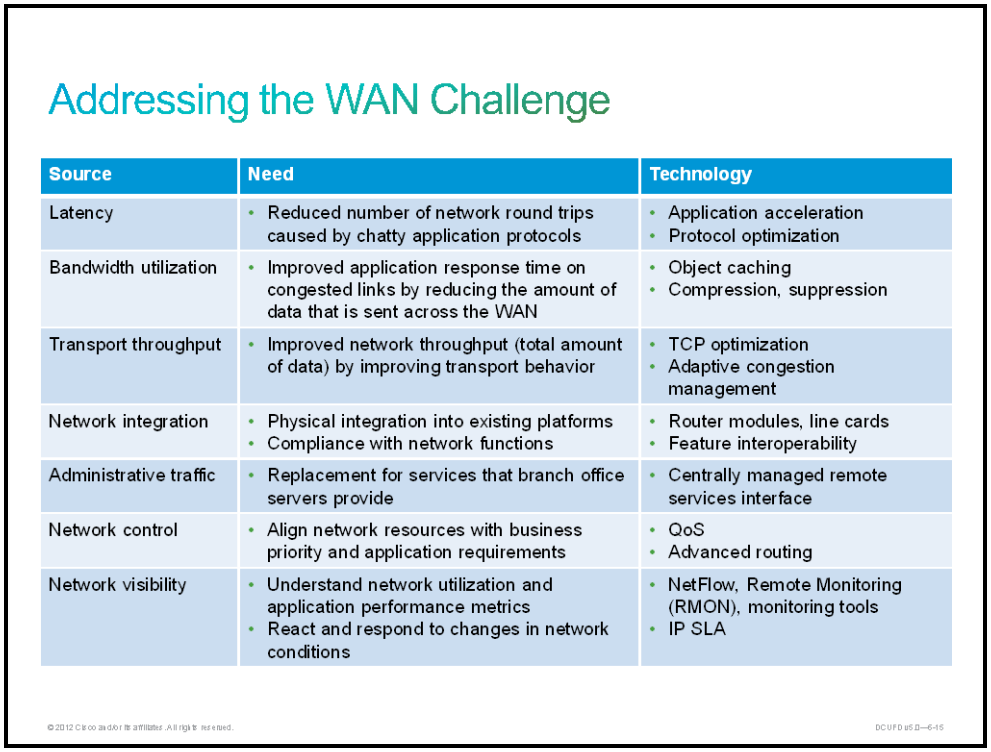
Many organizations have infrastructure silos in each of their remote, branch, and regional offices. These silos are typically carbon-copies of the infrastructure in the data center, including file servers, print servers, backup servers, application servers, email servers, web servers, and storage infrastructure. In any location where storage capacity is deployed with active data, that data must be protected with disk drives, tape drives, tape libraries, backup software, service with an offsite vaulting company, and perhaps even replication. The remote office infrastructure is costly to maintain.

The goal of the typical distributed enterprise is to consolidate as much of this infrastructure as possible into the data center, without overloading the WAN and without compromising the performance expectations of remote office users who are accustomed to working with local resources.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net



Latency is the most silent yet largest detractor of application performance over the WAN. Latency is problematic because of the volume of message traffic that must be sent and received. Some messages are very small, but even with substantial compression and flow optimizations, these messages must be exchanged between the client and the server to maintain protocol correctness, data integrity, and so on. The best way to mitigate latency is to deploy intelligent protocol optimizations, also known as application acceleration, in the remote office. This is done on a device that understands the application protocol well enough to make decisions on how best to manage application traffic as it occurs and, in many cases, can closely mimic the performance of a local server. On a per-message basis, the application accelerator examines messages to determine whether they can be suppressed or locally processed. If the request is for data, the application accelerator determines if the data is best served from the cache (if the object is valid, the user is authenticated, and the appropriate state is applied against the object on the origin server) or if a message must be sent to the origin server to maintain proper protocol semantics.

Bandwidth utilization also harms application performance. Transferring a file multiple times can consume significant WAN bandwidth. If a validated copy of a file or other object is stored locally in an application cache, it can be served to the user without using the WAN. Application caching is typically tied to an application accelerator and is specific to that application, but there are compression techniques that can be applied at the transport layer that are application-agnostic. One of these techniques is standards-based compression. Another technique is called data redundancy elimination (DRE), which is an advanced form of suppressing the transmission of redundant network byte streams. Compression and application caching provide another way to improve application performance by minimizing the amount of data that must traverse the network. Minimizing the amount of data on the network improves response time and leads to better application performance, while also freeing up network resources for other applications.

https://t.me/learningnets

certcollection.net

Another barrier to application performance in a WAN environment is transport throughput. Application protocols run on top of a transport mechanism that provides connection-oriented or non-connection-oriented delivery of data. In many cases, enterprise applications use TCP for its inherent reliability. Although it is reliable, TCP presents performance obstacles of its own. If TCP could be optimized to perform better in WAN environments, then application throughput, response time, and the user experience would all show improvement, due to better utilization of existing network capacity and better response to network conditions.

Two factors should be considered for all consolidation-enabling solutions. The first factor is network integration. Consolidation solutions should not disrupt the operation of existing network features such as quality of service (QoS), access lists, NetFlow, and firewall policies. By integrating with the network in a logical manner—that is, by maintaining service transparency (preserving information in packets that the network needs to make intelligent feature-oriented decisions) —fundamental network layer optimizations can continue to operate in the face of application acceleration or WAN optimization. Physical integration allows such technology to be directly integrated into existing network devices, thereby providing a far more effective total cost of ownership (TCO) and return on investment (ROI) model.

When possible, administrative services such as print services should be centrally managed but locally deployed in remote sites. This keeps such administrative traffic from needing to traverse the WAN.

The network should be aligned with business priority and application requirements to ensure the appropriate handling of traffic. QoS, for example, allows network administrators to configure network behavior in specific ways for specific applications. Because all applications are not created equal, the network must be prepared to process traffic in different ways based on how the application needs to be managed. This involves classification of data (seeing what application it is and who is talking to who, among other metrics), prequeuing operations (immediate actions, such as marking, dropping, or policing), queuing and scheduling (ensuring that the appropriate level of service and capacity are assigned to the flow), and postqueuing optimizations (such as link fragmentation and interleaving, and packet header compression). This set of four functions is known as the QoS Behavioral Model, which relies on visibility (service transparency) if acceleration technology is deployed to fully function. Also, the network should be able to make path routing decisions (advanced routing) in real time to ensure that the right path is taken for the right flows. This includes policy-based routing (PBR) and Optimized Edge Routing (OER).

Finally, the network should be visible. That is, administrators need to know how the network is performing, how the network is being used, and when network characteristics are performing as expected. Technologies such as NetFlow and collection or analysis tools allow administrators to see how the network is being used, top talkers, and so on. Functions such as Cisco IOS IP Service Level Agreements (IP SLAs) allow the network to alert administrators when conditions exceed thresholds and, furthermore, allow the network to react when such events occur.

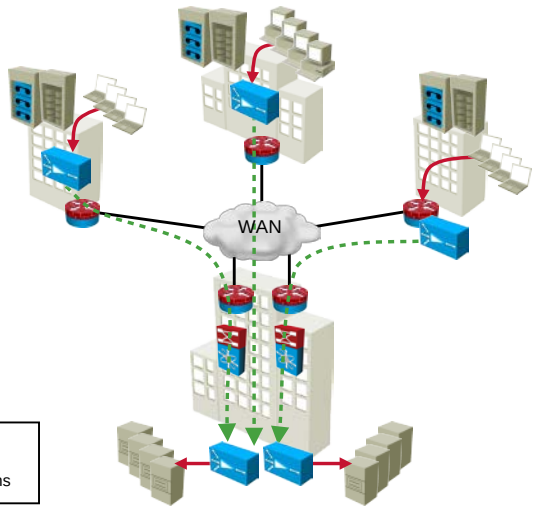
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

Cisco WAAS Overcomes the WAN Performance Barrier

Cisco WAAS leverages a hardware footprint (Cisco WAE) in the remote office and the data center to overcome application performance problems in WAN environments.



© 2012 Cisco and/or its affiliates. All rights reserved.

Cisco WAAS is a solution that overcomes the challenges that are presented by the WAN. Cisco WAAS is a software package that runs on the Cisco Wide Area Application Engine (WAE), which transparently integrates with the network to optimize applications without client, server, or network feature changes.

A Cisco WAE is deployed in each remote office, regional office, and data center of the enterprise. With Cisco WAAS, flows that are to be optimized are transparently redirected to the Cisco WAE, which overcomes WAN restrictions, including bandwidth disparity, packet loss, congestion, and latency. Cisco WAAS enables application flows to overcome restrictive WAN characteristics to enable the consolidation of distributed servers, save WAN bandwidth, and improve the performance of applications that are already centralized.

https://t.me/learningnets

certcollection.net

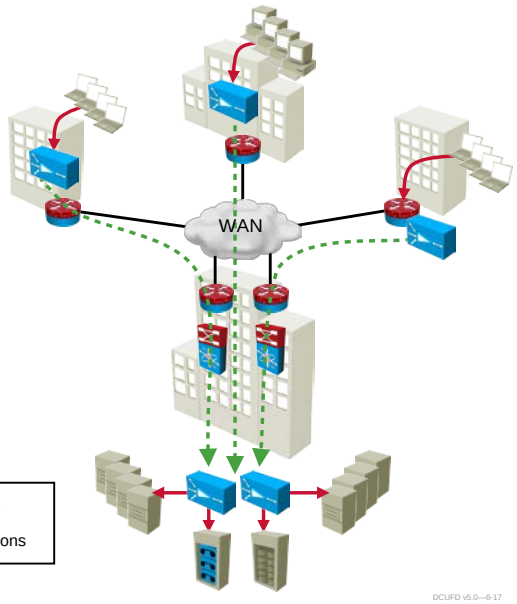
Cisco WAAS Enables Consolidation

Consolidation benefits:

- Remove costly branch servers
- Centralize storage
- Centralize data protection
- Conserve WAN resources

Improvements:

- Application acceleration
- WAN optimization
- Local infrastructure services



© 2012 Cisco and/or its affiliates. All rights reserved.

Cisco WAAS helps consolidate infrastructure from remote offices into the data center. Cisco WAAS has numerous features:

- Integrate transparently into the existing infrastructure
- Understand application protocols and how to optimize those applications
- Provide compression and flow optimizations to improve delivery of data that must traverse the WAN
- Simplify consolidation by providing policy-based configuration and automatic discovery

Aside from cost savings, the primary goal of infrastructure consolidation is to give users the same level of access that is available with a local infrastructure.

Maintaining performance while enabling consolidation entails various services:

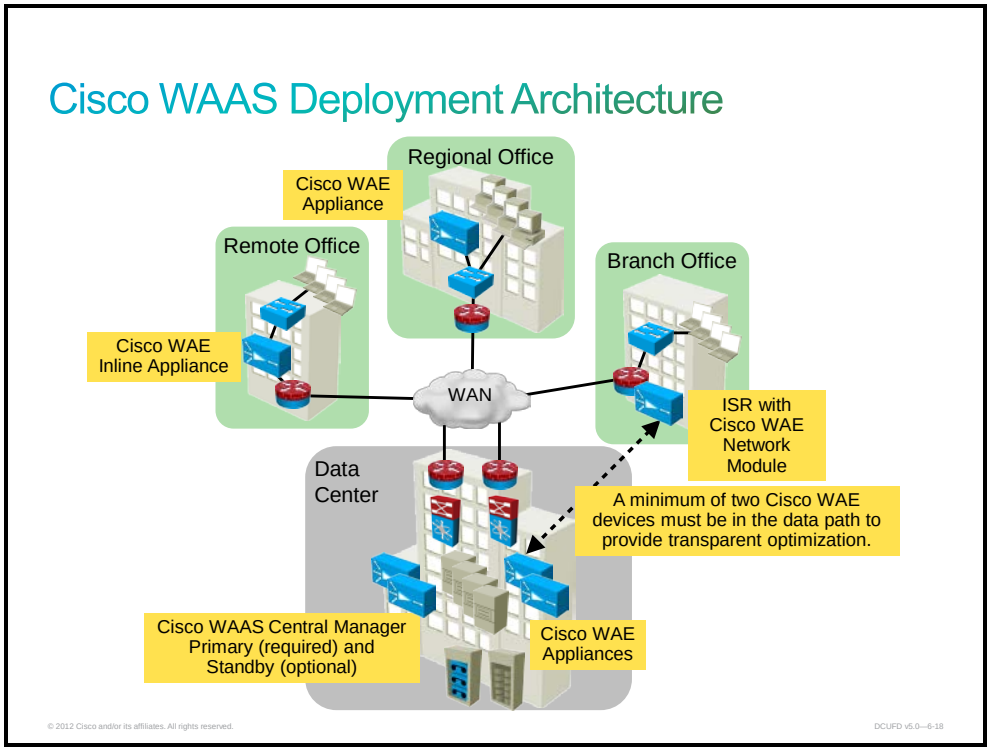
- Application-specific acceleration (file and print services)
- WAN optimizations such as transport flow optimization, DRE, and persistent Lempel-Ziv (LZ) compression

With Cisco WAAS, Cisco WAE devices automatically discover each other to minimize the administrative burden.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net



Cisco WAE devices are deployed at network entry and exit points of WAN connections. If multiple entry and exit points exist, you can deploy a single Cisco WAE that optimizes both connections by sharing the interception configuration across those entry and exit routers. To provide and support optimizations, Cisco WAAS requires that devices be deployed in two or more sites. To support redundancy, more than one Cisco WAE is typically deployed in the data center. Cisco WAE devices must also be deployed to host the Cisco WAAS Central Manager application, which can be made highly available by using two Cisco WAE devices. To provide transparent optimizations, Cisco WAAS requires two devices in the path of the connection to be optimized.

As shown in the figure, Cisco WAE devices can either be standalone appliances or network modules that integrate physically into the integrated services router (ISR).

https://t.me/learningnets

certcollection.net

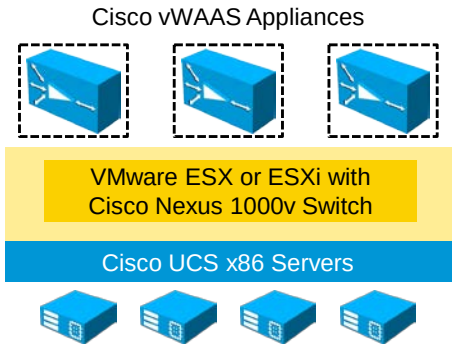
Cisco Virtual WAAS

Cisco vWAAS is cloud-ready WAN optimization:

- Virtual appliance that accelerates applications that are delivered from private and virtual private cloud infrastructures
- Runs on the VMware ESXi hypervisor and Cisco UCS x86 servers

Cisco vWAAS can be deployed in two ways:

- Transparently at the WAN network edge
- Within the data center along with application servers



© 2012 Cisco and/or its affiliates. All rights reserved.

Cisco Virtual WAAS (vWAAS) is a cloud-ready WAN optimization solution. Cisco vWAAS is a virtual appliance that accelerates business applications that are delivered from private and virtual private cloud infrastructures, helping to ensure an optimal user experience. Cisco vWAAS runs on the VMware ESXi hypervisor and Cisco Unified Computing System (UCS) x86 servers, providing an agile, elastic, and multitenant deployment.

Cisco vWAAS can be deployed in two ways:

- Transparently at the WAN network edge using out-of-path interception technology such as Web Cache Control Protocol (WCCP), similar to the deployment of a physical Cisco WAAS appliance
- Within the data center along with application servers, using a virtual network services framework based on Cisco Nexus 1000V Series Switches to offer cloud-optimized application service in response to instantiation of application server virtual machines

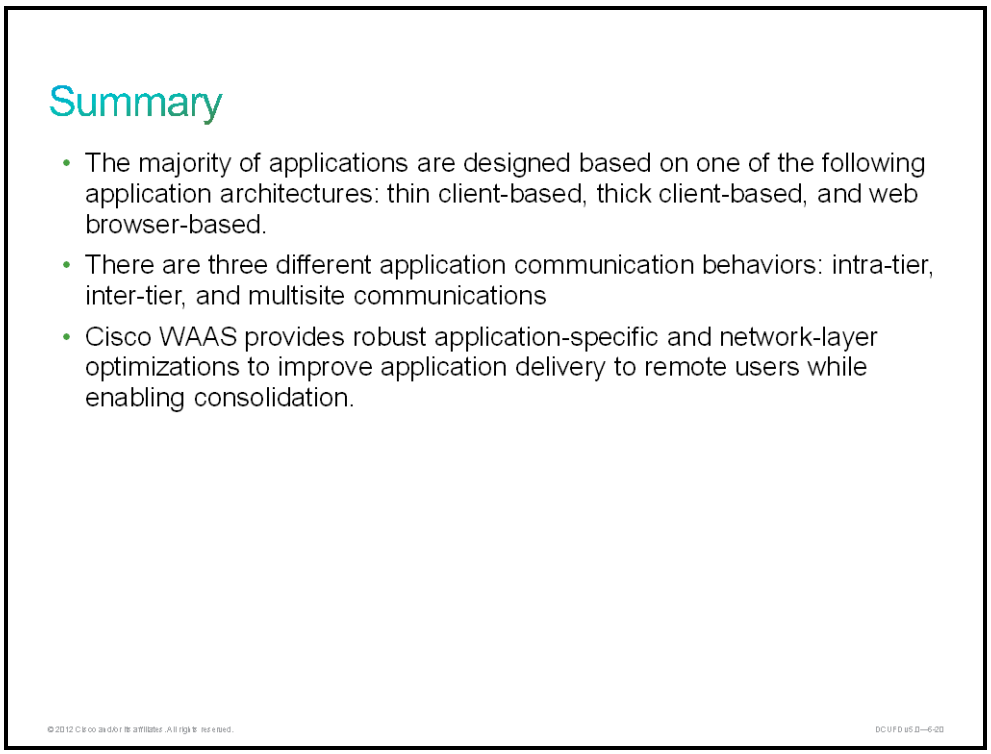
© 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

Summary

This topic summarizes the primary points that were discussed in this lesson.



https://t.me/learningnets

certcollection.net

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

Lesson 2

Designing Application Services

Overview

This lesson explains the effect of the server load-balancing technologies on the data center design. This lesson also explains Cisco Application Control Engine (ACE) Module and Cisco ACE appliance deployment topologies, including routed, bridge, and one-arm modes, as well as direct server return deployment topologies. In this lesson, you will also learn about Cisco ACE Module support for Secure Sockets Layer (SSL) protocol processing.

Objectives

Upon completing this lesson, you will be able to design the network infrastructure for application services. This ability includes being able to meet these objectives:

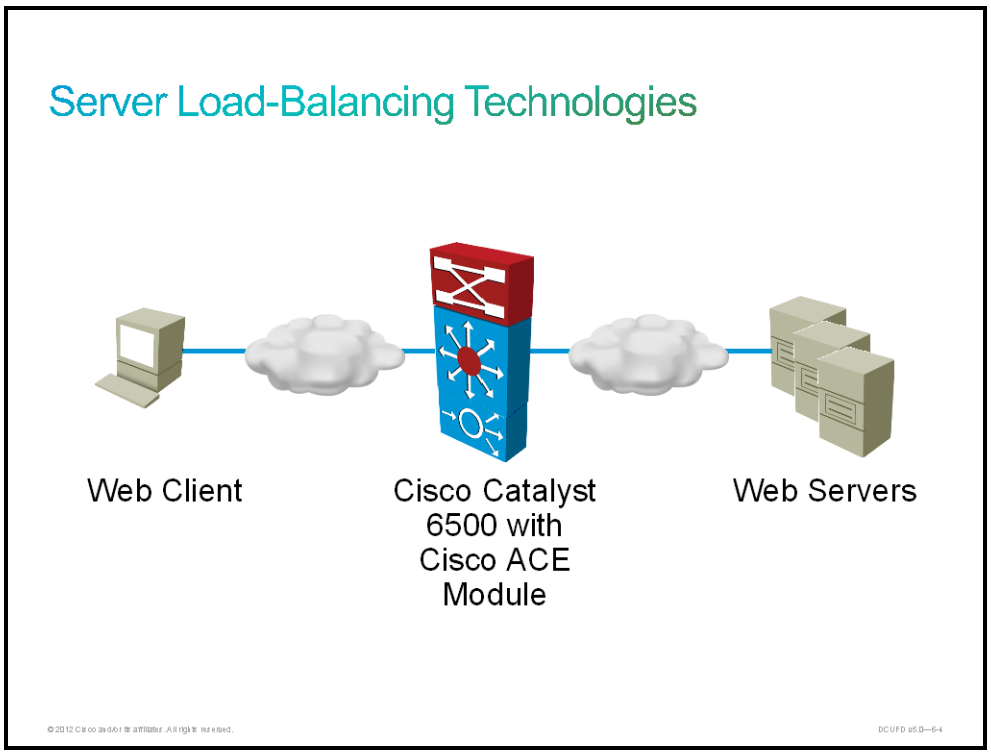
- Explain server load-balancing technologies
- Add application services to an existing data center
- Explain contexts
- Design secure application load-balancing solutions

https://t.me/learningnets

certcollection.net

Server Load-Balancing Technologies

This topic describes server load-balancing technologies.



Load balancing is a computer networking methodology to distribute workload across multiple computers or a computer cluster, network links, CPUs, disk drives, or other resources, to achieve optimal resource utilization, maximize throughput, minimize response time, and avoid overload. Using multiple components with load balancing, instead of a single component, may increase reliability through redundancy. The load-balancing service is usually provided by dedicated software or hardware, such as a multilayer switch or a Domain Name System (DNS) server.

Server load balancing is the process of deciding to which server a load-balancing device should send a client request for service. For example, a client request may consist of an HTTP GET for a web page or an FTP GET to download a file. The job of the load balancer is to select the server that can successfully fulfill the client request and do so in the shortest amount of time without overloading either the server or the server farm as a whole.

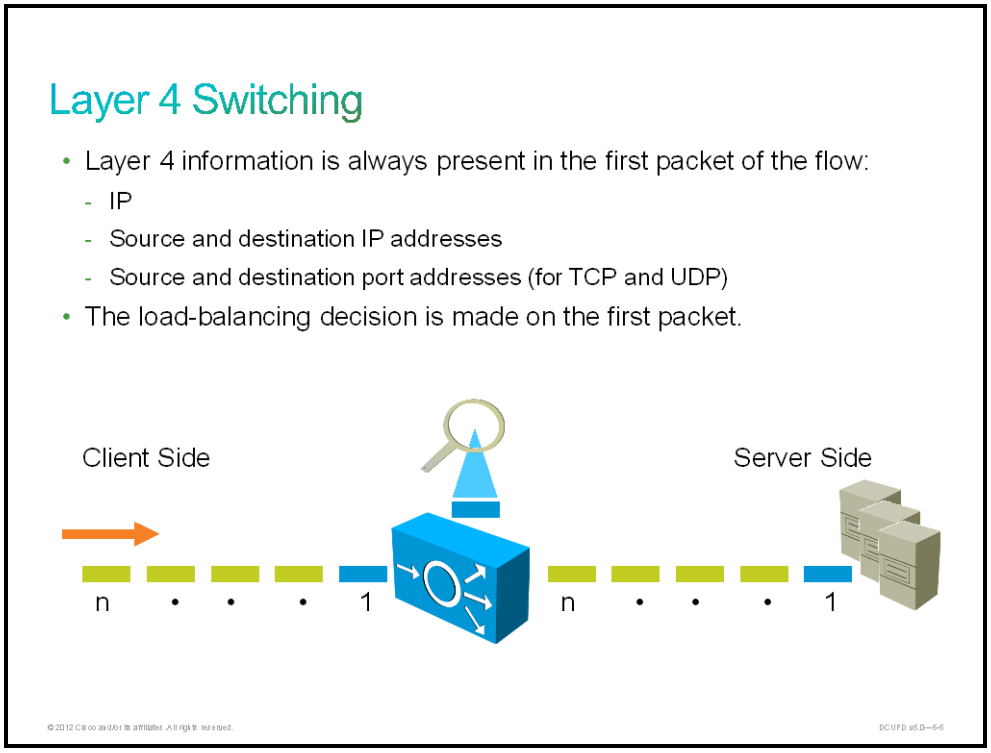
Depending on the load-balancing algorithm or predictor that you configure, the Cisco ACE Module performs a series of checks and calculations to determine the server that can best service each client request. The Cisco ACE Module bases server selection on several factors, including the server with the fewest connections with respect to load, source or destination address, cookies, URLs, or HTTP headers.

The Cisco ACE Module resides in a Cisco Catalyst 6500 chassis. The figure shows a basic network where a Cisco Catalyst 6500, equipped with a Cisco ACE Module, switches and routes traffic between a web client and a web server.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net



Layer 4 information in the packet includes the following fields:

- **IP:** This field is used to differentiate between the higher-level protocols that are supported by IP, such as UDP and TCP.
- **Source and destination IP addresses:** The IP address of the transmitting system and the IP address of the intended recipient.
- **Source and destination port:** The port number for the transmitting system and the port number for the intended recipient.

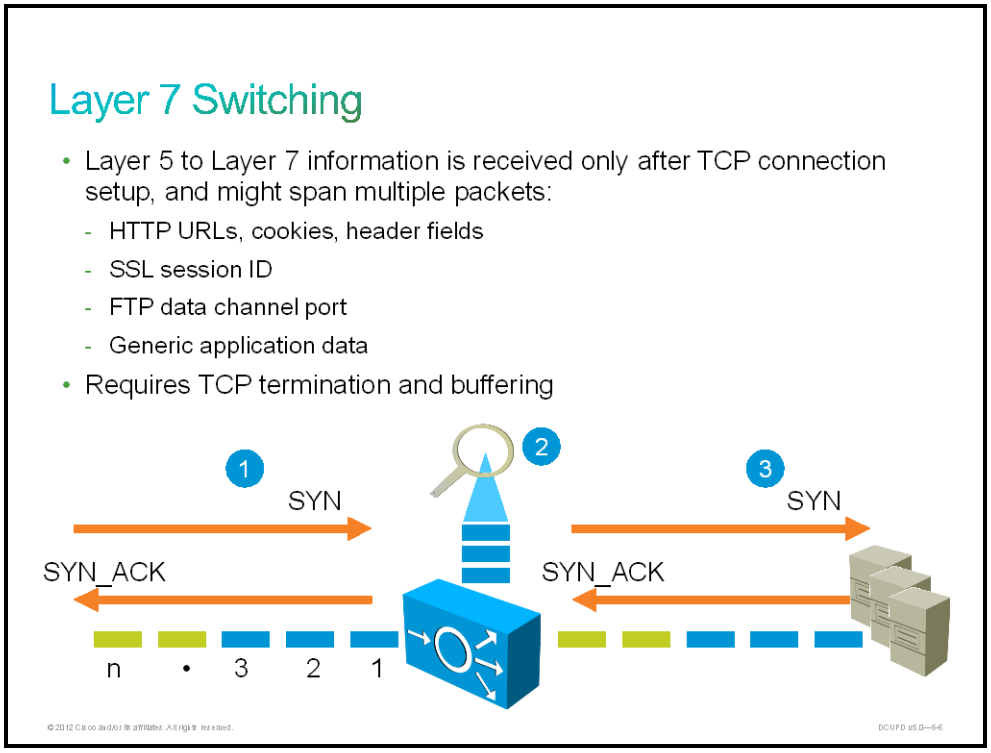
Note	Port numbers are used to direct the IP traffic to a particular application process, such as a web client or server. Well-known port numbers are defined for most IP-based services. For example, port 80 is used for HTTP.

Layer 4 content-switching decisions can be based on any of the Layer 4 fields listed here. With TCP connections, the Layer 4 information is consistent for all packets in the connection. The Layer 4 information is often said to define a flow, which is the communication path for a particular connection.

The figure shows a flow of packets coming from the client side of the network to a Cisco ACE Module. The Cisco ACE Module examines the first packet in a new flow or connection and a Layer 4 switching decision is made for the flow as a whole. The content switch makes this decision and then records the flow parameters and the switching decision. This table of switching decisions is used to switch every subsequent packet in the flow. Information is removed from the switching table when a connection is closed. For Layer 4 switching of TCP packets, these decisions are normally made based on SYN and FIN packets and are done at TCP connection setup and termination. Reset (RST) packets are also analyzed because they are used to refuse a connection when it is requested or to abort an existing connection.

https://t.me/learningnets

certcollection.net



Layer 7 information is available only after application data has been transmitted, but transmission requires that the TCP connection be fully functional, which causes a dilemma: A server needs to respond to the client to fully start the TCP connection before the client sends the Layer 7 information that the content switch needs to choose the server.

The content switch solves this problem by buffering client data and temporarily acting as a server. To do this, the content switch responds to the incoming SYN packet with its own SYN_ACK. The content switch then buffers packets until it has enough Layer 7 information to make a load-balancing decision.

After a destination server is selected, the content switch makes a connection to the server on behalf of the client. To establish the TCP connection to the server, a SYN packet is sent to the server and then the Cisco ACE Module waits for the SYN_ACK packet to be sent from the server. At this point, all buffered packets that were received from the client are sent to the server.

After the buffered packets have been sent, the two TCP connections can be spliced together by the content switch. This splicing is performed by receiving packets from one connection and retransmitting them to the other.

Because there are two different TCP connections from the content switch—one to the client and one to the server—there are probably two sets of sequence numbers in use, one on each connection. The content switch translates the sequence numbers from one connection to the other.

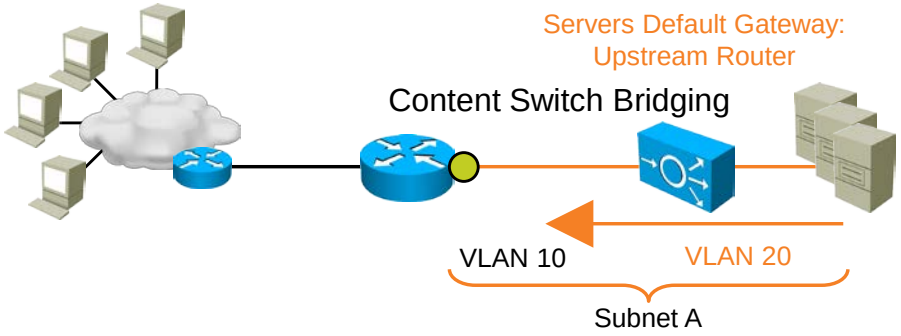
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

Cisco ACE Module Bridged Mode

- Client and server VLANs are part of the same IP subnet.
- The Cisco ACE Module uses the ARP table to track which VLAN contains which physical devices.
- Servers in this environment use the IP address of the upstream router interface as their default gateway.
- Static and dynamic routing (Open Shortest Path First [OSPF]) can be used in bridged mode.



© 2012 Cisco and/or its affiliates. All rights reserved.

The Cisco ACE Module can be configured in bridge mode. In this mode, the client and server VLANs are part of the same IP subnet. The Cisco ACE Module uses an Address Resolution Protocol (ARP) table to track which VLAN contains which physical devices.

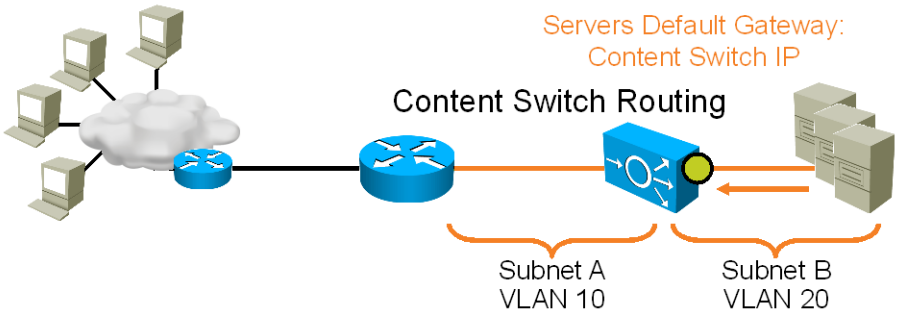
In this figure, VLAN 10 is used as the client-side VLAN, while VLAN 20 is the server-side VLAN. The same IP subnet is used on both VLANs. The physical port that is attached to the upstream router is assigned to VLAN 10. Physical ports that are connected to the servers are assigned to VLAN 20. The servers in a bridge mode environment are configured to use the IP address of the upstream router interface as their default gateway.

https://t.me/learningnets

certcollection.net

Cisco ACE Module Routed Mode

- Client and server VLANs are part of different subnets (VLAN 10 as the client-side VLAN, and VLAN 20 as the server-side VLAN).
- Servers in this environment use the IP address of the ACE Module as their default gateway.
- Static and dynamic routing (Internal Border Gateway Protocol [IBGP]) can be used in routed mode.



© 2012 Cisco and/or its affiliates. All rights reserved.

The Cisco ACE Module can be configured in routed mode. In this mode, the client and server VLANs are part of different IP subnets. This breaks the old rule of one VLAN per subnet and one subnet per VLAN.

In this figure, VLAN 10 is configured as the client-side VLAN, while VLAN 20 is the server-side VLAN. Different IP subnets are associated with each VLAN. The physical port that is attached to the upstream router is assigned to VLAN 10. Physical ports that are connected to the servers are assigned to VLAN 20. The servers in a routed mode environment are configured to use the IP address of the Cisco ACE Module as their default gateway.

 © 2012 Cisco Systems, Inc.

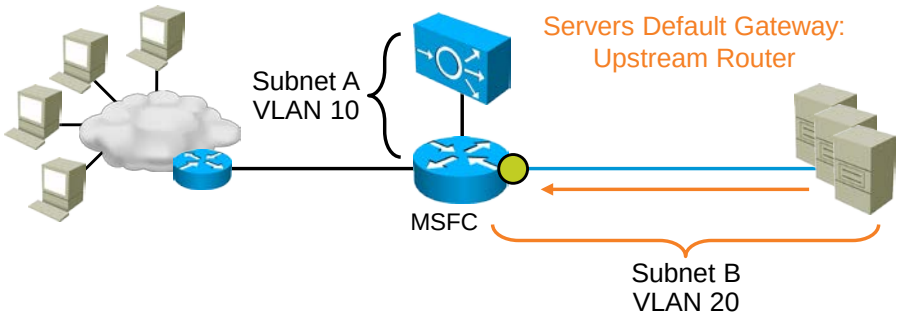
https://t.me/learningnets

certcollection.net

Cisco ACE Module One-Arm Mode

There are two ways to get the traffic to flow through the Cisco ACE Module:

- SNAT
- PBR



© 2012 Cisco and/or its affiliates. All rights reserved.

The one-arm mode removes the Cisco ACE Module from a position that is directly in the transit path for all traffic to the server farms. An advantage of this configuration is that the Cisco ACE Module does not need to process traffic that is not affected by Cisco ACE Module features. In this figure, VLAN 10 is used for traffic between the Cisco ACE Module and the Multilayer Switch Feature Card (MSFC), while VLAN 20 is used for traffic to the server farms. A VLAN 10 interface is configured on the MSFC, and an IP address from Subnet A is configured on the Cisco ACE Module. Additional IP addresses from Subnet A are used to configure the virtual server IP addresses. A VLAN 20 interface is configured on the MSFC and is used by the servers as their default gateway.

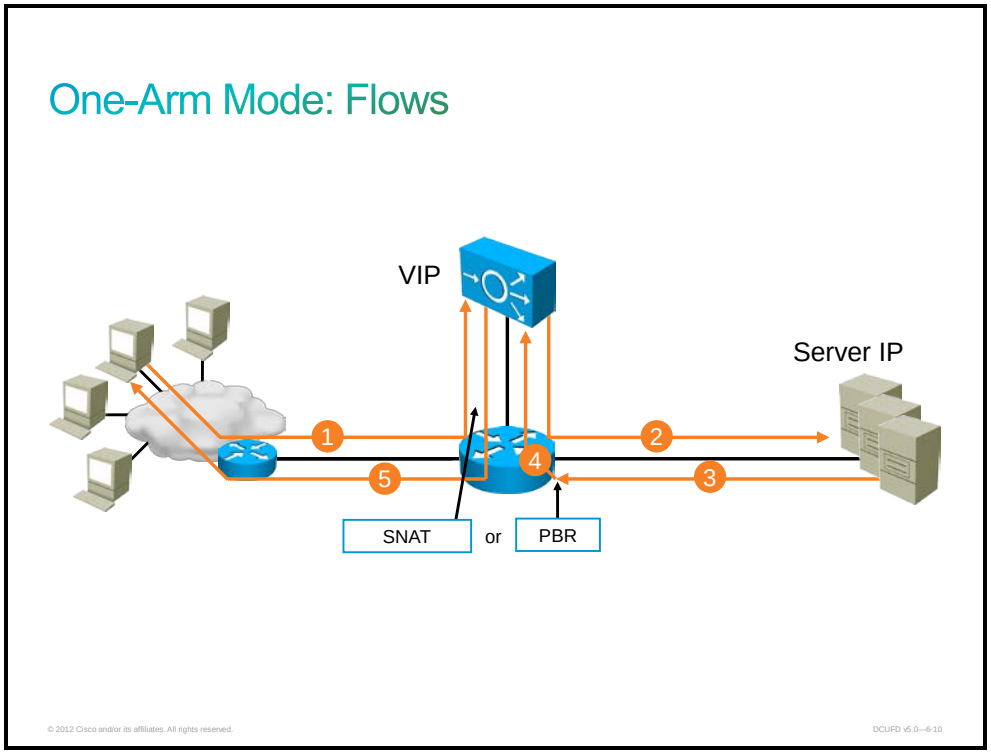
Note A Cisco ACE Module in one-arm mode has only one VLAN.

Return traffic that is generated by the servers in response to load-balanced requests is still needed by the Cisco ACE Module for full functionality. Getting this traffic to flow through the ACE Module is more complicated than with an inline configuration. There are two ways to address this situation:

- **Source Network Address Translation (SNAT):** Source-based NAT is configured by creating a pool of IP addresses. Client IP addresses are translated to IP addresses from the client pool. These translated addresses are used as the source address in the source packet that is sent to the server.
- **Policy-based routing (PBR):** PBR is a router feature that is available on Cisco IOS-based routers, including the Cisco Catalyst 6500 MSFC. PBR allows the router to be configured to select a next hop for a packet based on a configured policy. This policy overrides the routing decision that would have been made by consulting the routing database. A routing policy is attached to the ingress interface on the router. Access lists can be used to limit the traffic to which the policy is applied. For example, web responses that are sent to clients can be load balanced and redirected via PBR, while Simple Network Management Protocol (SNMP) responses from the servers are routed normally.

https://t.me/learningnets

certcollection.net



The traffic flow for load-balanced requests is shown in this figure. Packets are processed as follows:

Step 1 Traffic from the client to the virtual IP (VIP) is routed normally by the MSFC.

Step 2 Traffic from the Cisco ACE Module to the server is routed normally by the MSFC. If SNAT is used, the source IP address is in the client NAT pool. Otherwise, the source IP address remains the client IP address.

Step 3 Traffic from the server is returned to the MSFC because the MSFC is the server default gateway.

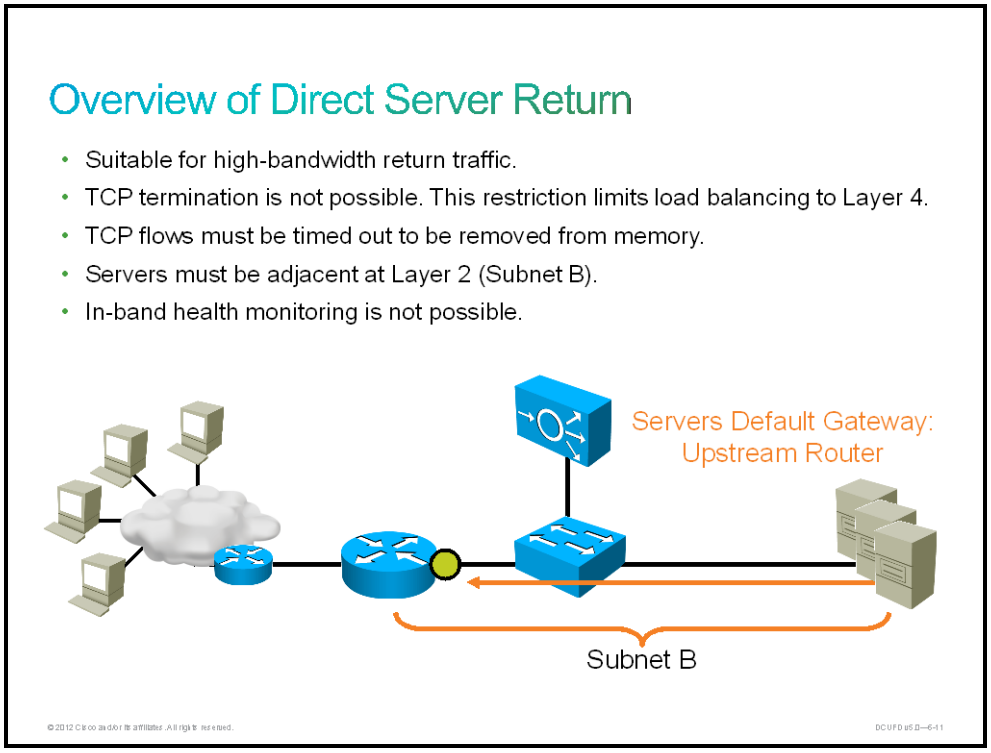
Step 4 If SNAT is used, the destination IP address in the server response is routed normally to the Cisco ACE Module. If SNAT is not used, PBR must be used on the MSFC interface that is used as the server default gateway. The policies that are configured must match any traffic that is being sent in response to a load-balanced request. The IP address that is specified for the Cisco ACE Module is set as the next-hop address by PBR.

Step 5 Traffic from the Cisco ACE Module to the client is routed normally by the MSFC. If SNAT is used, the Cisco ACE Module translates the destination IP address from the NAT pool IP address to the client IP address. If PBR is used, the Cisco ACE Module does not need to modify the destination IP address because the client IP address is already in the packet.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net



A variation of one-arm mode is a direct server return. The figure shows the architecture of this variation.

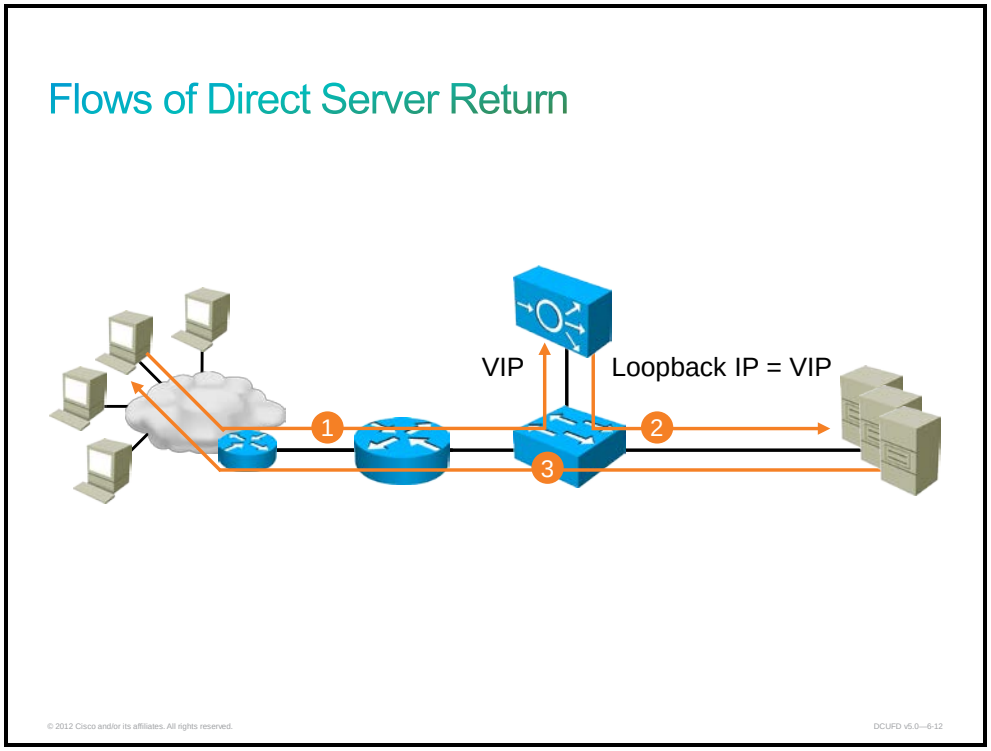
The Cisco ACE Module and the servers are placed in the same VLAN and IP subnet. An interface on that VLAN is defined on the MSFC and is the default gateway for the Cisco ACE Module and the servers. NAT is turned off for the server destination address. Return traffic does not flow through the Cisco ACE Module but returns directly to the client.

The advantage of a direct server return is that web servers can return higher-bandwidth traffic than can be processed by the Cisco ACE Module. Because the return traffic is not processed by the Cisco ACE Module, these restrictions apply:

- TCP termination is not possible. This restriction limits load balancing to Layer 4.
- TCP flows must be timed out to be removed from memory.
- Servers must be adjacent at Layer 2.
- In-band health monitoring is not possible when using this logical topology.

https://t.me/learningnets

certcollection.net



The traffic flow for load-balanced requests is shown in this figure. The packets are processed as follows:

Step 1 Incoming client requests are routed to the server VLAN. The packet is switched to the Cisco ACE Module.

Step 2 The Cisco ACE Module rewrites the Layer 2 destination MAC address and returns the packet to the switch processor. The packet is switched to the server. The server uses a loopback interface that is configured with the VIP address so that the server accepts a packet destined for the VIP.

Step 3 The server responds directly to the client. This traffic is routed normally because the MSFC is the default gateway for the server.

When no more traffic is generated by the client on this TCP connection, the connection goes idle. After the idle timeout, the Cisco ACE Module removes the connection from its session table.

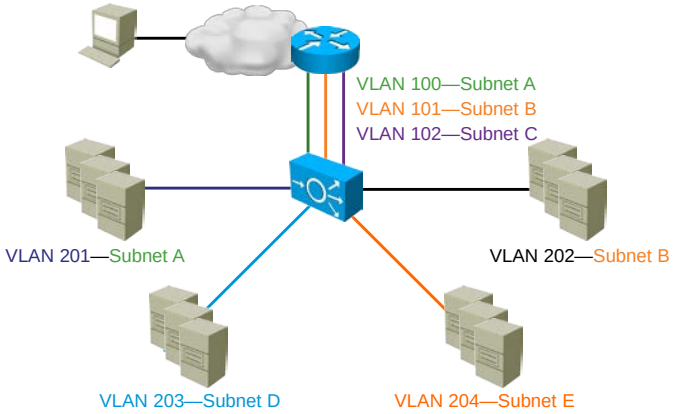
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

Mixed Modes

- The Cisco ACE Module can handle mixed modes between contexts—bridged and routed.
- Interfaces that are used in bridged contexts cannot be shared.



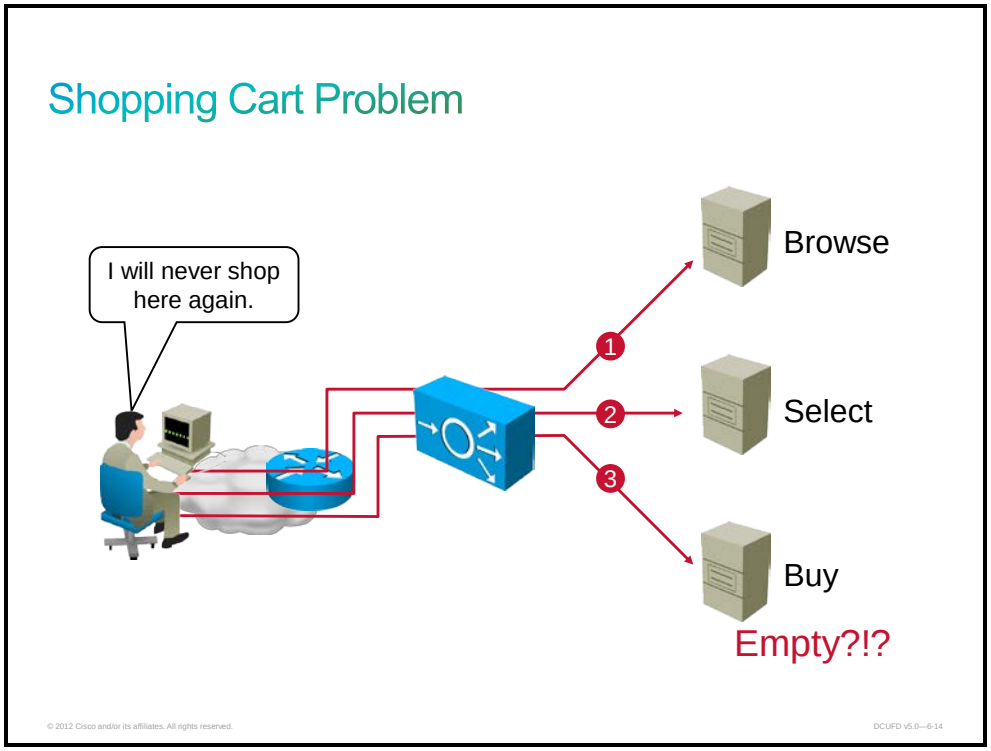
© 2012 Cisco and/or its affiliates. All rights reserved.

The Cisco ACE Module can manage multiple pairs of VLANs and mixed modes. This figure shows one Cisco ACE Module managing several VLANs. The following mode configurations are possible:

- Subnet A bridged between VLAN 100 and VLAN 201
- Subnet B bridged between VLAN 101 and VLAN 202
- Subnet C on VLAN 102 routed to Subnet D on VLAN 203 or Subnet E on VLAN 204

https://t.me/learningnets

certcollection.net



Many web applications require multiple interactions between the client and the server. The challenge with these applications is to distinguish which client is which when a request is received by the server. Often the solution is to establish a session ID that is transmitted by the client with each request. This session ID is then used by the server to retrieve stored information about former interactions with this client.

Load-balancing applications, such as the Cisco ACE Module, create a potential problem with this approach to multiple interactions. For example, the shopper in this figure is using an e-commerce application to purchase an item from a website. Simple round-robin load balancing can result in the following sequence of interactions:

Step 1 The shopper retrieves a page with details about a product of interest. Load balancing assigns this request to the top server. The server creates a session ID and sends it along with the rest of the response to the client.

Step 2 The shopper presses the “Buy Now” button. The resulting request contains the session ID and is assigned to the middle server. A record is created in the shopping cart database, associating the item that was selected to the session ID. A page is built and returned to the client with confirmation of the buying decision and checkout link.

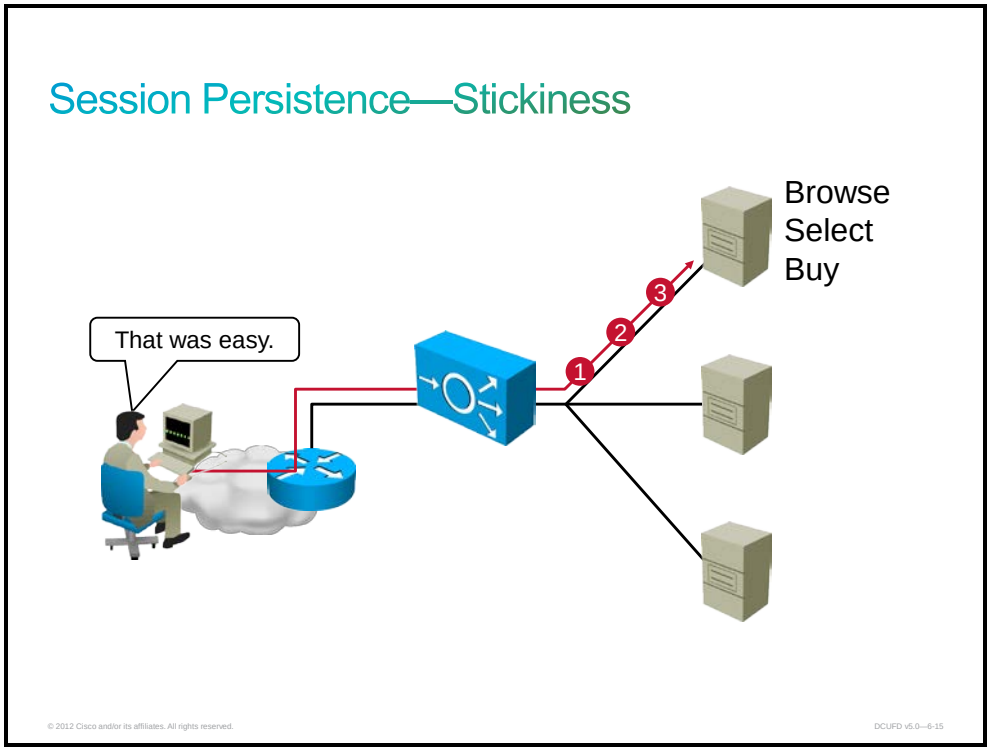
Step 3 The shopper presses the checkout link. The resulting request is assigned to the bottom server. This server uses the session ID in the client request to retrieve information about what items are in the shopping cart. Finding no entries in the shopping cart database, the server includes an indication to the client that the cart is empty.

Note The session ID can be carried in various places, including cookies and the URL.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net



The solution to the shopping cart problem and similar problems is session persistence, also known as “stickiness.” Stickiness modifies the content-switching decision process. When a connection first matches certain configured criteria, an entry is made in the “sticky” database by the Cisco ACE Module. This entry stores the connection criteria that were matched and the results of the load-balancing decision. Stickiness criteria can be matched on traffic in either direction. For example, if a cookie is being used for stickiness, the Cisco ACE Module can match the set cookie portion of the response from the server or in the cookie portion of the request from the client.

The shopper in this figure is using an e-commerce application to purchase an item from a website. With stickiness, the following sequence of interactions can result:

Step 1 The shopper retrieves a page with details about a product of interest. Load balancing assigns this request to the top server. The server creates a session ID and sends it along with the rest of the response to the client. The Cisco ACE Module detects the session ID and creates an entry that associates the session ID with the top server in the sticky database.

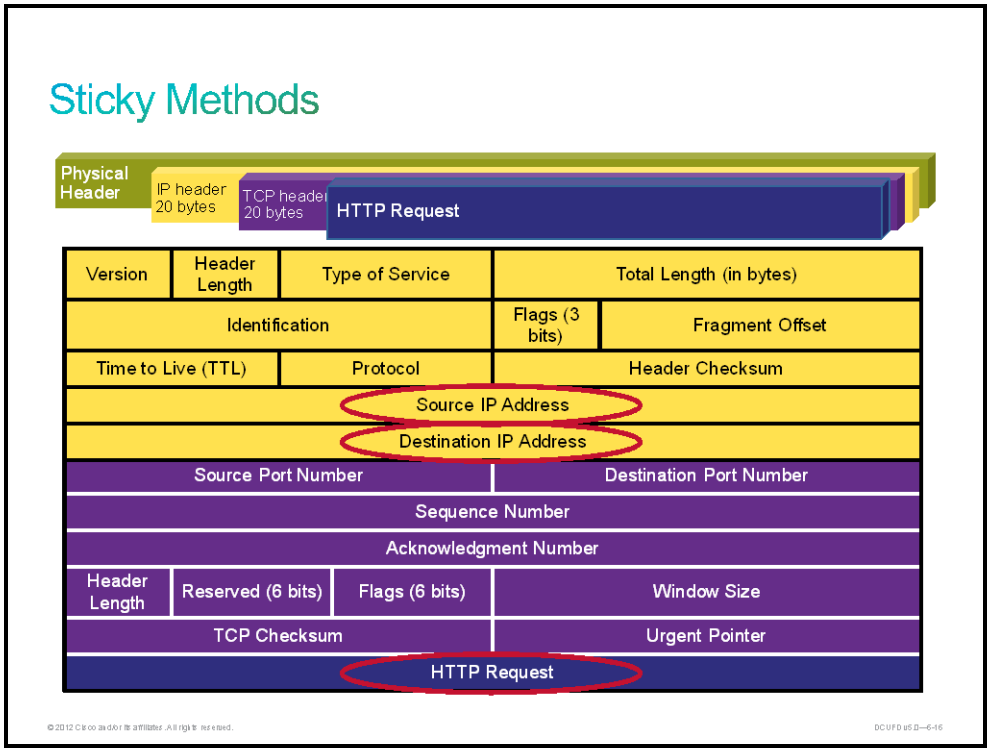
Step 2 The shopper presses the “Buy Now” button. The resulting request contains the session ID. The Cisco ACE Module finds the session ID in the sticky database and the request is assigned to the top server. A record is created in the shopping cart database, associating the item that was selected to the session ID. A page is built and returned to the client with confirmation of the buying decision and a checkout link.

Step 3 The shopper presses the checkout link. Again, the Cisco ACE Module finds the session ID in the sticky database and the request is assigned to the top server. This server uses the session ID in the client request to retrieve the list of items in the shopping cart and continues with the transaction.

Note Session persistence lasts longer than a single TCP connection.

https://t.me/learningnets

certcollection.net



Three different methods of stickiness can be configured with the Cisco ACE Module:

- IP address stickiness tracks the source IP address, the destination IP address, or both IP addresses in the request packets.
- HTTP header stickiness tracks the value of an HTTP header field in the HTTP request.
- Cookie stickiness tracks the values of cookies in the HTTP request and response.

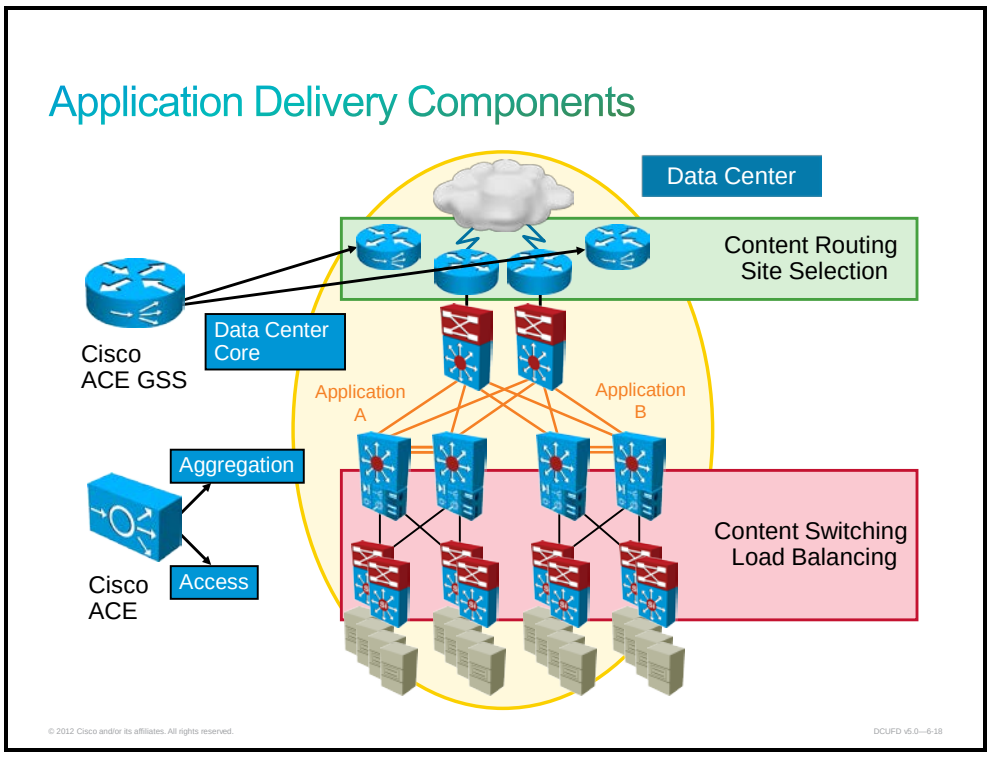
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

Application Delivery Services

This topic explains how to add application services to an existing data center.



This figure illustrates the application delivery components in the data center network. At the content routing layer, site selection is provided by Cisco ACE Global Site Selector (GSS). At the content switching layer, load balancing is provided by the Cisco ACE Module or appliance. The Cisco ACE appliance is deployed in the data center access layer as well.

https://t.me/learningnets

certcollection.net

Content Routing

- Provides global redundancy among the redundant data center sites
- Used for site selection in GSLB
- Deployed in data center core layer
- Different site selection technologies:
  - HTTP redirect
  - Route health injection (RHI)
  - DNS-based solution
- DNS-based solution = GSS



Cisco
ACE GSS

© 2012 Cisco and/or its affiliates. All rights reserved.

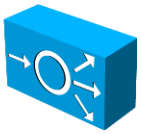
This figure introduces the content routing role and functions in the data center. Content routing provides global redundancy among the redundant data center sites and is used for site selection in global server load balancing (GSLB).

https://t.me/learningnets

certcollection.net

Cisco ACE Module Content Switching

- Also referred to as SLB for the group of servers or cache (Cisco WAE) farms
- TCP connection must be served by the same server unless TCP is split across members of the server cluster
- Session persistence ensures that many TCP and UDP connections are served by the same server
- Provides local application and server access redundancy within the single data center
- Can be deployed in conjunction with GSS to provide the following:
  - Load-sharing of client requests among servers between data centers
  - Disaster recovery
  - Redundancy between data centers

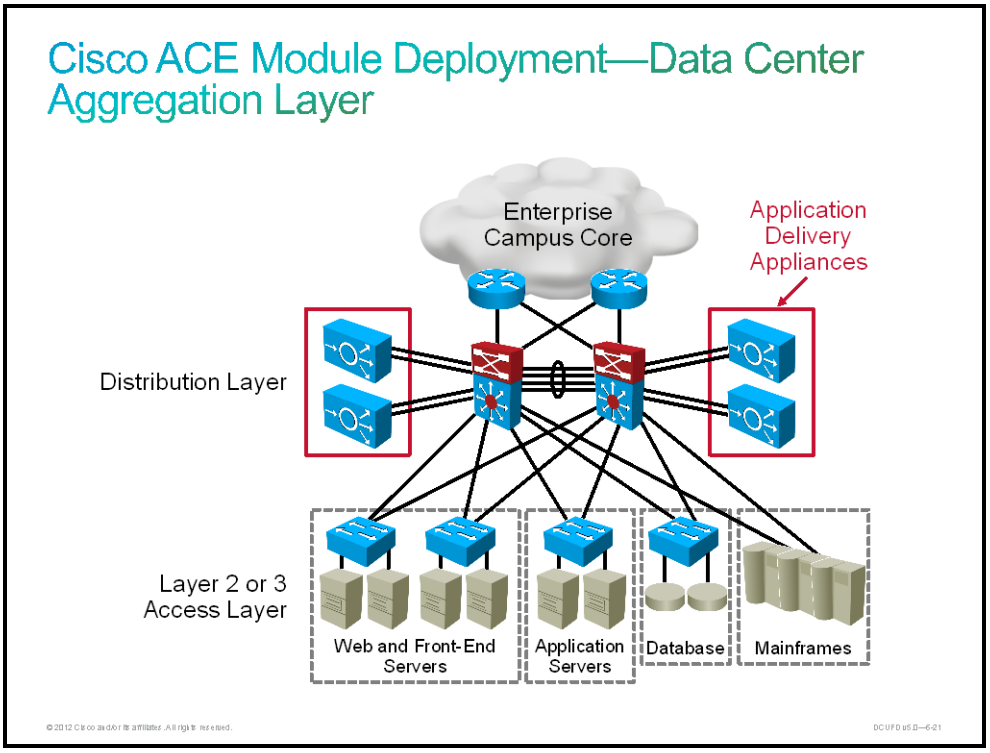


Cisco ACE Module

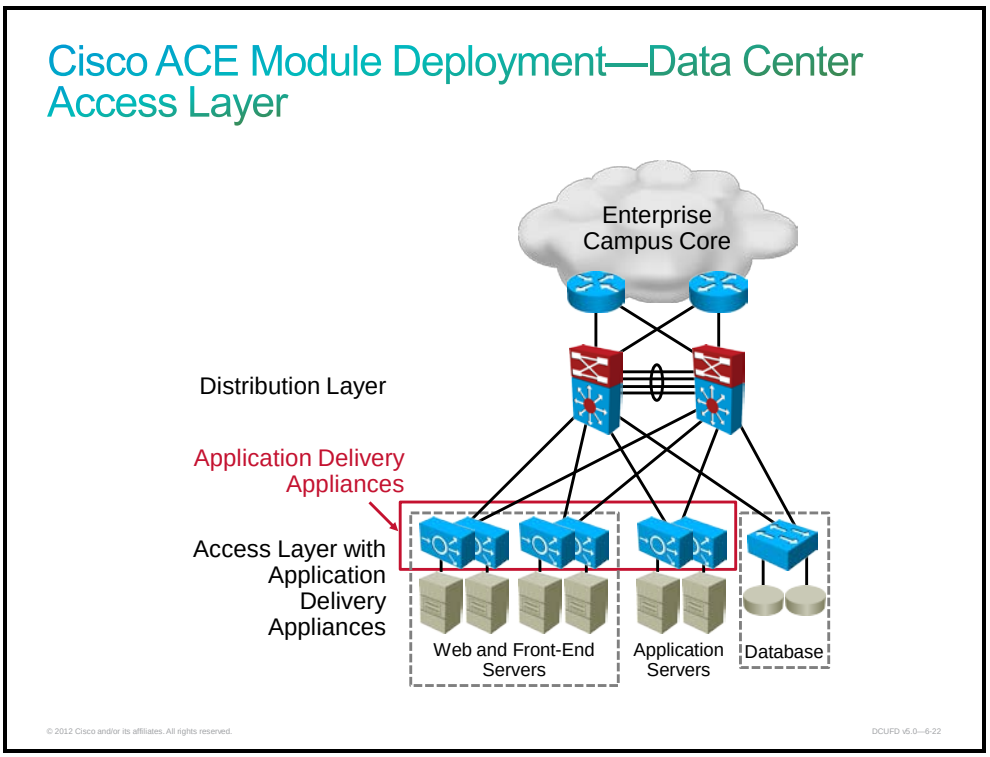
© 2012 Cisco and/or its affiliates. All rights reserved.

This figure introduces the content switching role and functions in the data center. Cisco ACE Module content switching is also referred to as server load balancing (SLB) for the group of servers or cache (Cisco Wide Area Application Engine [WAE]) farms. TCP connections must be served by the same server unless TCP is split across the members of the server cluster. Session persistence ensures that many TCP or UDP connections are served by the same server. SLB provides local application and server access redundancy within a single data center.

https://t.me/learningnets



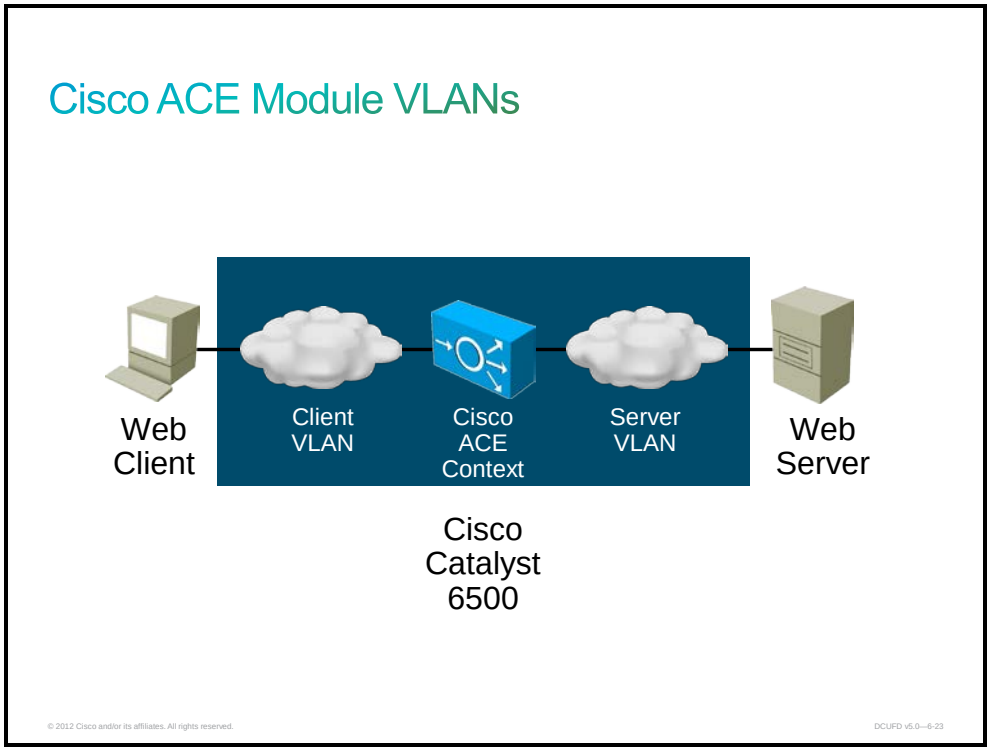
Cisco ACE Module SLB can be deployed in the data center distribution layer using Cisco Catalyst 6500 Series Switches and ACE Module services or by using an external Cisco ACE appliance. The Cisco ACE appliance can be connected to the Cisco Nexus 7000 Series Switches.



Cisco ACE Module SLB can be deployed in the data center access layer using Cisco Catalyst 6500 Series Switches with integrated Cisco ACE Module services or by using an external Cisco ACE appliance. The Cisco ACE appliance can be connected to the Cisco Nexus 7000 Series Switches or to the Cisco Nexus 5000 Series Switches.

https://t.me/learningnets

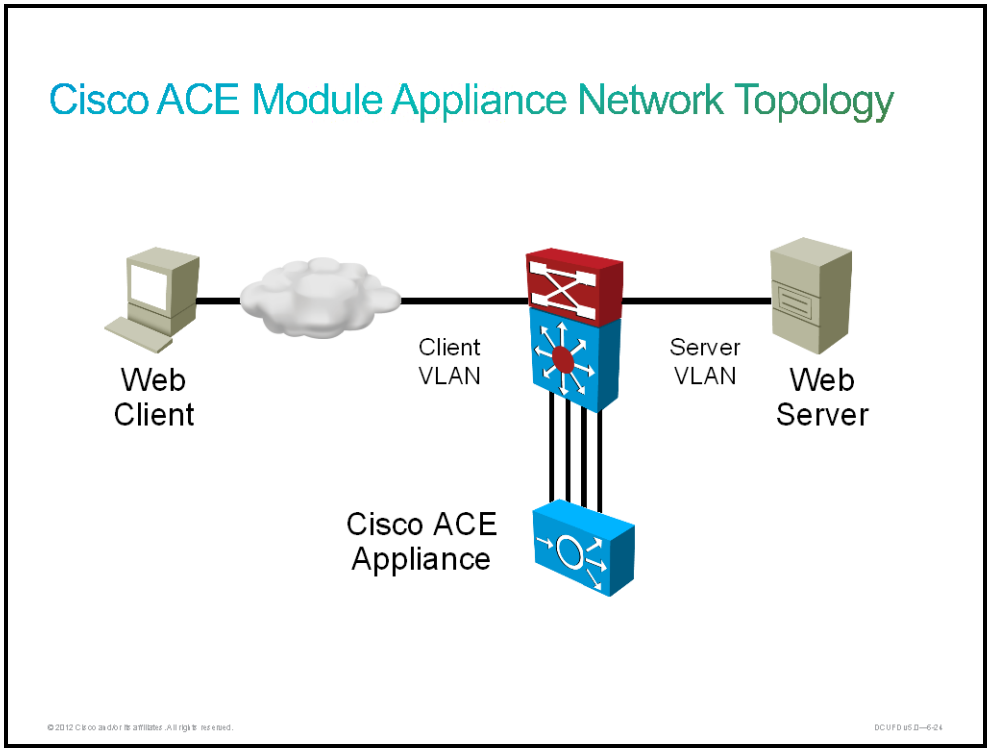
certcollection.net



This figure shows how a Cisco Catalyst 6500 Series Switch that contains a Cisco ACE Module connects to a network. In this example, the Cisco ACE Module connects to the network using two VLANs—one for a connection to clients (a web client in this figure) and the other for a connection to servers (a web server in this figure). Diagramming the network within the chassis is often necessary to completely understand and document a network topology with Cisco Catalyst 6500 service modules. As a result, the Cisco ACE Module is shown diagrammed as a standalone component of the network.

https://t.me/learningnets

certcollection.net

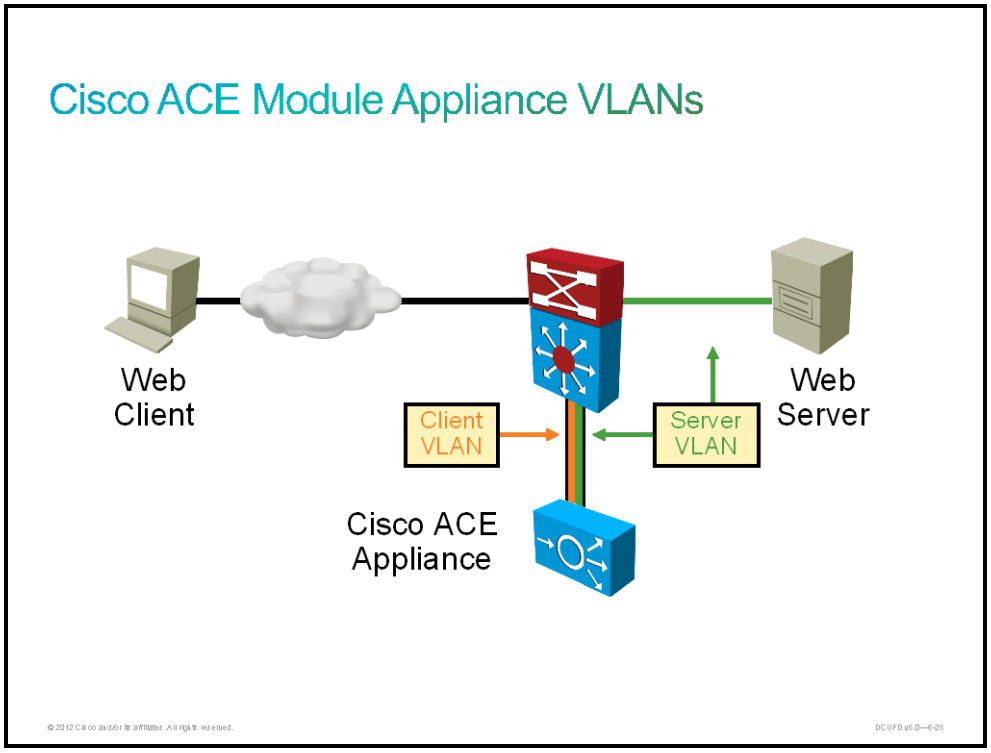


This figure shows a basic network where a Cisco ACE Module appliance is physically connected to a router using Gigabit Ethernet, port channels, and VLAN trunking to communicate with the servers in the network.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net



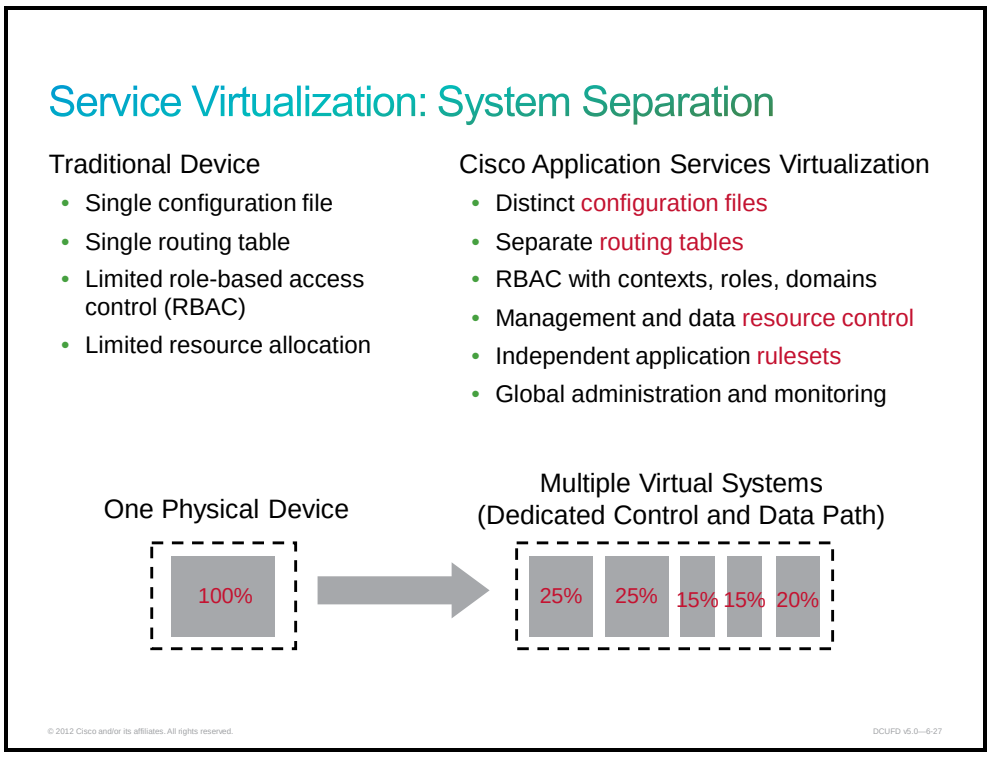
In this example, the Cisco ACE Module connects to the network over all four Gigabit Ethernet links that are logically bonded together using a port channel link. Two VLANs are used—one for a connection to clients (a web client in this figure) and the other for a connection to servers (a web server in this figure). Diagramming the individual VLAN connections is often necessary to completely understand and document a network topology. As a result, the Cisco ACE Module appliance is shown diagrammed as a standalone component of the network.

https://t.me/learningnets

certcollection.net

Cisco ACE Virtualization

This topic describes contexts.

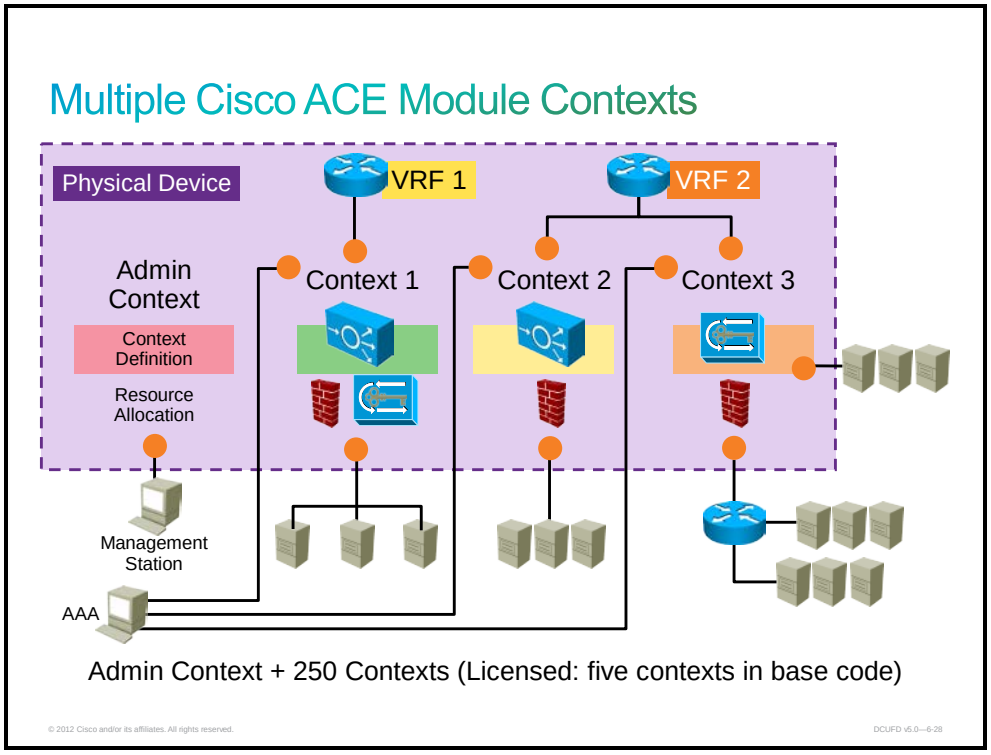


The Cisco ACE Module supports the creation of virtual Cisco ACE Module images, called “contexts.” Each context has its own configuration file and operational data, providing complete isolation from other contexts on both the control and data levels. Hardware resources are shared among the contexts on a percentage basis.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net



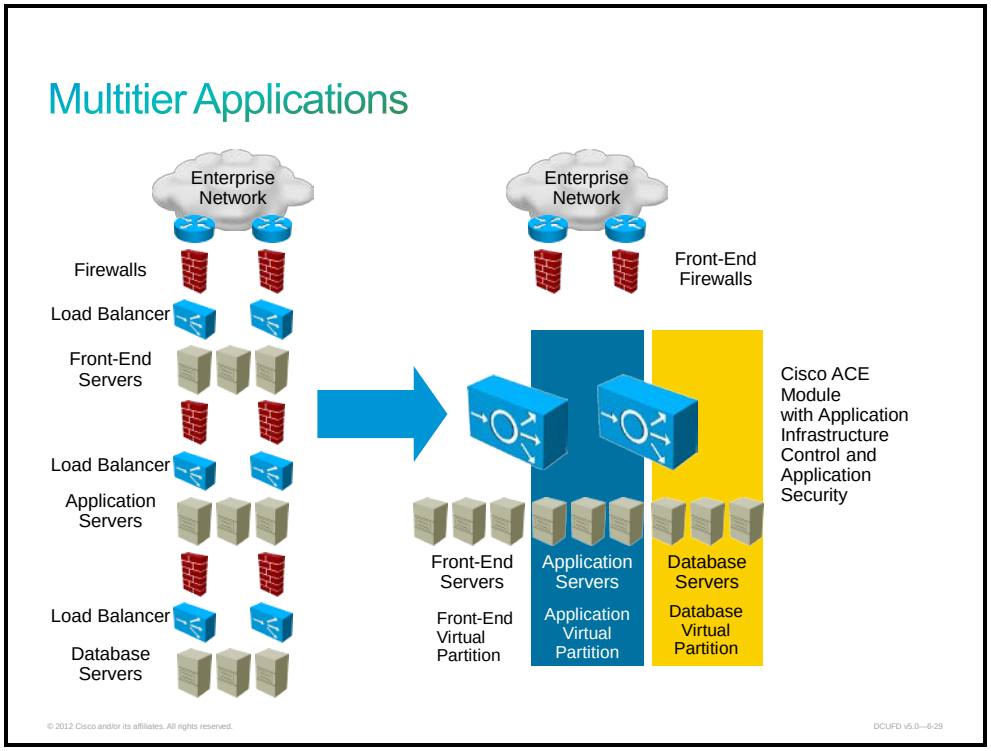
The Cisco ACE appliance supports virtualization through the extension of the logic to the application delivery space of the Layer 2 and Layer 3 VLANs and virtual routing and forwarding (VRF) instances that the Cisco Catalyst 6500 Series natively supports. It is simple to map Cisco ACE Module virtual devices to VLANs and VRF instances, thereby associating a separate network instance on the Cisco Catalyst 6500 Supervisor Engine with a completely independent application delivery instance.

Each virtual device can be dedicated to a set of applications, to an organization within the enterprise, or to a customer in a hosted environment. Overlapping IP addresses are supported and each virtual device benefits from independent network management and policies, as well as from a dedicated virtual routing instance with full Cisco IOS routing protocol support.

Network resources can be dedicated to a single context or shared between contexts. By default, a context named “Admin” is created by the Cisco ACE Module. This context cannot be removed or renamed. Additional contexts and the resources to be allocated to each context are defined in the configuration of the Admin context. The number of contexts that can be configured is controlled by licensing on the Cisco ACE Module. The base code allows five contexts to be configured. Licenses are available that expand the virtualization to 250 contexts. The Admin context does not count in the license limit for the number of contexts.

https://t.me/learningnets

certcollection.net



One use of Cisco ACE Module contexts is to provide application controls at multiple levels of a multitier application architecture. On the left side of this figure is a typical multitier architecture with front-end web servers, application or middleware servers, and back-end database servers. Typically, load-balancing and firewall services are required between layers. Each layer can be implemented using a Cisco ACE Module context, which maintains separate data flows and security controls while minimizing the number of devices to be managed.

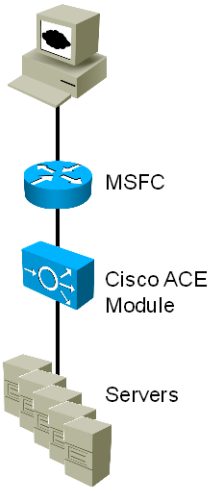
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

Designing Cisco ACE Module Contexts

- Always use at least one context in addition to Admin.
- Identify network segments where multiple flows are in transit.
- Allocate contexts to topology points with common functional, management, resource, and redundancy requirements.
- Split contexts for ease of configuration.
- Design topology changes.



© 2012 Cisco and/or its affiliates. All rights reserved.

The process of designing a Cisco ACE Module solution includes determining the number of contexts to use. After the number of contexts has been determined, topological changes to the network can be designed. There are some guidelines to consider in determining the number of Cisco ACE Module contexts:

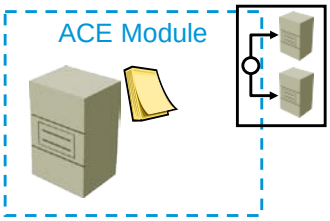
- Always use at least one non-Admin context for functional configuration. This allows a second functional context to be added as required, without the need to move the production configuration from Admin to another context.
- Identify the network segments where multiple flows to be processed are in transit.
- Contexts can be effectively allocated to points in the network topology where the flows in transit have common processing and management requirements.
- Contexts can be split as a mechanism to segment the size of a configuration file if the network topology allows it.

https://t.me/learningnets

certcollection.net

Sticky Database Resources

- Entries must be allocated with resource class.
- Entries cannot be oversubscribed.
- Static entries are configurable.
- One free nonstatic entry is needed for dynamic sticky.
- Oldest entries are replaced, if needed.
- HTTP content sticky.



© 2012 Cisco and/or its affiliates. All rights reserved.

Session persistence is implemented by tracking load-balancing decisions in a sticky database. The memory resources for entries in this database are allocated via the resource management mechanism. By default, no sticky database entries are available to a context; therefore, they must be allocated by putting the context in a resource class. Database entries cannot be oversubscribed.

Static and dynamic sticky entries are stored in the database. At least one entry that is not used for a static entry must be available in order for dynamic sticky to work. If a new dynamic sticky database entry needs to be created, the oldest database entry is replaced.

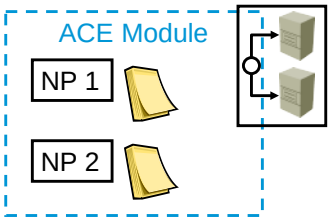
Note	If entries are removed from a context through changes in the resource management definitions, the oldest sticky database entries are removed. This can take some time.

https://t.me/learningnets

certcollection.net

Multinetwork Processor in Sticky Processing

- Sticky database pool split between NPs; CDE does not rehash because one NP hits maximum entries.
- Communication between NPs used to share sticky information.
- Two million sticky entries per NP.
- Four million sticky entries per module.



© 2012 Cisco and/or its affiliates. All rights reserved.

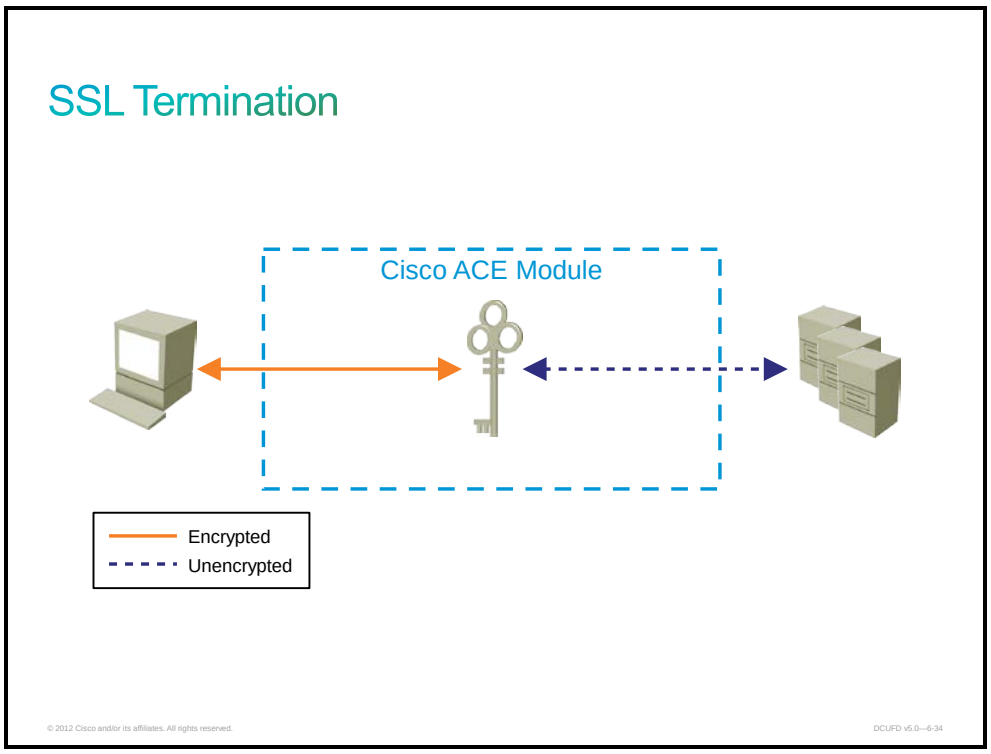
The Cisco ACE Module supports four million sticky database entries, with two million available to each network processor (NP). Sticky processing code that is running on one NP checks for a relevant sticky entry on the other NP without replicating the information in both NPs. As with other resource and processing constraints that are per NP, the Classification Distribution Engine (CDE) does not rehash a connection because of sticky database usage levels on the NPs.

https://t.me/learningnets

certcollection.net

Secure Load-Balancing Design

This topic describes how to design secure application load-balancing solutions.

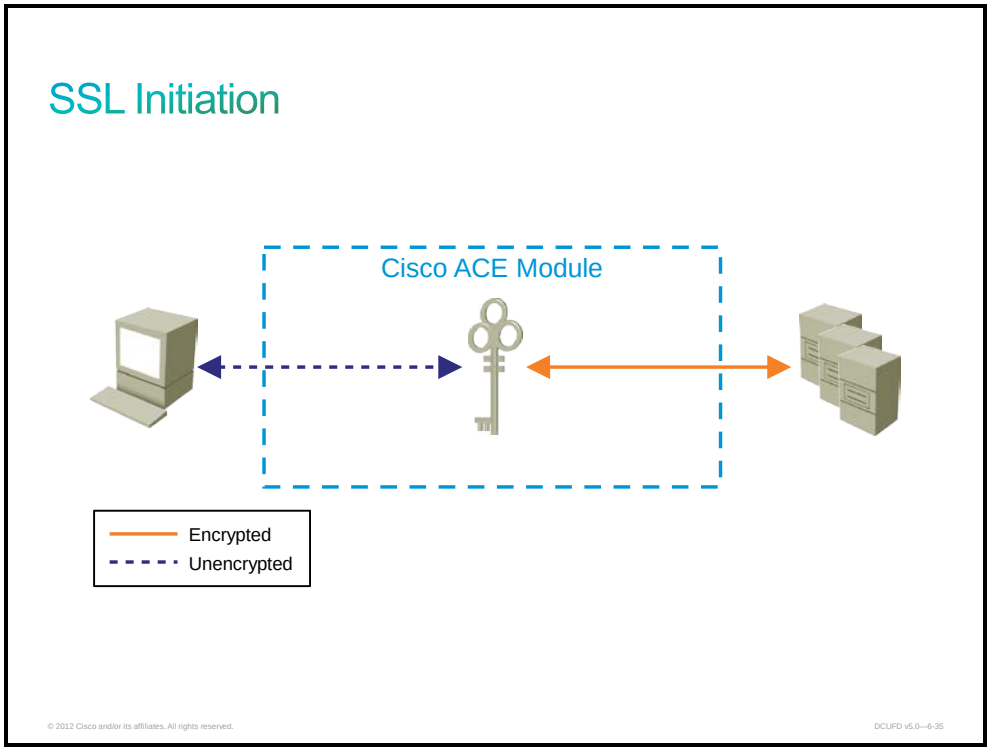


SSL termination is the Cisco ACE Module terminology for deploying the Cisco ACE Module as an SSL offload device. When configured for SSL termination, the Cisco ACE Module terminates the SSL connection from the client, decrypts the request from the client, and sends it as plaintext to the real servers. Notice that the real servers are selected through the normal load-balancing functions of the Cisco ACE Module. Responses from the real server are received by the Cisco ACE Module in plaintext, encrypted, and sent back over the SSL connection to the client.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

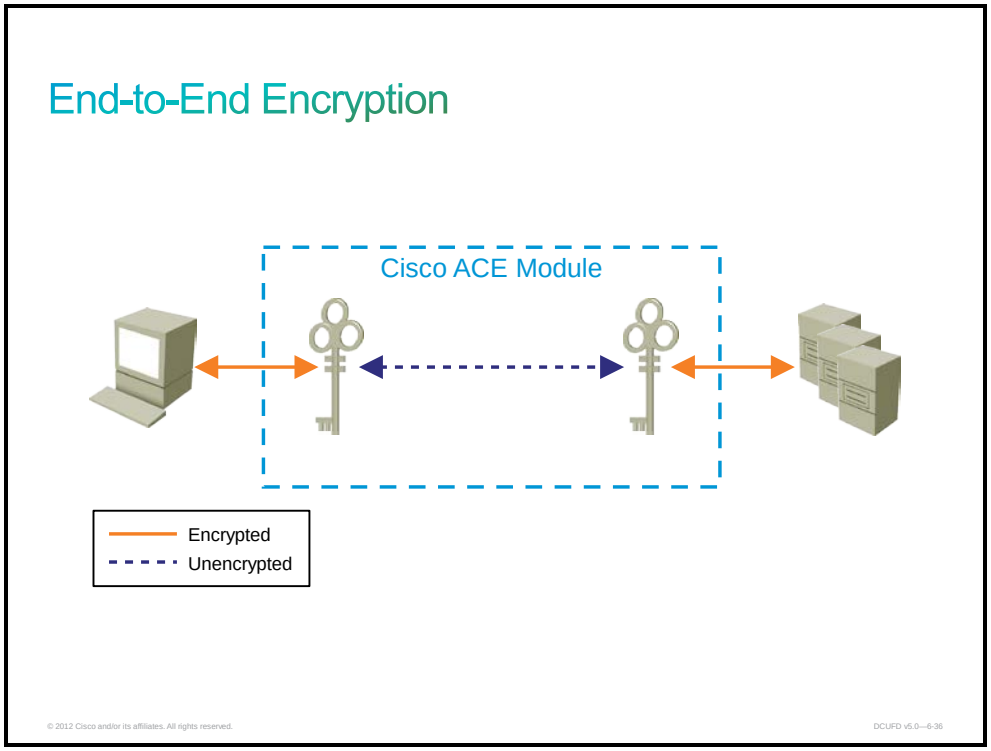
certcollection.net



SSL initiation is used to implement a network design that is often called “back-end SSL,” in which the interaction between the client and the Cisco ACE Module is in plaintext, while the traffic between the Cisco ACE Module and the real servers is encrypted SSL traffic. In SSL initiation, the Cisco ACE Module takes the role of the SSL client when dealing with the real servers.

https://t.me/learningnets

certcollection.net



End-to-end encryption combines SSL termination and SSL initiation in one Cisco ACE Module configuration. This deployment model is often used when highly sensitive data needs to be load-balanced based on Layer 7 criteria but the data is not allowed to exist on any network segment as plaintext. In this situation, the data is only unencrypted within the Cisco ACE Module.

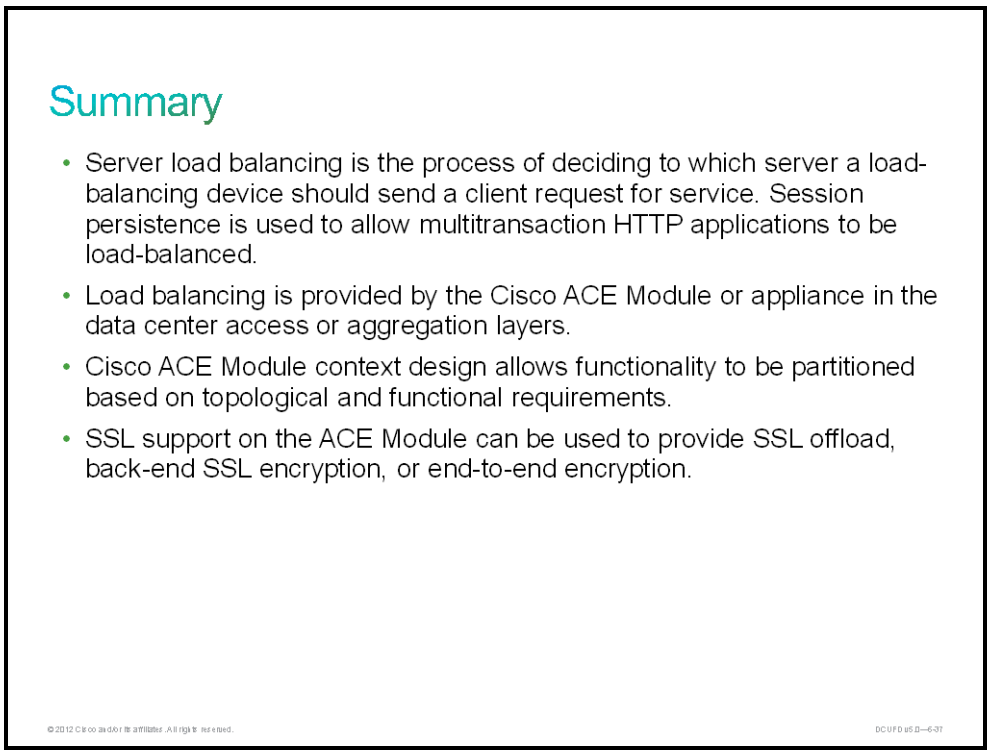
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

Summary

This topic summarizes the primary points that were discussed in this lesson.



https://t.me/learningnets

certcollection.net

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

Lesson 3

Designing Global Load Balancing

Overview

This lesson provides an overview of global server load balancing (GSLB) design. GSLB complements server load balancing (SLB) mechanisms that are used in data centers. GSLB allows you to control how you provide your applications globally and for which data center you choose to fulfill application requests.

Objectives

Upon completing this lesson, you will be able to design GSLB solutions. This ability includes being able to meet these objectives:

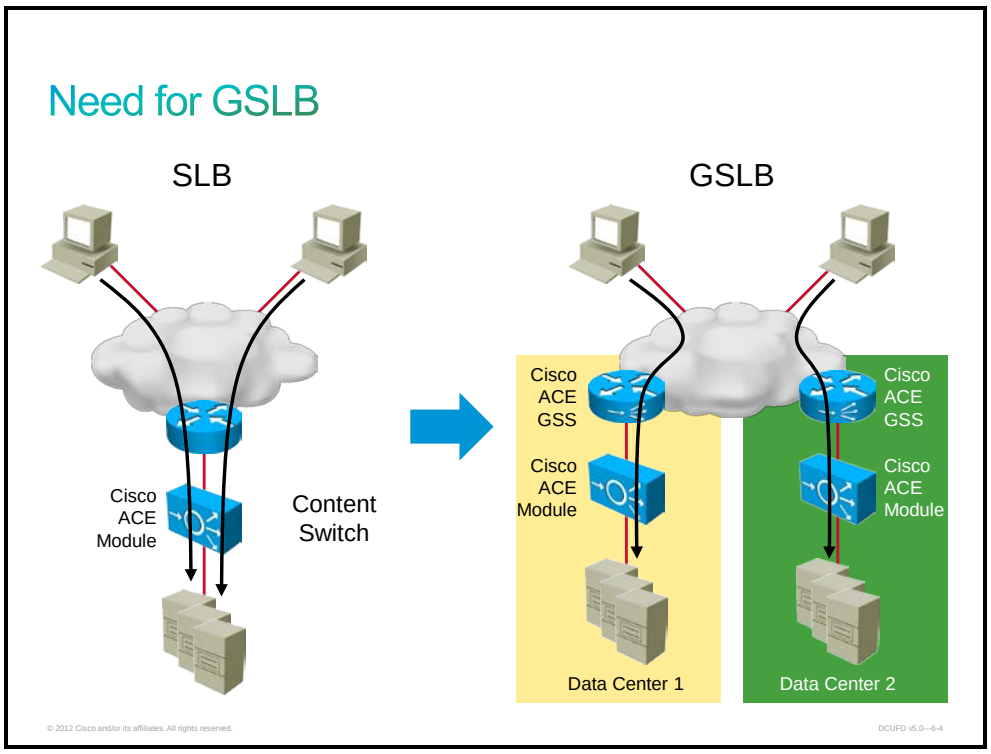
- Explain the need for GSLB
- Design a GSLB solution
- Explain protocols that are used for site selection and site monitoring
- Explain the site selection process

https://t.me/learningnets

certcollection.net

Need for GSLB

This topic explains the need for GSLB.



This figure illustrates the differences between SLB and GSLB.

Server Load Balancing

SLB provides you with the means to load balance between servers, serving the same content at the same site.

SLB is provided by a server load balancer (or a redundant set of them), Cisco Application Control Engine (ACE) Module, or Cisco ACE appliance.

Global Server Load Balancing

GSLB provides you with the means to load-balance client requests between two data centers that serve the same content, and then to load-balance within that site—between servers—serving the same content.

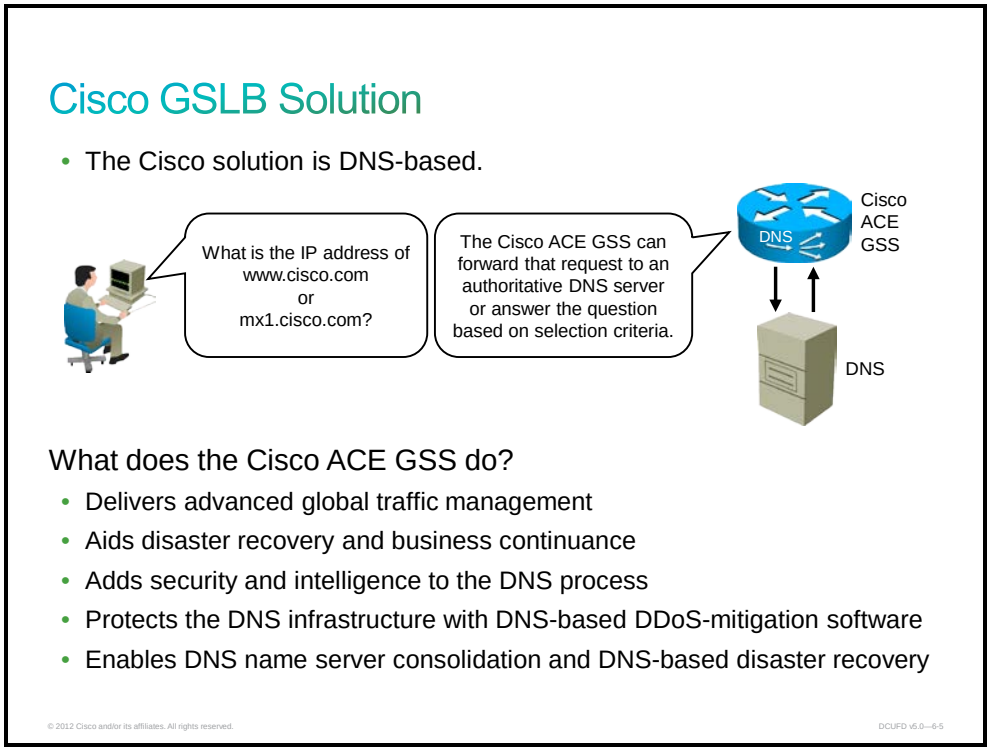
By definition, GSLB balances user requests to available server load balancers (virtual IPs) or hosts that are hosted at different locations.

Note Typically, the locations are geographically dispersed and GSLB is used either in a disaster recovery design or in a site load-balancing design where users are directed to different locations based on a specific request load-balancing algorithm or a proximity discovery method.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net



Cisco ACE Module GSS Features

The Cisco ACE Module Global Site Selector (GSS) delivers advanced global traffic management by working with the Domain Name System (DNS) servers and providing different DNS replies to queries from multiple sites.

GSLB supports geographically dispersed server load balancers and caches:

- Capable of load-balancing any device that uses DNS to get to a data center:
  - Layer 4 to Layer 7 switches
  - Origin servers
  - Mainframe and webframe
- Connect clients to the best server based on the following:
  - Network topology
  - Server load
  - Availability of content and devices

GSS adds security and intelligence to the DNS process along with DNS consolidation.

GSS protects the DNS infrastructure with DNS-based distributed denial of service (DDoS) mitigation software.

GSS enables DNS name server consolidation:

- Supports complete IP management system (DHCP and TFTP).
- With a Cisco Network Registrar license, GSS can replace any existing DNS name server.

It provides universal DNS-based disaster recovery:

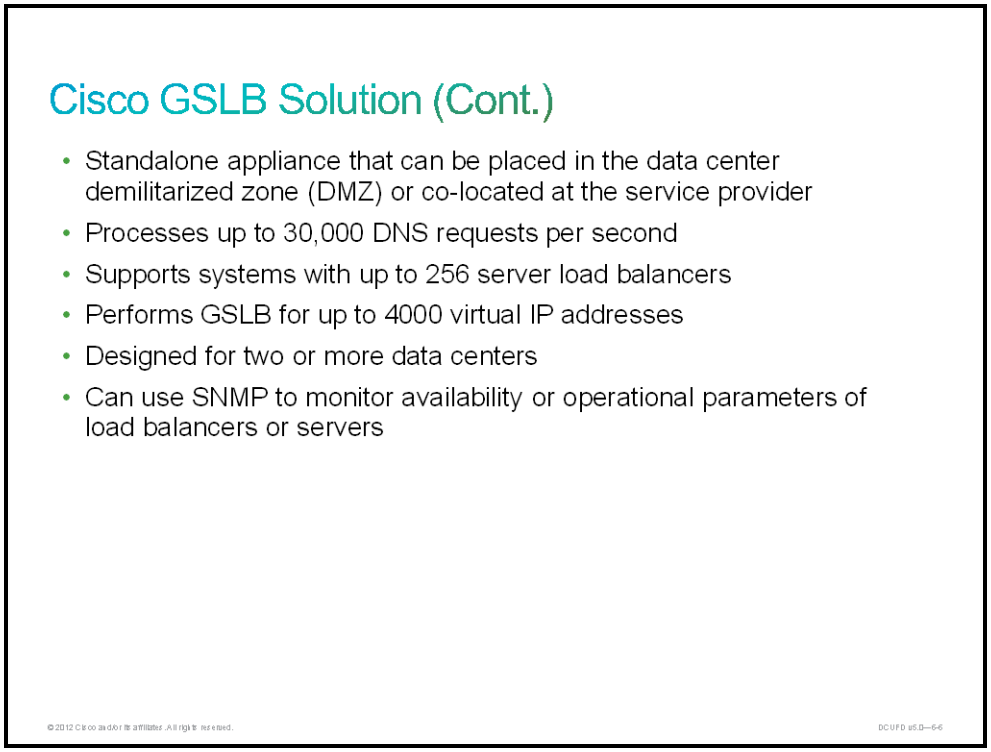
- Redirects clients to the backup data center for any device that supports Simple Network Management Protocol (SNMP) MIB and uses DNS.

https://t.me/learningnets

certcollection.net

Cisco ACE Module GSS Functions

The Cisco ACE Module GSS takes control of the DNS control plane. This enables the ability to globally load-balance all web-based traffic, for example, across multiple data centers in real time. GSS can also verify reachability of sites to ensure that, in a site failure, all traffic is rerouted automatically for continuous site accessibility. In short, GSS is capable of load-balancing any traffic that uses DNS to reach a data center.



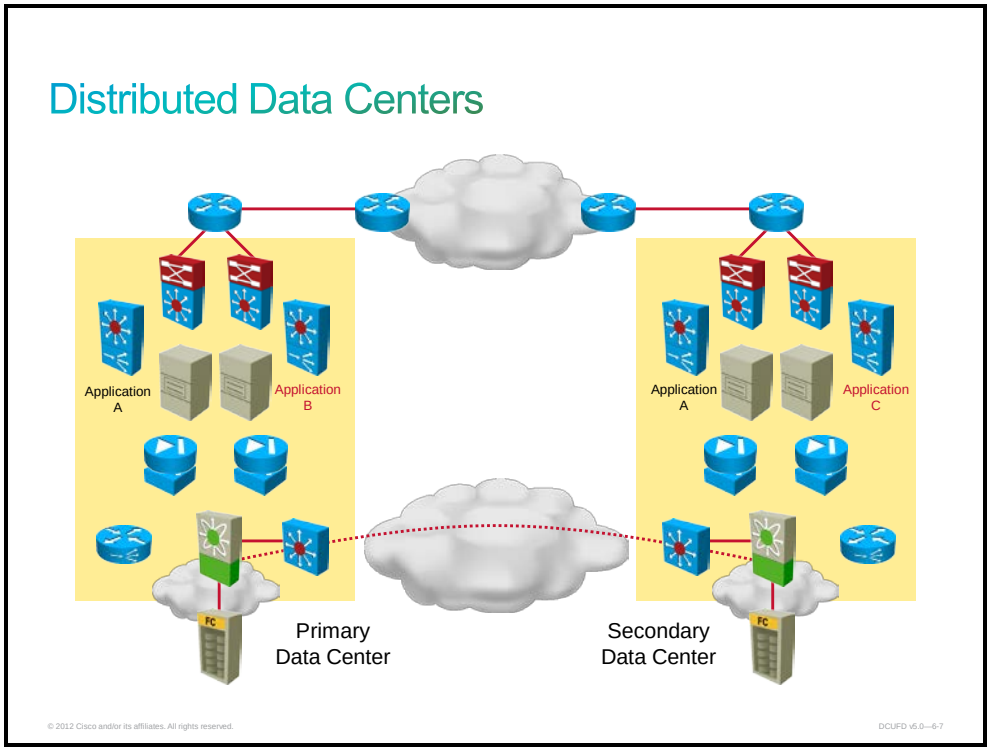
Dedicated GSS uses the GSS to load-balance traffic between multiple data centers, providing the ability to scale and optimize reliability of existing DNS or third-party server load-balancing infrastructures, and thereby providing a robust business-continuance architecture:

- The DNS resolution process is dedicated to a single device.
- Dedicated GSS is used for disaster recovery and multisite data center deployments. It is capable of massive scalability.
- The value of the dedicated solution justifies the management of an additional device in the data center.
- With centralized command and control, the number of sites and server load balancers has a very small impact on complexity.
- This approach provides heterogeneous support for all Cisco server load balancers (Cisco ACE Module, Cisco Content Services Switch, Cisco Content Switching Module, LocalDirector [LD]) and third-party server load balancers. It is the only approach for mixed deployments of Cisco server load balancers.
- It is dedicated to processing DNS requests, and delivers high performance and scaling.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net



Why do companies require distributed data center environments? For any enterprise or service provider environment, there is a requirement to ensure that data is available anywhere and anytime that anyone requires it. If there were only one data center, this would become the single point of failure and, in case of failure, data would no longer be available for customers to access. Therefore, there is a need for multiple data centers to service the guaranteed availability to both internal and external customers.

Other requirements include the need for application scalability and security. Again, how can you provide this if there is only one single data center on which everyone relies? In addition, other considerations are regulatory, along with how to avoid data loss in the event of a disaster.

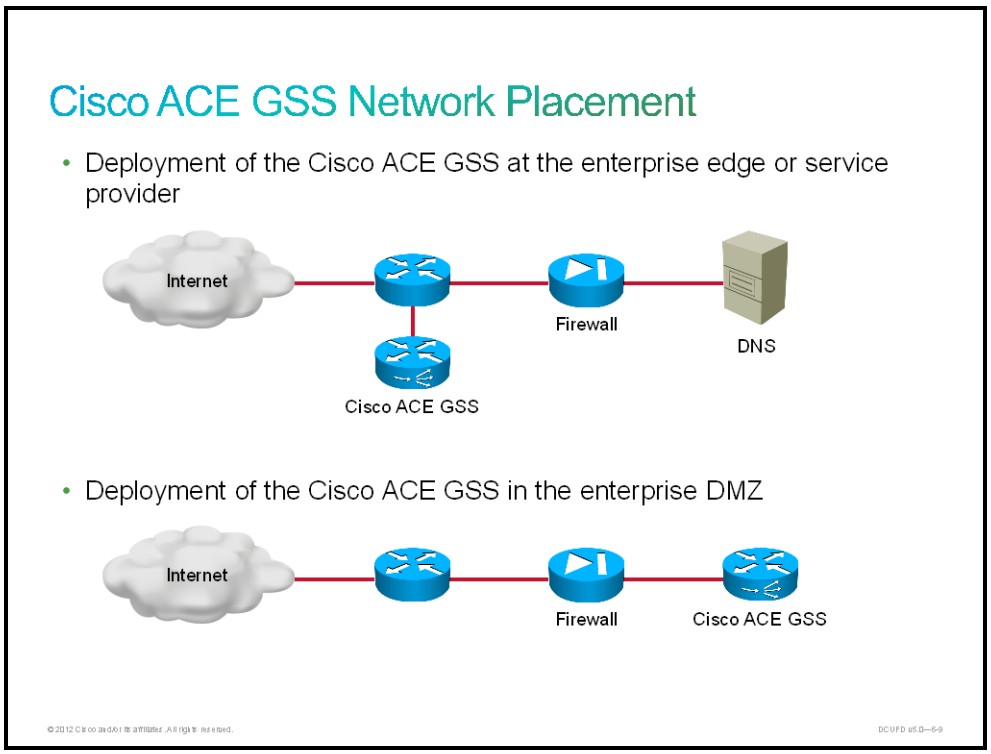
All these issues are paramount in any company, and a company must ensure that there is no single point of failure that would cost money if that data center fails.

https://t.me/learningnets

certcollection.net

GSLB Solution Design

This topic describes how to design a global load-balancing solution.



This figure shows GSS topology options.

GSS serves as an authoritative name server for one or more domains and must be available either publicly or privately in your network, depending on whether GSS will serve clients in a public or a private network. DNS proxy servers must be able to connect to GSS to resolve domain name requests for the client DNS proxy servers.

GSS can be deployed in the same location in which an enterprise would normally deploy its DNS servers. Depending on the configuration, GSS can assume all name server functions, or can assume a subset of the name server functions and forward other requests to other authoritative GSSs and name servers in the enterprise.

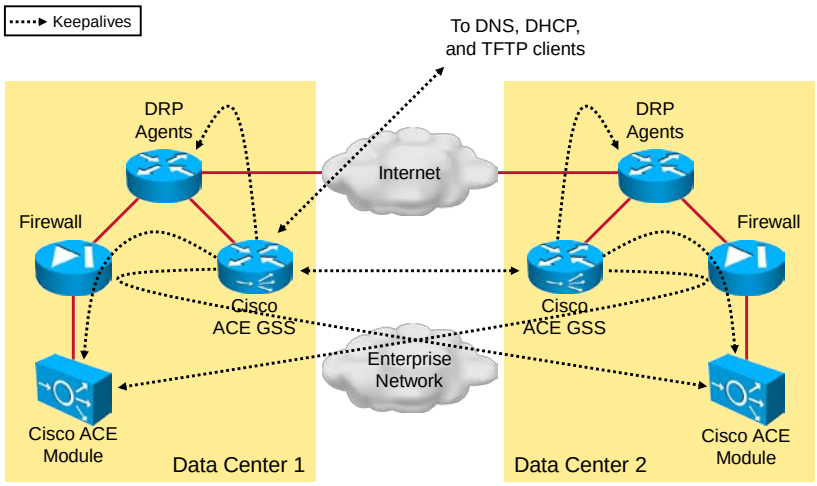
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

Cisco ACE GSS Flows

- Deployment of the Cisco ACE GSS requires open sessions through the enterprise and data center firewalls



© 2012 Cisco and/or its affiliates. All rights reserved.

This figure illustrates the communication to and from a GSS.

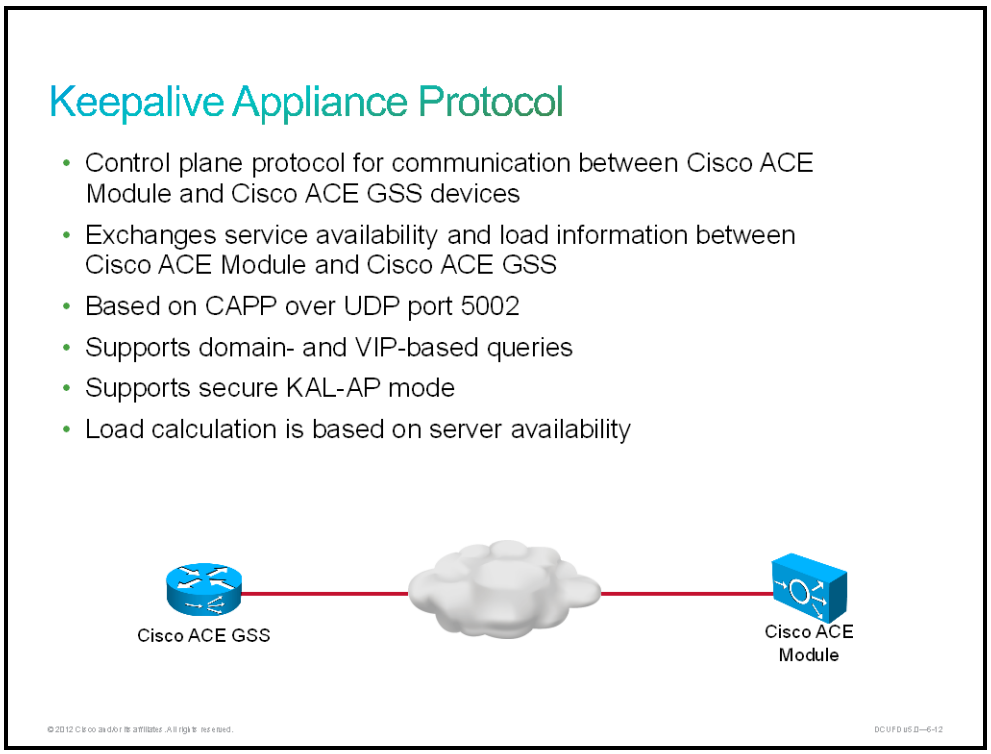
Name servers are often deployed behind firewalls to prevent unauthorized access and DDoS attacks. This is a good practice with GSSs, as well. When deploying a GSS behind a firewall, the firewall must be configured not only to accept DNS queries, but also to allow keepalive protocols (Keepalive-Appliance Protocol [KAL-AP], Internet Control Message Protocol [ICMP], HTTP, and user-defined TCP ports) or Director Response Protocol (DRP) from server load balancers and servers, router agents, and inter-GSS communication. Other possible protocols that might be allowed access to GSS through the firewall are FTP, Telnet, Secure Shell (SSH), and SNMP.

https://t.me/learningnets

certcollection.net

Site Selection Protocols

This topic describes protocols that are used for site selection and site monitoring.



The KAL-AP is the control plane protocol for the Cisco GSLB solution.

KAL-AP support is outlined in this figure.

Through KAL-AP, Cisco ACE Module returns availability information to the global server load balancer in the form of a percentage of the server farm available.

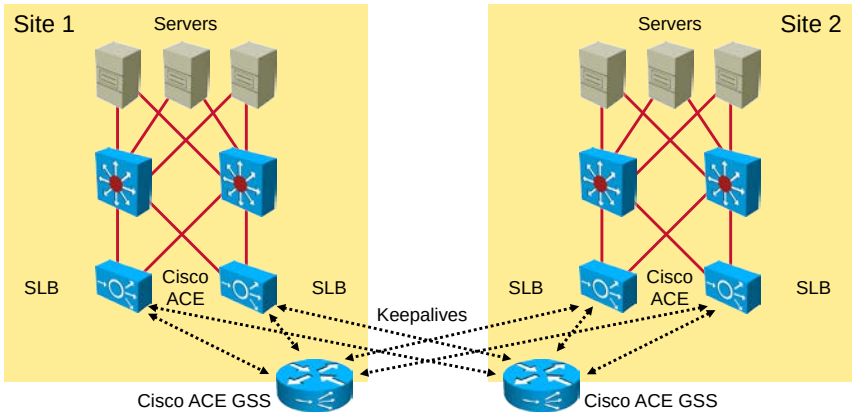
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

KAL-AP Keepalives

- The firewalls need to be configured to permit KAL-AP flows between Cisco ACE Module devices and Cisco ACE GSS.
- Cisco ACE Module establishes the health, load, and availability of the server farm.
- Data from Cisco ACE Module signaled to Cisco ACE GSS.



© 2012 Cisco and/or its affiliates. All rights reserved.

Keepalives on GSS are back-end processes that are used to gather state and load information from devices within the data center, such as local server load balancers and origin servers. This information can then be used by GSS to choose sites based on their current loading so that client requests are not forwarded through to sites that are currently overloaded.

GSS keepalive types include the following:

- Simple (this verifies availability):
  - **Layer 3:** An ICMP ping is used for device online status.
  - **Layer 4:** The TCP three-way handshake is used to identify online status of a device.
  - **Layer 5:** An HTTP head request is sent through to the target device with GSS checking for “200 OK” responses from the web page.
- Advanced (this verifies both availability and load):
  - **KAL-AP:** GSS uses this to check the loading and virtual IP (VIP) online status.
  - **Name server query:** GSS requests a name server record from the DNS server to check availability of the local name server.

https://t.me/learningnets

certcollection.net

KAL-AP Load Calculation

- Server list is built based on query type:
  - The domain query uses an object list of the management domain to find servers.
  - The VIP query finds servers through Layer 3 and Layer 4 class maps that reference VIP and relevant Layer 3 and Layer 4 policy map actions.
  - Backup servers are considered if operational.
- Load value equals 255 – (percentage of operational servers * 255)
- Zero (“0”) is returned if service (server VIP address) is not operational.



© 2012 Cisco and/or its affiliates. All rights reserved.

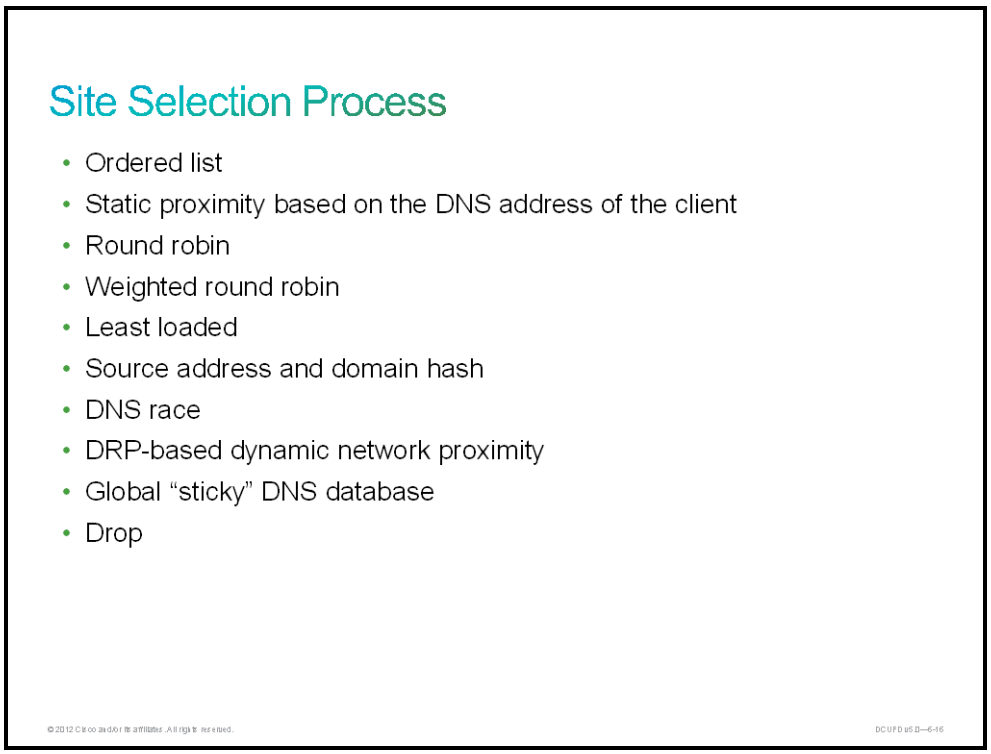
The KAL-AP load value is computed by finding all the relevant servers for a query and determining a percentage of servers that are operational. This percentage is then scaled to a number between 0 and 255 and subtracted from 255. For example, if 6 servers out of 10 are operational, the load value that is returned is 255 – (6/10 * 255), which is 102.

https://t.me/learningnets

certcollection.net

Site Selection Process

This topic describes the site selection process.



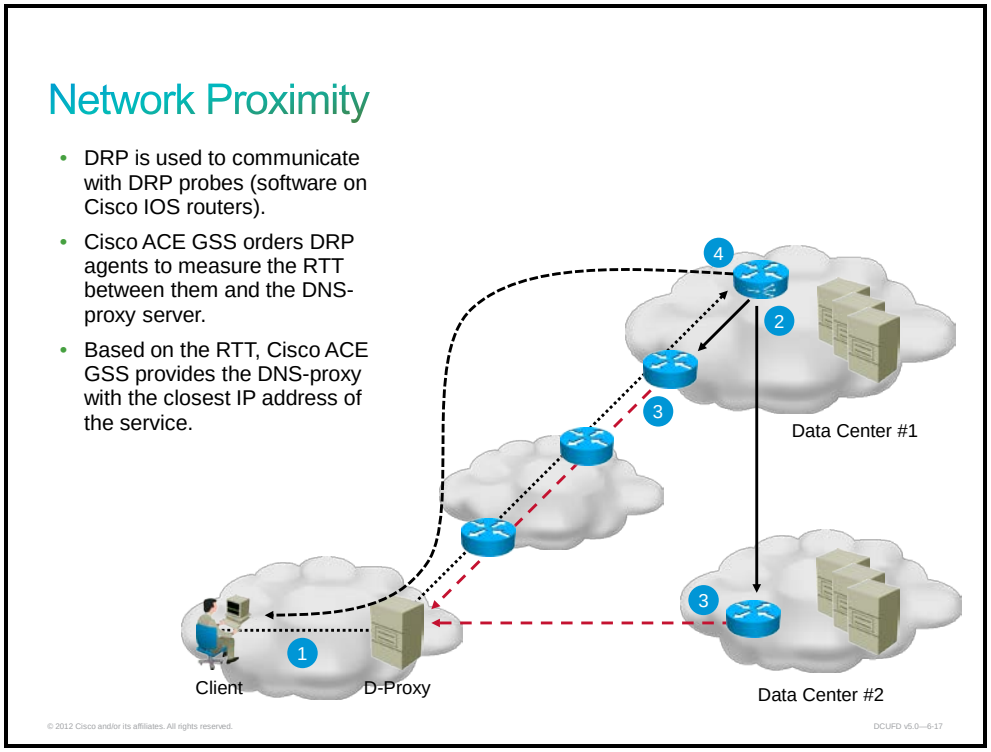
GSS has several ways to load-balance:

- **Ordered list:** A list of locations is configured in order of preference.
- **Static proximity based on the DNS address of the client:** Static proximity provides a location address based on the client DNS address along with an optional mask.
- **Round robin:** Each request cycles through the available answers in order.
- **Weighted round robin (WRR):** A weighting is applied to the various sites, causing them to be chosen based on the weighting value, in a round-robin manner.
- **Least loaded:** Loading information is sent back to GSS using the Content and Application Peering Protocol (CAPP) UDP. With this detail, GSS can load-balance based on the loading at a specific site.
- **Source address and domain hash:** The IP address of the client DNS proxy (D-proxy) and destination domain are used to identify the destination site of the request. This also provides answer “stickiness.”
- **DNS race:** In this instance, GSS initiates a race of A-record responses to the client, thereby finding the closest site to the client D-proxy.
- **DRP-based dynamic network proximity**: GSS localizes client traffic by probing the client DNS name servers and routing the client to the closest data center, based on the lowest round-trip time (RTT) measurement.
- **Global “sticky” DNS database:** GSS dynamically tracks where clients are sent and then ensures that they are sent to the same device for subsequent requests. Entries are based on the IP address of the client name server and the domain name being requested, as well as which sticky answers are being shared between GSSs.

https://t.me/learningnets

certcollection.net

- **Drop:** GSS silently discards the DNS request.



GSS uses DRP to communicate with probing devices, called DRP agents, in any given zone. DRP is a general UDP-based query and response information exchange protocol that was developed by Cisco. Any Cisco router can be used as the probing device in a zone that is capable of supporting the DRP agent software and can measure ICMP, TCP, or path-probe RTT.

GSS transmits DRP queries to one or more probing devices in the GSS network, instructing the DRP agent in the probing device to probe specific DNS-proxy IP addresses. Each probing device responds to the query by using a standard protocol, such as ICMP or TCP, to measure the RTT between the DRP agent in the zone and the IP address of the D-proxy device of the requesting client.

When GSS receives a request from a D-proxy, it decides if it can provide a proximate answer from its proximity database (PDB). If not, GSS sends a probe to one or more probing devices to get proximity information between those probing devices and the new D-proxy. This information is then added to the PDB.

This is the process:

Step 1 The client sends a DNS request via his D-proxy, which is forwarded to GSS for resolution.

Step 2 GSS sends a DRP message to the routers in the data centers.

Step 3 The routers send a TCP or ICMP message to the D-proxy to ascertain RTT information so that GSS can choose the closest site to the client.

Step 4 Based on the information that is received from the routers, GSS selects an A-record of the site that is closest to the user for content requests.

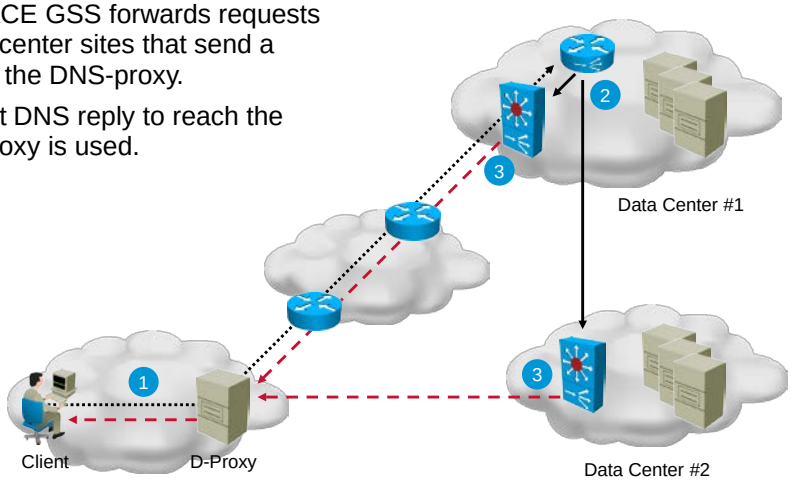
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

DNS Race

- Client DNS-proxy requests name resolution to the Cisco ACE GSS.
- Cisco ACE GSS forwards requests to data center sites that send a reply to the DNS-proxy.
- The first DNS reply to reach the DNS-proxy is used.



© 2012 Cisco and/or its affiliates. All rights reserved.

Proximity without probing is, in effect, a DNS race. GSS sets up a race between content routing agent (CRAs) on Cisco Content Services Switches. The Cisco Content Services Switches respond to the D-proxy and whichever agent has the lowest response time wins the race and is considered the best location for the content request of the client.

GSS measures the latency between data centers to prepare for DNS race conditions. The latency from GSS to each data center is used to send requests for DNS resolution so that they arrive in each data center simultaneously.

This is the process:

Step 1 The client sends a DNS request via his D-proxy, which is forwarded to GSS for resolution.

Step 2 GSS forwards this request through to the Cisco Content Services Switches and initiates a race.

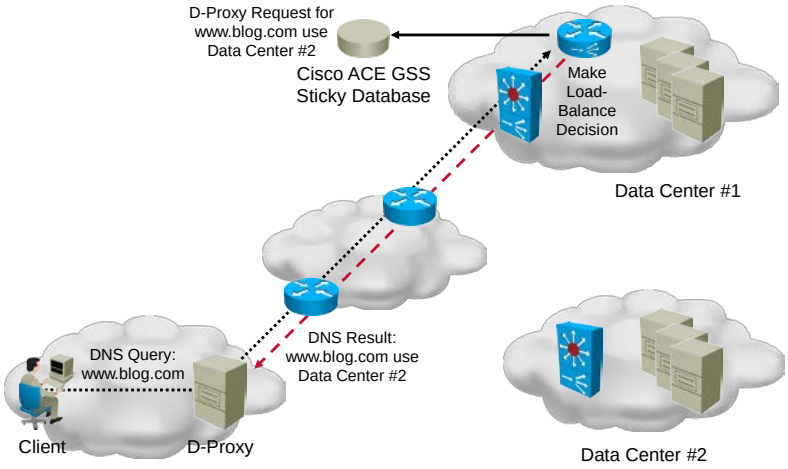
Step 3 The Cisco Content Services Switches respond to the DNS proxy of the client and whichever response is received first is deemed to be the closest site for content requests.

https://t.me/learningnets

certcollection.net

Sticky DNS Database

Stickiness ensures that a DNS-proxy server always gets served by the same data center.



© 2012 Cisco and/or its affiliates. All rights reserved.

Stickiness enables a GSS to remember the DNS response that was returned for a client D-proxy and to later return that same answer when the client D-proxy makes the same request. When stickiness is enabled in a DNS rule, GSS makes a best effort to always provide identical A-record responses to the requesting client D-proxy, assuming that the original VIP address continues to be available.

When users browse a site, any redirection to a new site is transparent. However, if the user is performing e-commerce-type transactions, a break in the connection might occur when that redirection occurs, which results in a loss of the e-commerce transaction. With DNS sticky enabled on GSS, the e-commerce clients can remain connected to a particular server for the duration of the transaction, even when the client browser refreshes the DNS mapping.

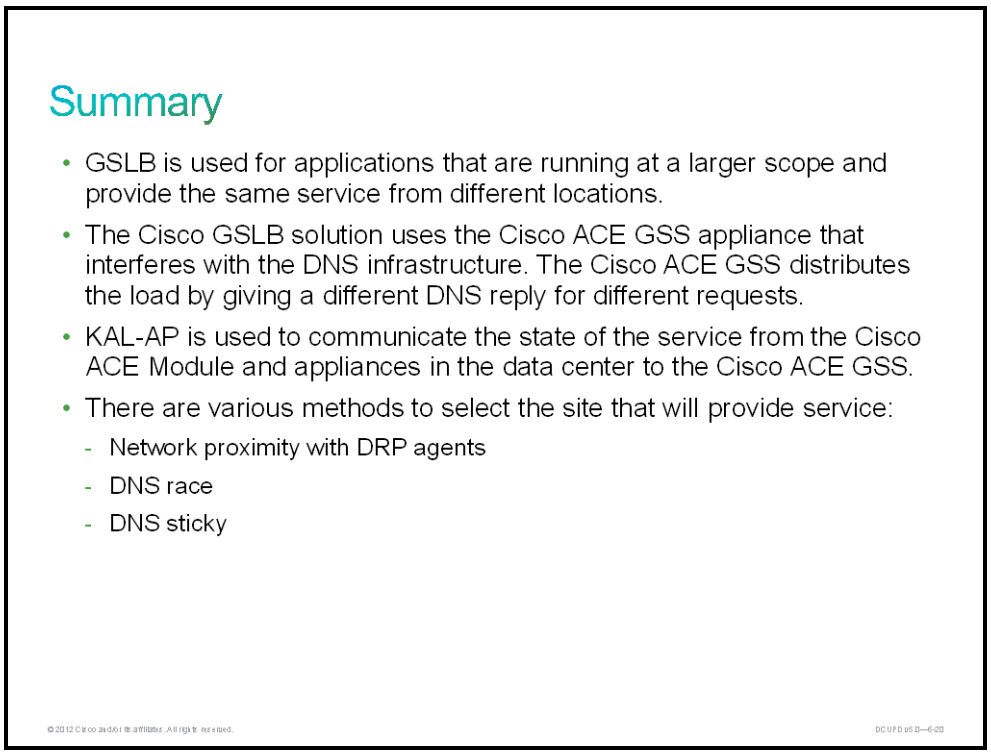
Some browsers impose a connection limit of 30 minutes before requiring a DNS re-resolution. This timeframe might be too short for the client to be able to complete the e-commerce transaction. DNS sticky helps to ensure that the client completes the transaction on the same server even if a DNS re-resolution occurs.

https://t.me/learningnets

certcollection.net

Summary

This topic summarizes the primary points that were discussed in this lesson.



https://t.me/learningnets

certcollection.net

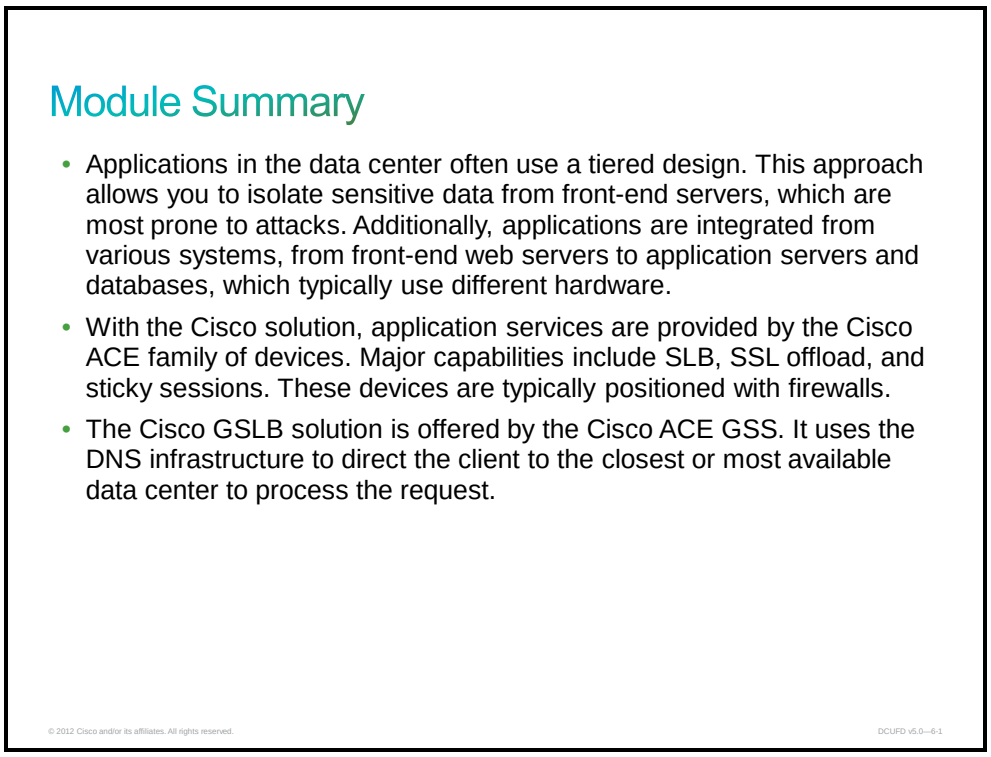
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

Module Summary

This topic summarizes the primary points that were discussed in this module.



Application services are one of many services that are performed in the data center. Application services are mainly provided by the Cisco Application Control Engine (ACE) Module and Cisco Wide Area Application Services (WAAS) families of products. Two important features of the Cisco ACE Module solution are server load balancing (SLB) and Secure Sockets Layer (SSL) services offloading.

https://t.me/learningnets

certcollection.net

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

Module Self-Check

Use these questions to review what you learned in this module. The correct answers and solutions are found in the Module Self-Check Answer Key.

Q1) A thick client-based computer depends heavily on another computer to fulfill its traditional computational roles. (Source: Designing Data Center Application Architecture)

A) true
B) false

Q2) In which application design is scalability increased? (Source: Designing Data Center Application Architecture)

A) single-tier model
B) dual-tier model
C) triple-tier model
D) thick-client application model

Q3) Match the application communication behavior to its correct description. (Source: Designing Data Center Application Architecture)

A) intra-tier communications
B) inter-tier communications
C) multi-site communications

_____ 1. Vertical communication between the front end to the application, the application to the back end, the client to the application, or the client to the back end

_____ 2. Horizontal communication between server farms, clusters, grids, and blades

_____ 3. Communication between distributed data centers

Q4) In which two ways can Cisco WAAS be deployed? (Choose two.) (Source: Designing Data Center Application Architecture)

A) on the Cisco WAE devices as standalone appliances
B) as Cisco WAE network modules that integrate physically into the ISR
C) on the Cisco WAAS devices as standalone appliances
D) as Cisco WAAS network modules that integrate physically into the ISR

Q5) What are the three valid Cisco ACE Module modes of operation? (Choose three.) (Source: Designing Application Services)

A) bridged mode
B) switched mode
C) routed mode
D) one-arm mode
E) bound mode

https://t.me/learningnets

certcollection.net

Q6) Which problem is solved by session persistence? (Source: Designing Application Services)

A) shopping cart problem
B) DoS attack
C) server CPU overload
D) server memory overload

Q7) Which option shows the data center layer or layers at which load balancing can be provided by the Cisco ACE Module or appliance? (Source: Designing Application Services)

A) access layer
B) access or core layers
C) access or aggregation layers
D) core layer
E) aggregation layer

Q8) Each Cisco ACE Module context has its own configuration file and operational data, providing complete isolation from other contexts on the data level only. (Source: Designing Application Services)

A) true
B) false

Q9) What are three secure load-balancing solutions that are supported by Cisco ACE Module? (List three.) (Source: Designing Application Services)

Q10) Which option is the primary protocol that is used in the Cisco global load-balancing solution? (Source: Designing Global Load Balancing)

A) ARP
B) DRP
C) DNS
D) TCP

Q11) Which two flows does the Cisco GSS need to establish? (Choose two.) (Source: Designing Global Load Balancing)

A) sessions to the firewall
B) KAL-AP flow from the GSS to the Cisco ACE Module
C) sessions to D-proxy servers
D) sessions to physical servers

Q12) Which three algorithms are used to select the best possible site to perform the client request? (Choose three.) (Source: Designing Global Load Balancing)

A) network proximity
B) round robin
C) server response time
D) ordered list
E) client browser version

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

Module Self-Check Answer Key

Q1) B

Q2) C

Q3) A-2
B-1
C-3

Q4) A, B

Q5) A, C, D

Q6) A

Q7) C

Q8) B

Q9) SSL termination, SSL initiation, and a combination of SSL termination and SSL initiation

Q10) C

Q11) B, C

Q12) A, B, D

https://t.me/learningnets

certcollection.net

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

Module 7

Data Center Management

Overview

In this module, you will learn about data center management and monitoring. When data centers grow large and the number of devices reaches the hundreds or thousands, good management tools are essential to successfully configure and operate the data center. Management tools provide consistent configuration and an aid in troubleshooting.

Network monitoring is a building block that gives you insight regarding the traffic that goes through network devices. These devices are physical or virtual and use similar mechanisms to report about data traffic and load.

Module Objectives

Upon completing this module, you will be able to design a data center management solution to facilitate monitoring, managing, and provisioning data center equipment and applications. This ability includes being able to meet this objective:

- Present data center management software and solutions

https://t.me/learningnets

certcollection.net

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

Lesson 1

Designing Data Center Management Solutions

Overview

Network management is one of the crucial elements that helps keep the network operational and under control. There are several tools that are available to manage and monitor data center equipment. Network management software can also collect statistics that you can use for network planning and other functions.

Objectives

Upon completing this lesson, you will be able to explain data center management software and solutions. This ability includes being able to meet these objectives:

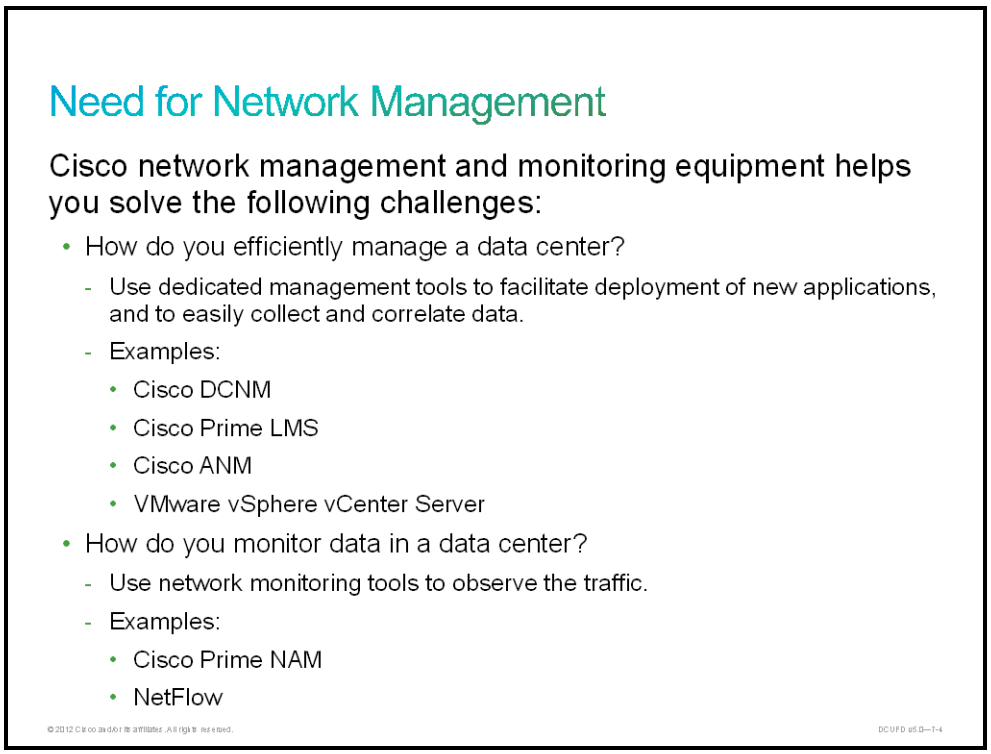
- Describe the need for network management
- Describe Cisco Data Center management products
- Describe scalability limitations
- Secure management in multitenant environments

https://t.me/learningnets

certcollection.net

Need for Network Management

This topic describes the need for network management.



To efficiently manage a data center, you should use dedicated management tools, such as Cisco Prime Data Center Network Manager (DCNM), Cisco Prime LAN Management Solution (LMS), Cisco Application Networking Manager (ANM), or VMware vSphere vCenter Server, to facilitate deployment of new applications and to easily collect and correlate data.

To monitor data in a data center, you should use monitoring tools, such as Cisco Prime Network Analysis Module (NAM) or NetFlow, to observe the traffic.

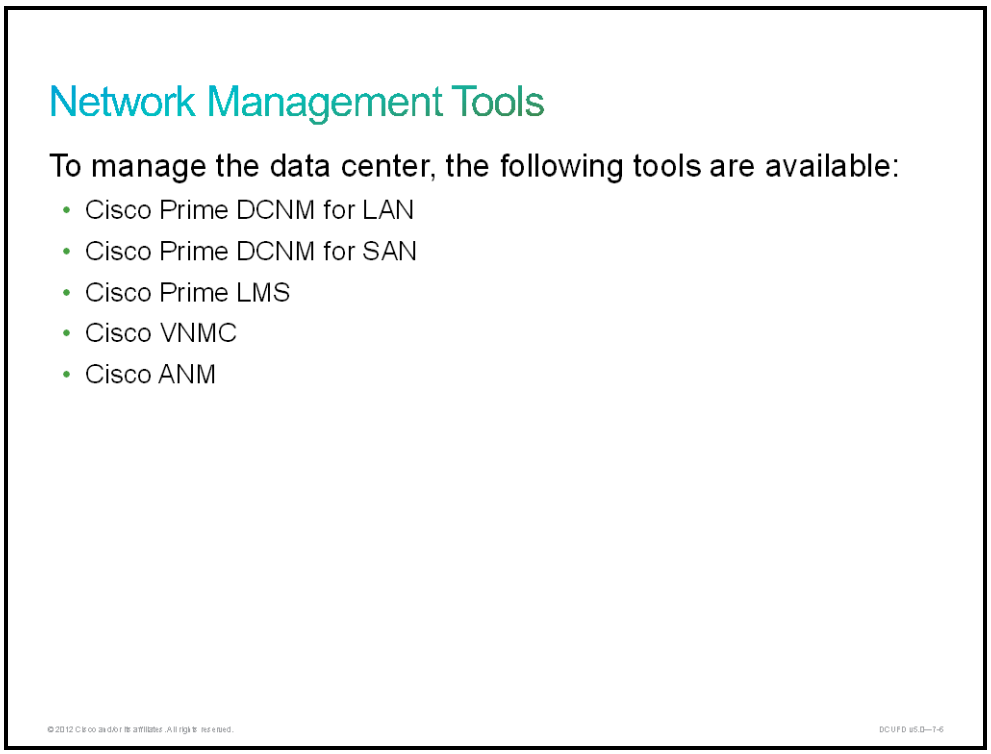
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

Cisco Data Center Management Tools

This topic describes Cisco Data Center management products.



Network management itself is not difficult until you manage a few devices. When the number of devices grows, you need a tool to consistently manage your devices.

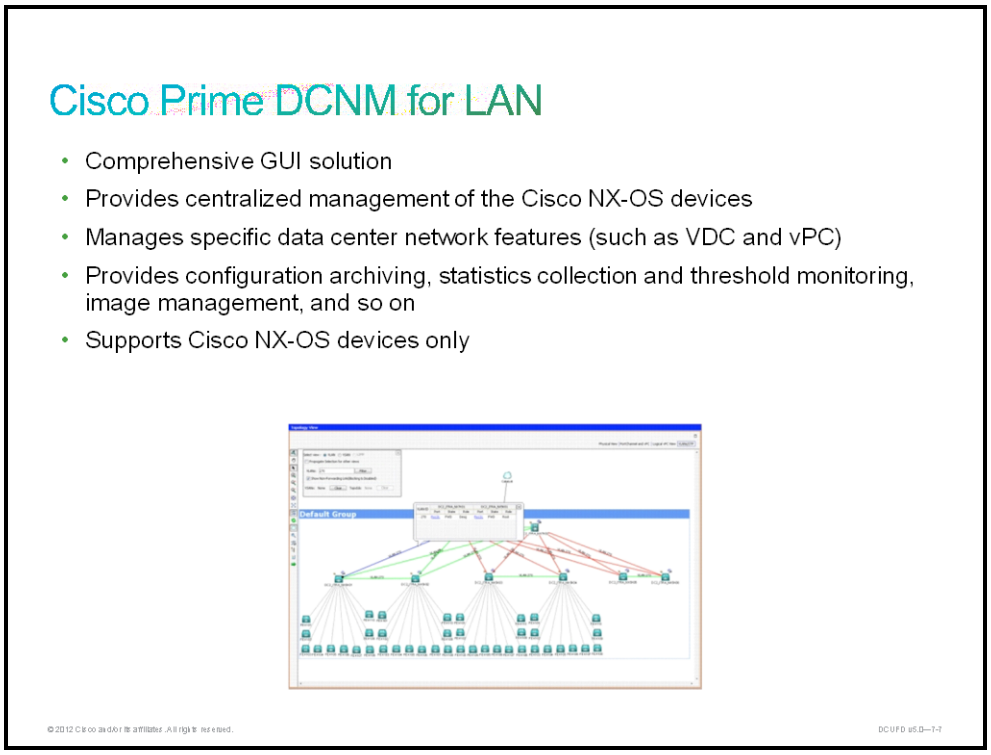
There are several tools to manage the data center:

- **Cisco Prime DCNM for LAN:** Cisco Prime DCNM is a Cisco management solution that increases overall data center infrastructure uptime and reliability, which improves business continuity. Focused on supporting efficient operations and management of the data center network, Cisco Prime DCNM provides a robust framework and plentiful feature set that meets the routing, switching, and storage administration needs of present and future data centers. In particular, Cisco Prime DCNM automates the provisioning process, proactively monitors the SAN and LAN by detecting performance degradation, streamlines the diagnosis of dysfunctional network elements, and secures the network. Offering an exceptional level of visibility and control through a single pane to Cisco Nexus and Cisco MDS 9000 family products, Cisco Prime DCNM is the Cisco recommended solution for mission-critical data centers.
- **Cisco Prime DCNM for SAN (previously known as Cisco Fabric Manager):** Cisco Prime DCNM for SAN is the management tool for storage networking across all Cisco SAN and unified fabrics.
- **Cisco Prime LMS (previously known as CiscoWorks LMS):** Cisco Prime LMS is a suite of powerful management tools that simplifies the configuration, administration, monitoring, and troubleshooting of Cisco networks.
- **Cisco Virtual Network Management Center (VNMC):** Cisco VNMC is a virtual appliance that provides centralized device and security policy management for Cisco Virtual Security Gateway (VSG) for Cisco Nexus 1000V Series Switches.

https://t.me/learningnets

certcollection.net

- **Cisco ANM:** Cisco ANM software is part of the Cisco Application Control Engine (ACE) Module product family and is a critical component of any data center or cloud-computing architecture that requires centralized configuration, operation, and monitoring of Cisco Data Center networking equipment and services. Cisco ANM provides this management capability for the Cisco ACE appliances, as well as operations management for the Cisco Content Services Switch (CSS), Cisco Content Switching Module (CSM), Cisco CSM with SSL (CSM-S), and Cisco ACE Global Site Selector (GSS). It also integrates with VMware virtual data center environments, providing continuity between the application server and network operator and increasing the application network services awareness and capabilities of the operators, while reducing the burden of operating and managing those services.



The Cisco Prime DCNM for LAN is an easy way of utilizing the management application for the Cisco Nexus Operating System (Cisco NX-OS)-based devices only. It is designed to provide centralized management of Cisco NX-OS-based data center networking devices.

Cisco Prime DCNM is able to manage Cisco NX-OS specific features, such as Cisco Nexus 7000 virtual device contexts (VDCs), virtual port channels (vPCs), and so on.

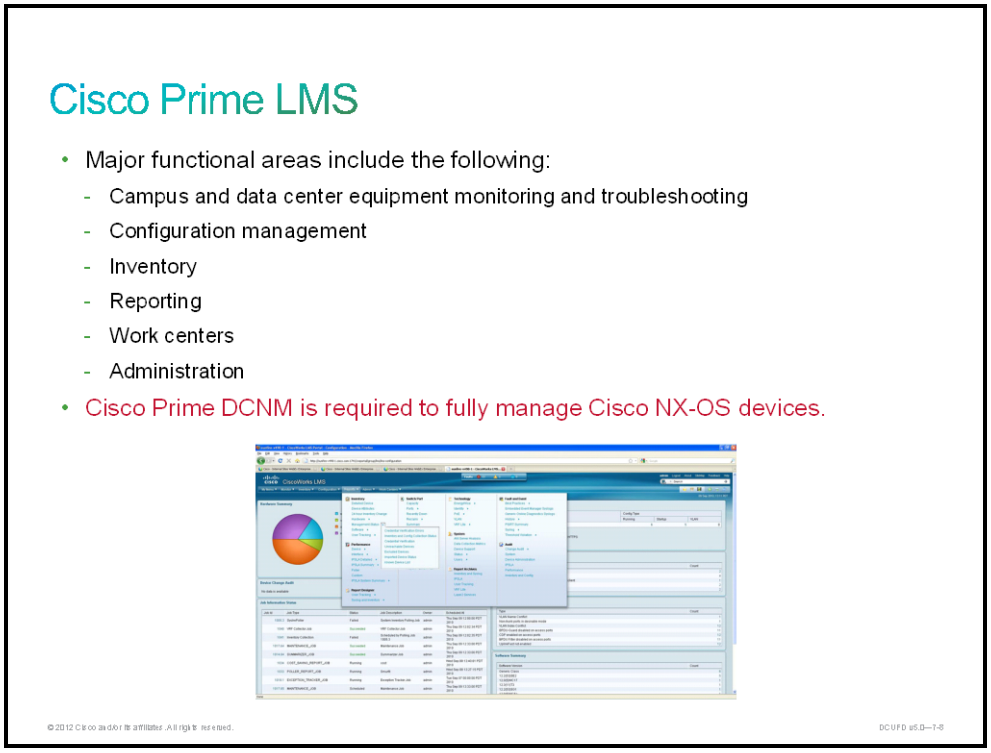
Cisco Prime DCNM provides fault management, configuration management, accounting, performance, and security management functions like Fault, Configuration, Accounting, Performance, and Security (FCAPS).

Cisco Prime DCNM follows the corresponding Cisco NX-OS releases. For example, Cisco Prime DCNM 6.0 is the appropriate version to use with Cisco NX-OS version 6.0. Cisco Prime DCNM 5.1 cannot manage Cisco NX-OS 6.0 devices.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net



Cisco Prime LMS has evolved from a collection of individual products into a seamless set of integrated management functions that is based upon the way network managers do their work. Organizing the product based on management function simplifies the overall user experience by reducing the need to cross application boundaries to complete a specific management task. Workflows are self-contained and all required functionality is maintained within a functional area. The major functional areas include the following:

- **Monitoring and troubleshooting:** Quickly and proactively identify and fix network problems before they affect end users or services.
- **Configuration management:** Configuration backup, software image management, compliance, and change management are required to maintain and update network devices.
- **Inventory:** Complete a thorough inventory of all Cisco equipment details, such as chassis, module, and interface.
- **Reporting:** All reports are centralized in a single menu, simplifying navigation and access to detailed reports and information.
- **Work centers:** End-to-end life-cycle management of Cisco value-added technologies, such as deployment, monitoring, and management of Cisco EnergyWise, Cisco TrustSec Identity, Cisco Auto Smartports, and Cisco Smart Install.
- **Administration:** Getting started and improved workflows simplify application setup and administration.

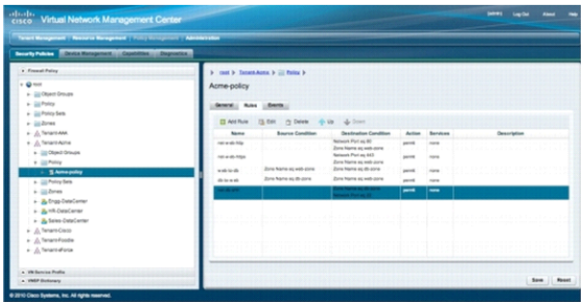
Note	Cisco Prime LMS recognizes the Cisco NX-OS devices (Cisco Nexus switches), but does not manage them fully. Cisco NX-OS devices need to be managed using Cisco Prime DCNM.
Note	Many of the management features are licensed. You need a license to manage Cisco Nexus 7000 VDCs, vPCs, Cisco Nexus 5000 Series Switches, and so on.

https://t.me/learningnets

certcollection.net

Cisco VNMC

- Management tool for Cisco VSG for Cisco Nexus 1000V switch
- Centralized security policy definition for a multitenant VMware-virtualized environment
- Deployment of policies into the virtual access layer to the VSG and then to the Cisco Nexus 1000V VEM
- Simplifies collaboration across security and server teams



© 2012 Cisco and/or its affiliates. All rights reserved.

Designed for enterprise and multitenant cloud deployments, Cisco VNMC offers transparent, scalable, and automation-centric management for securing virtualized data center and cloud environments. With both a built-in GUI and an XML application programming interface (API), centralized management of Cisco VSG can be performed by an administrator or programmatically.

Cisco VNMC provides these main benefits:

- Rapid and scalable deployment through dynamic, template-driven policy management that is based on security profiles
- Policies get applied to multiple VSGs, belonging to security domains that apply the security policy to a particular port group to which a virtual machine connects
- Collaboration across security and server teams while maintaining administrative separation and reducing errors via a consistent and repeatable deployment model

https://t.me/learningnets

certcollection.net

Cisco ANM

- Cisco ANM helps customers perform the following processes:
  - Managing data center application services effectively
  - Configuring Cisco ACE Module
  - Troubleshooting
  - Maintenance
  - Operations
  - Monitoring
- Standalone product for Cisco ACE Module; can manage the Cisco ACE 4710 appliance as well



© 2012 Cisco and/or its affiliates. All rights reserved.

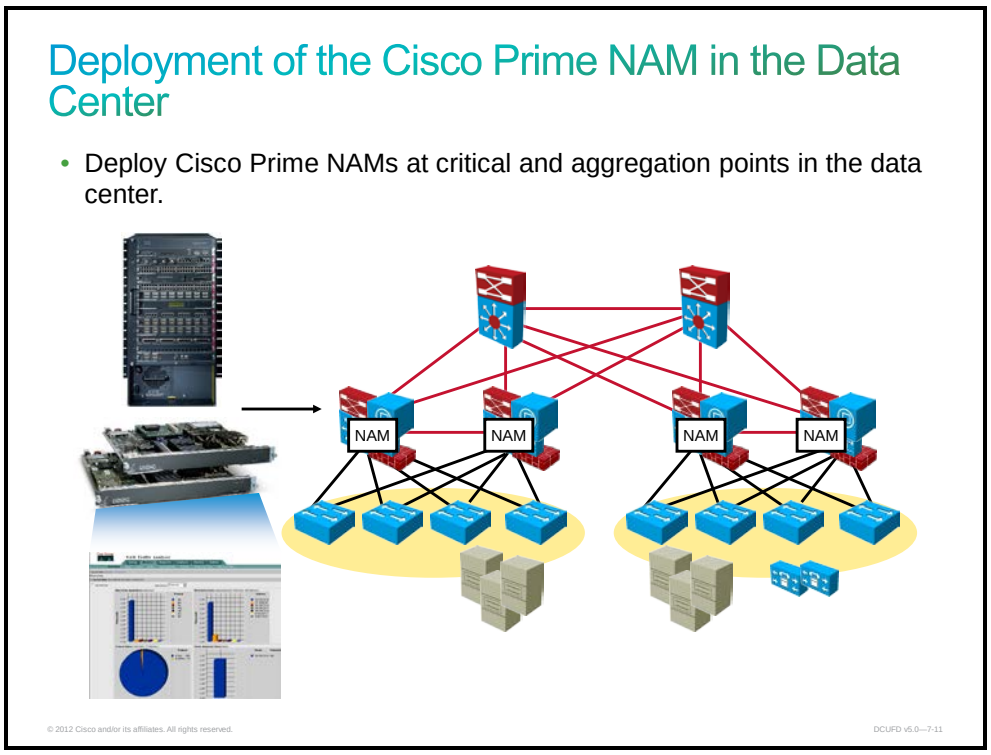
DCUFD v5.0—7-10

Cisco ANM helps customers manage multidevice data for Cisco ACE Module troubleshooting, maintenance, operations, and monitoring. It also unifies the operations center network services effectively. By using Cisco ANM, customers can simplify the deployment and ongoing maintenance of their Cisco ACE Module virtualized environment, providing a unified interface management and monitoring of real and virtual servers spanning a load-balancing infrastructure of Cisco ACE Module, CSS, CSM, and CSM-S devices. Cisco ANM also centralizes operations management of virtual IP answers and Domain Name System (DNS) rules for Cisco ACE GSS devices.

Cisco ANM is ideal for enterprises and service providers that implement Cisco ACE Module and provides additional value to customers using Cisco CSS, CSM, CSM-S, and Cisco ACE GSS devices. These customers include data center infrastructure providers, application service providers, large enterprises, and e-business data centers. Even small and medium-sized enterprises with small deployments of Cisco ACE appliances can benefit from Cisco ANM through the entry-point offering.

https://t.me/learningnets

certcollection.net



Collecting the data that you need is made easier by the flexibility of Cisco Prime NAM to be placed where it is needed and where it can gather data from either local or remote switches and routers. Typical deployment places for Cisco Prime NAM include LAN aggregation points, where it can collect the most data, service points (server farms, data centers, and so on), where performance is critical, and important access points. Of course, actual placement depends on the problems that you are trying to solve with Cisco Prime NAM. As shown in the figure, the Cisco Catalyst 6500 Series Switch NAM can be complimented with the Cisco Branch Routers Series NAM and the network module NM-NAM for monitoring WANs.

Cisco Catalyst 6500 Series Switch NAMs

The Cisco Catalyst 6500 Series Switches can host NAM-1, NAM-2, or NAM-3. These Cisco Catalyst 6500 Series Switch NAM modules can collect and display per-port Layer 2 statistics with mini-Remote Monitoring (mini-RMON) on every interface. You can achieve more in-depth analysis of LAN ports by spanning or copying traffic from ports, VLANs, or EtherChannel to the embedded Cisco Catalyst 6500 Series Switch NAM, or by using VLAN access control lists (VACLs) to mirror data to the Cisco Catalyst 6500 Series Switch NAM if no spanning sessions are available.

You can analyze remote switches using the Remote Switched Port Analyzer (RSPAN) feature of Cisco Catalyst switches. You can achieve a detailed analysis of WAN ports by using VACLs on a local device or by forwarding NetFlow data from either the local or a remote device.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

The Cisco Catalyst 6500 Series Switch NAMs are vital tools that provide high performance to monitor traffic that is running at sub-gigabit speeds (NAM-1) and gigabit speeds (NAM-2). The Cisco Catalyst 6500 Series Switch NAMs can be deployed in the following areas:

- distribution or core layer trunk ports
- service points (for example, in data centers, server farms, or Cisco Communications Manager clusters in IP telephony) where performance is critical
- critical access points

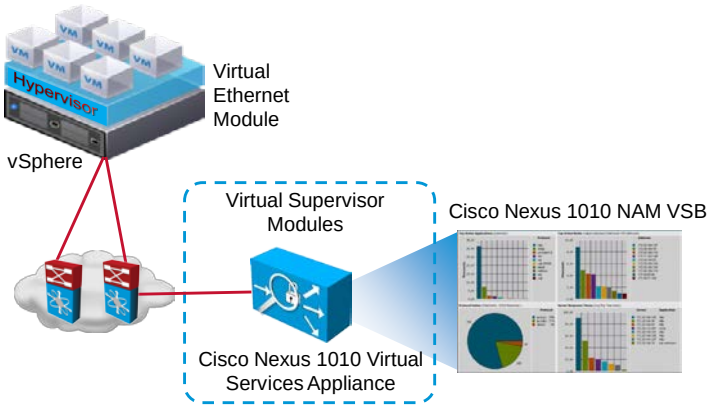
Placement and intended use can dictate the need for the higher-performance NAM-2 or NAM-3.

https://t.me/learningnets

certcollection.net

Cisco Nexus 1010 NAM VSB

- Simplifies manageability of the virtual switching infrastructure
- Comprehensive visibility into the virtual environment
- Requires the Cisco Nexus 1010 Virtual Services Appliance



© 2012 Cisco and/or its affiliates. All rights reserved.

The Cisco Nexus 1010 NAM Virtual Service Blade (VSB) allows network administrators to extend operational visibility into Cisco Nexus 1000V switch deployments.

Integrated with the Cisco Nexus 1010 Virtual Services Appliance, this virtual service blade simplifies manageability of the virtual switching infrastructure. It offers comprehensive visibility into the virtual environment to meet the service delivery challenges in next-generation data centers.

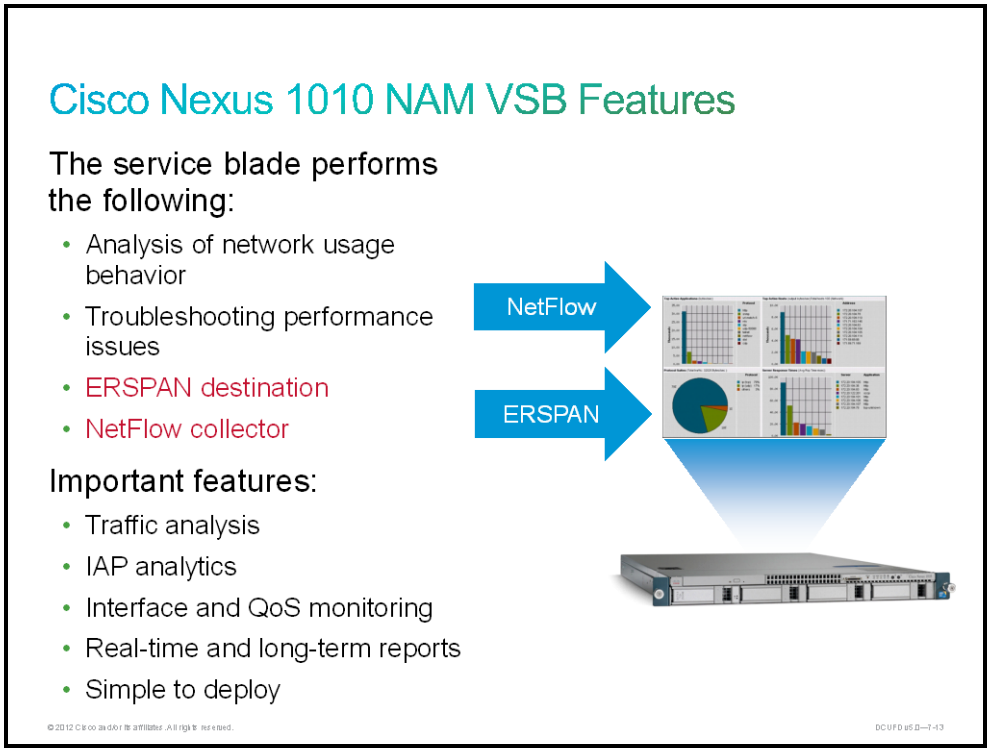
As flexible advanced Cisco instrumentation, the Cisco Catalyst 6500 Series Switch NAMs can be deployed at places in the network that are necessary for end-to-end network and application performance visibility. For example, a Cisco Nexus 1010 NAM VSB is deployed with the Cisco Nexus 1010 appliance in the data center for operational visibility into Cisco Nexus 1010 deployments. This integrated solution allows you to monitor virtual network behavior and analyze communication across virtual machines to gain performance visibility into applications that are deployed in a virtual computing environment.

The intelligence from the Cisco Nexus 1010 NAM VSB can optionally be combined with other NAM form factors such as the Cisco Catalyst 6500 Series Switch NAM, the Cisco NAM appliance, or Cisco Branch Routers Series NAM that are deployed in the data center, campus, or remote sites to provide enterprise-wide visibility.

The Cisco Catalyst 6500 Series Switch NAM can export computed performance information to third-party and homegrown applications to meet end-to-end performance reporting needs. Third-party applications gather application and network performance information from Cisco Catalyst 6500 Series Switch NAMs that are deployed across the network for consolidated networkwide reporting. Such applications complement the granular performance visibility that is offered by Cisco Catalyst 6500 Series Switch NAMs to help enable you to monitor how applications are being delivered enterprise-wide, yet isolate and resolve delivery problems proactively and promptly at their source.

https://t.me/learningnets

certcollection.net



The Cisco Catalyst 6500 Series Switch NAM VSB allows you to effectively use embedded management features, such as Encapsulated RSPAN (ERSPAN) and NetFlow, on the Cisco Nexus 1000V Switch to perform the following:

- Analyze conversation and network usage behavior by application, host, or virtual machine (VM) to identify bottlenecks that may affect performance and availability
- Troubleshoot performance issues with extended visibility into VM-to-VM traffic, virtual interface statistics, and transaction response times
- Improve the efficiency of the virtual infrastructure and distributed application components with deeper operational insight

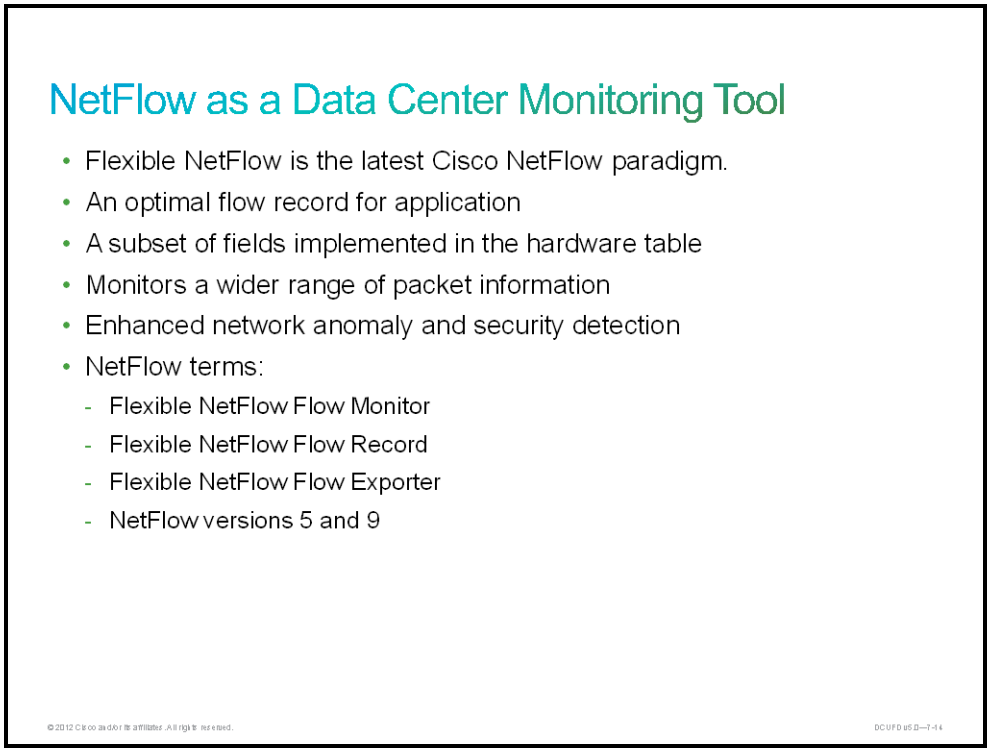
Note The Cisco Nexus 1010 NAM VSB can be a NetFlow *collector* device. The Cisco Nexus 1000V Switch can be a NetFlow *source*.

There are various important features of the Cisco Nexus 1010 NAM VSB:

- Traffic analysis
- Intelligent application performance (IAP) analytics
- Interface and quality of service (QoS) monitoring
- Real-time and long-term reports
- Simple deployment

https://t.me/learningnets

certcollection.net



Flexible NetFlow is the most recent Cisco NetFlow paradigm. It is a very flexible way of configuring NetFlow in the network or to define a flow record that is optimal for a particular application. Definition is done by selecting the keys from a large collection of predefined fields. Not all of the fields are supported. A subset of Flexible NetFlow key and nonkey fields, based on support, is provided by the forwarding engine. Therefore, only subsets of fields, which are implemented in the hardware table, are supported.

The Flexible NetFlow-based configuration model includes the following:

- Create flow record
- Create exporter
- Combine flow record and exporter in flow monitor
- Tie monitor to interface

NetFlow offers the ability to monitor a wider range of packet information, producing new information about network behavior. Enhanced network anomaly and security detection is available as well. NetFlow is configured on the interface, not globally, which is the preferred way to verify just a few interfaces.

The NetFlow terms are Flexible Flow Monitor, Flexible NetFlow Flow Record, Flexible NetFlow Flow Exporter, and NetFlow versions 5 and 9.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

Flexible NetFlow Flow Monitor

A Flexible NetFlow Flow Monitor is essentially a NetFlow cache. The Flexible NetFlow Flow Monitor has two major components: the Flexible NetFlow Flow Record and the Flexible NetFlow Flow Exporter. The Flexible NetFlow Flow Monitor can track both ingress and egress information. The Flexible NetFlow Flow Record contains the information that is being tracked by NetFlow (that is, IP address, ports, protocol, and so on). The Flexible NetFlow Flow Exporter describes the NetFlow export. Flexible NetFlow Flow Monitors can be used to track IP version 4 (IPv4) traffic, IP version 6 (IPv6) traffic, multicast or unicast traffic, Multiprotocol Label Switching (MPLS) traffic, or bridged traffic. Multiple Flexible NetFlow Flow Monitors can be created and attached to a specific physical or logical interface. Flexible NetFlow Flow Monitors can also include packet sampling information if sampling is required.

Flexible NetFlow Flow Record

A Flexible NetFlow Flow Record defines what information NetFlow will track. The Flexible NetFlow Flow Record can be user-defined or a predefined scheme that is available in Cisco IOS Software. The Flexible NetFlow Flow Record is defined as a set of key and nonkey fields. Typical NetFlow key fields are IP addresses and ports, and, if the set of key fields is unique, a new flow is created. The nonkey field information is collected and attached to the flow. Typical nonkey fields include time stamps, packet and byte counters, and TCP flag information. Essentially, the Flexible NetFlow Flow Record tells NetFlow what information to obtain from the packets that are being forwarded.

Flexible NetFlow Flow Exporter

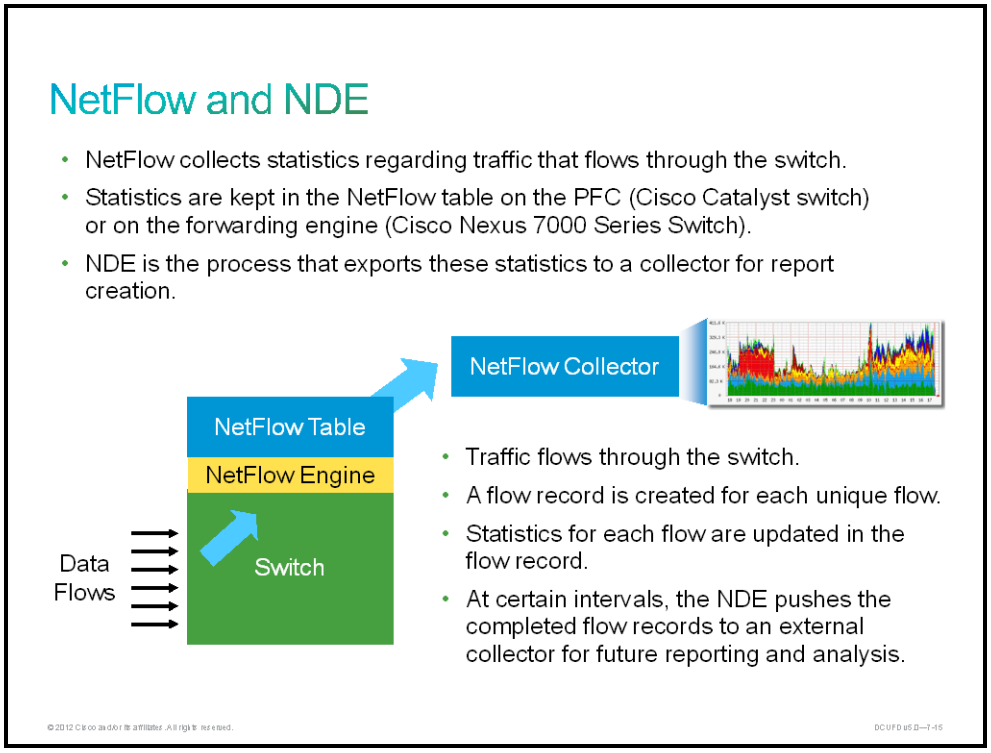
A Flexible NetFlow Flow Exporter describes information about the NetFlow export that is sent to the reporting server or NetFlow collector. The Flexible NetFlow Flow Exporter includes the destination address of the reporting server, the type of transport (that is, UDP or Stream Control Transmission Protocol [SCTP]), and the export format (that is, version 9). There can be multiple exporters per Flexible NetFlow Flow Monitor. The Flexible NetFlow Exporter is QoS-aware. The export stream is prioritized with other traffic based on its class of service (CoS) or differentiated services code point (DSCP) value.

NetFlow Versions 5 and 9

NetFlow exports information to reporting servers in various formats, including NetFlow versions 5 and 9. NetFlow version 5 is used with traditional NetFlow and is a fixed export format with a limited set of information being exported. NetFlow version 9 is a flexible and extensible NetFlow format that is used by Flexible NetFlow. NetFlow version 9 includes a template to describe what is being exported and the export data. The template is periodically sent to the NetFlow collector to tell it what data to expect from the router or switch. The data is then sent for the reporting system to analyze. Because NetFlow version 9 is extensible and flexible, any data that is available in the device can theoretically be sent in NetFlow version 9 formats. Flexible NetFlow allows the user to configure and customize the information that is exported using NetFlow version 9. NetFlow version 9 is the basis for the IETF standard IP Flow Information Export (IPFIX) that is associated with the IP Flow and Information working group in IETF.

https://t.me/learningnets

certcollection.net



NetFlow collects global statistics from traffic that flows through the switch and stores those statistics in the NetFlow table.

The NetFlow table is populated within the forwarding engine: the PFC3C or PFC3CXL on the Catalyst 6500 Switches, and on the M1 forwarding engine on the Cisco Nexus 7000 Series Switch.

The Cisco F2 and M2 forwarding engines support NetFlow as well, with up to 256 programmable sampling rates.

The Cisco Nexus 5000 and 5500 Series Switches do not support NetFlow. Generally, this is not a significant issue because most of the traffic monitoring using NetFlow is done at the core layer.

The Cisco Nexus 1000V Switch can also run NetFlow, and the collection process runs in software (Virtual Supervisor Module [VSM] and Virtual Ethernet Module [VEM]).

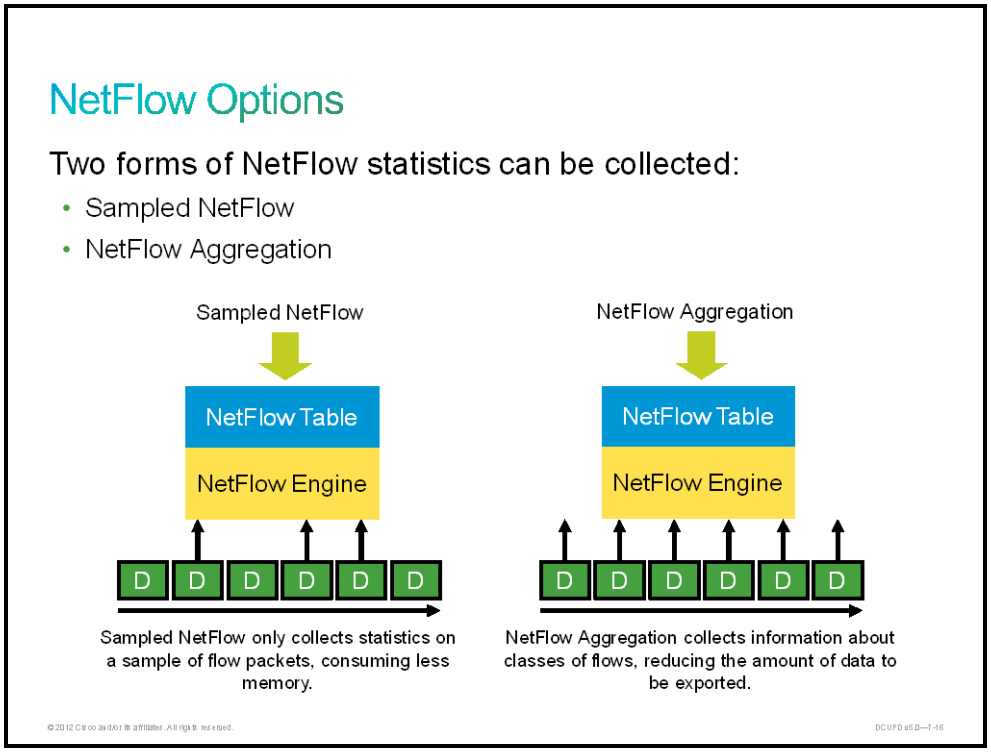
NetFlow Data Export (NDE) makes traffic statistics available for analysis by an external data collector.

Several external data collector addresses can be configured to provide redundant data streams to improve the probability of receiving complete NetFlow data.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net



Two NetFlow options are available to reduce the volume of statistics being collected:

- Sampled NetFlow reduces the number of statistics collected.
- NetFlow Aggregation merges collected statistics.

Sampled NetFlow

The Sampled NetFlow feature captures a subset of traffic in a flow, instead of all packets within a flow on Layer 3 interfaces. Sampled NetFlow substantially decreases the supervisor engine CPU utilization.

NetFlow Aggregation

The NetFlow Aggregation feature allows limited aggregation of NDE streams on a Cisco Catalyst 6500 Series Switch. This is achieved by maintaining one or more extra flow caches called “aggregation caches.”

There are benefits of using NetFlow Aggregation:

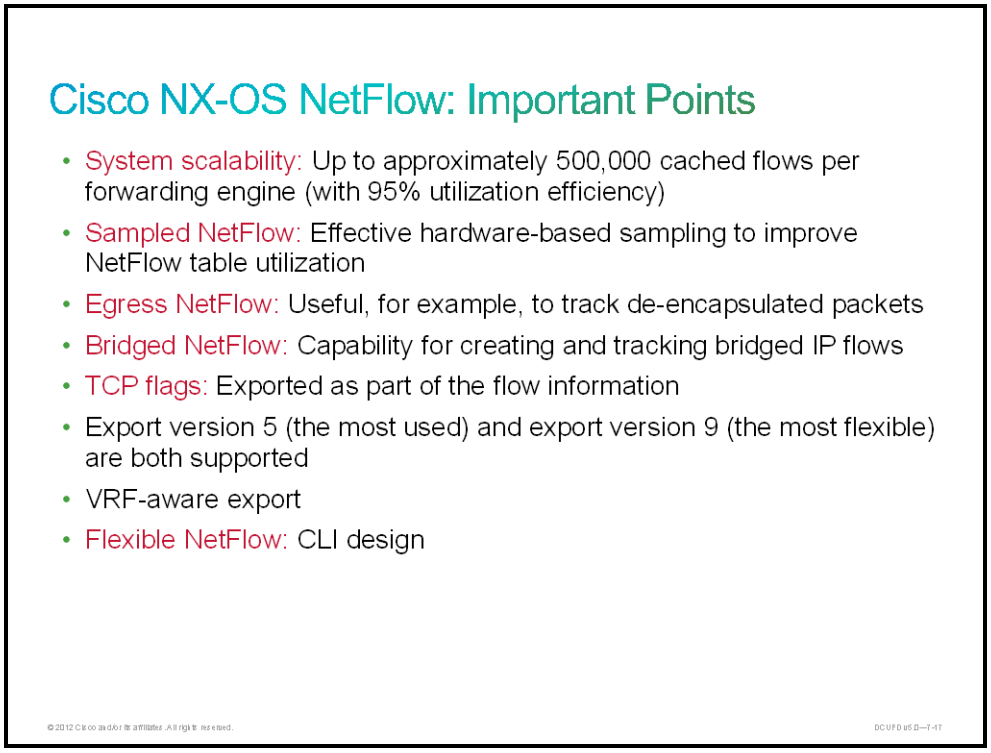
- **Reduced bandwidth requirement:** NetFlow Aggregation caches reduce the bandwidth that is required between the switch and the NetFlow management station.
- **Reduced NetFlow workstation requirements:** NetFlow Aggregation caches reduce the number of NetFlow management workstations that are required.
- **Improved scalability:** NetFlow Aggregation caches improve scalability for high-flow-per-second devices such as the Cisco Catalyst 6500 Series Switch.

Each aggregation cache can be configured with its own individual cache size, cache timeout parameter, export destination IP address, and export destination UDP port.

Note	NetFlow Aggregation uses NetFlow packets version 8 for exporting. You must verify compatibility with the collector.

https://t.me/learningnets

certcollection.net



Systems scalability is up to 500,000 cached flows per forwarding engine. With a fully loaded chassis, the Cisco Nexus 7010 Switch is able to cache four million flows. This is a significant step compared to Cisco Catalyst 6500 Series Supervisor Engine 720, where 128,000 entries were supported. In the NetFlow table, the system stores only those packets that are sampled, so the NetFlow table is not populated with meaningless information. Effective hardware-based sampling is used to improve NetFlow table utilization. This is one of the advantages compared to the Cisco Catalyst 6500 Series Switch, where all the flows go into the table and only the flows that are sampled are exported.

Egress NetFlow and bridged NetFlow are supported. Egress NetFlow is used to track de-encapsulated packets. Bridged NetFlow is used to create and track bridged IP flows. Additionally, TCP flags are supported, and they are exported as part of the flow information. This information is very useful to understand TCP flow directions and to detect denial of service (DoS) attacks.

Note	In terms of export version format, export versions 5 and 9 are supported. These are the most used (version 5) and the most flexible (version 9) formats. Export version 5 is the default version.

NetFlow exporting is virtual routing and forwarding (VRF)-aware. Specified destinations can be sent for the export and it is possible to define a VRF to which the administrator wants to send the export.

The programming method is important as well. It is based on Flexible NetFlow, which is a new paradigm in the Cisco NetFlow progression.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

NetFlow and the Cisco Prime NAM

- The Cisco Prime NAM can be a NetFlow collector device.
- The Cisco Prime NAM provides application-level visibility into network traffic:
  - Capture
  - Reduce
  - Analyze
  - Trend
- NetFlow provides metering of traffic statistics.



© 2012 Cisco and/or its affiliates. All rights reserved.

The Cisco Catalyst 6500 Series Switch NAM is an integrated traffic monitoring solution for the Cisco Catalyst 6500 Series Switches, Cisco 7600 Series Routers, and some branch routers. The Cisco Catalyst 6500 Series Switch NAM enables network managers to gain application-level visibility into network traffic to improve performance and reduce failures.

The Cisco Catalyst 6500 Series Switch NAM facilitates these functions:

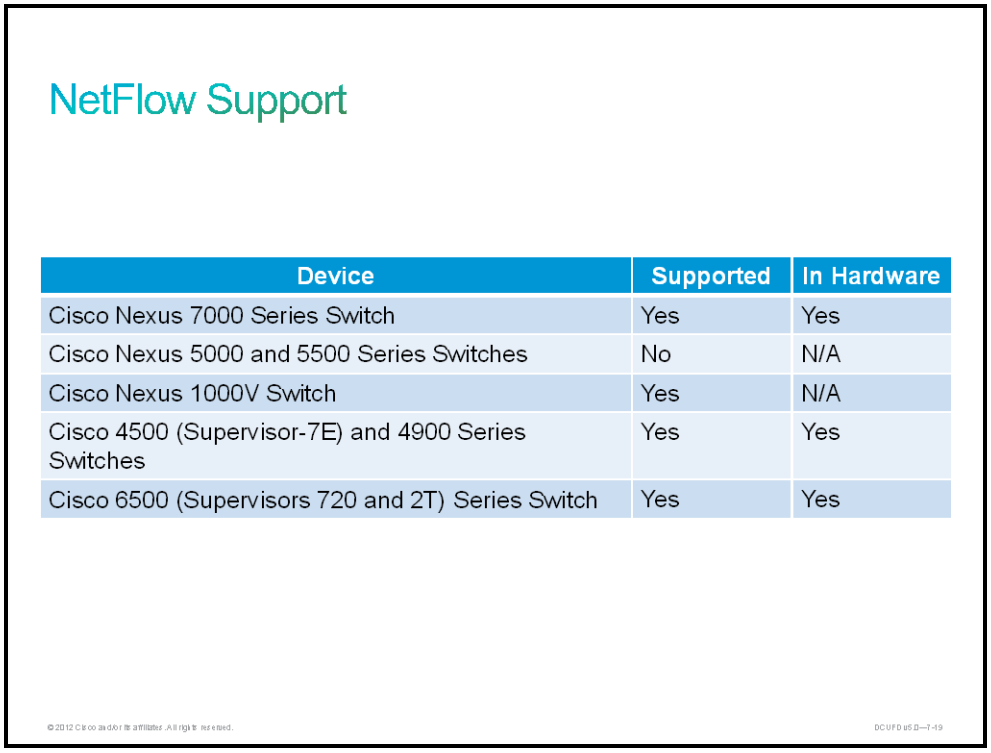
- **Capture:** Performs raw data capture
- **Reduce:** Reduces captured data to useful information
- **Analyze:** Assists in drawing conclusions about reduced data
- **Trend:** Maintains ongoing statistics on incremental data captures for long-term planning

NetFlow technology provides the metering base for an important set of applications including network traffic accounting, usage-based network billing, and network planning, as well as DoS monitoring capabilities, network monitoring, outbound marketing, and data-mining capabilities. Cisco provides a set of NetFlow applications to collect NetFlow export data, perform data volume reduction, and do post-processing.

The Cisco Catalyst 6500 Series Switch NAM and NetFlow work together. NetFlow traffic statistics are exported to the Cisco Catalyst 6500 Series Switch NAM without affecting network device performance, and the Cisco Catalyst 6500 Series Switch NAM performs data reduction.

https://t.me/learningnets

certcollection.net



The table in this figure lists the platforms that support NDE:

- The Cisco Nexus 7000 Series Switch supports NDE, and data about the traffic is collected in hardware on the M1 or M1-XL forwarding engine.
- The Cisco Nexus 5000 and 5500 Series Switches do not support NetFlow.
- The Cisco Nexus 1000V Switch supports NetFlow and can collect data about the traffic on a per-VM basis.
- The Cisco Catalyst 4500 Series Switches support NetFlow in hardware with the latest Supervisor Engine 7-E. The Supervisor Engine 6 does not support NetFlow, and the Supervisor Engine 5 supports NetFlow only in software.
- The Cisco Catalyst 4900 Series Switches support NetFlow in hardware.
- The Cisco Catalyst 6500 Series Switches support NetFlow in hardware. The NDE is performed on the Policy Feature Card (PFC) if centralized forwarding is used, or on the distributed forwarding cards (DFCs) if distributed forwarding is used.

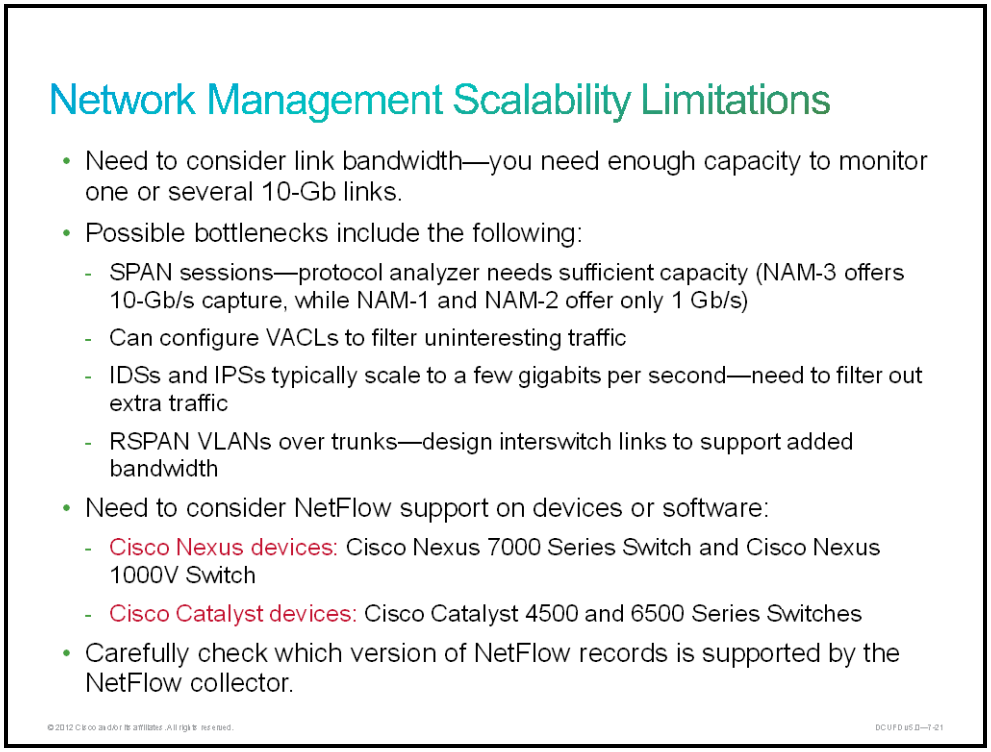
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

Network Management Scalability Limitations

This topic describes network management scalability limitations.



When using network monitoring and management tools in the data center, you need to consider scalability limitations, such as link bandwidth, resources availability, and so on.

For example, if you want to capture and analyze traffic on a 10-Gb link, you need a device with enough capacity to process flows with such a high bandwidth.

There are some possible bottlenecks:

- **SPAN sessions:** The protocol analyzer needs sufficient capacity (NAM-3 offers 10-Gb/s capture; NAM-1 and NAM-2 offer only 1 Gb/s). If you do not have enough capacity, you can configure a VACL to filter uninteresting traffic.
- **Intrusion detection systems (IDSs) and intrusion prevention systems (IPSs) typically scale to a few gigabits per second:** You need to filter out extra traffic (that is, traffic for IP-based storage [network-attached storage (NAS) or Internet Small Computer Systems Interface (iSCSI)]) or streaming video traffic.
- If you are using RSPAN VLANs over regular, production trunks, you need to provision enough bandwidth on Inter-Switch Links.

When designing a network monitoring solution using NetFlow, verify NetFlow support on devices or software:

- **Cisco Nexus devices:** Cisco Nexus 7000 Series Switch and Cisco Nexus 1000V Switch
- **Cisco Catalyst devices:** Cisco Catalyst 4500, 4900, and 6500 Series Switches

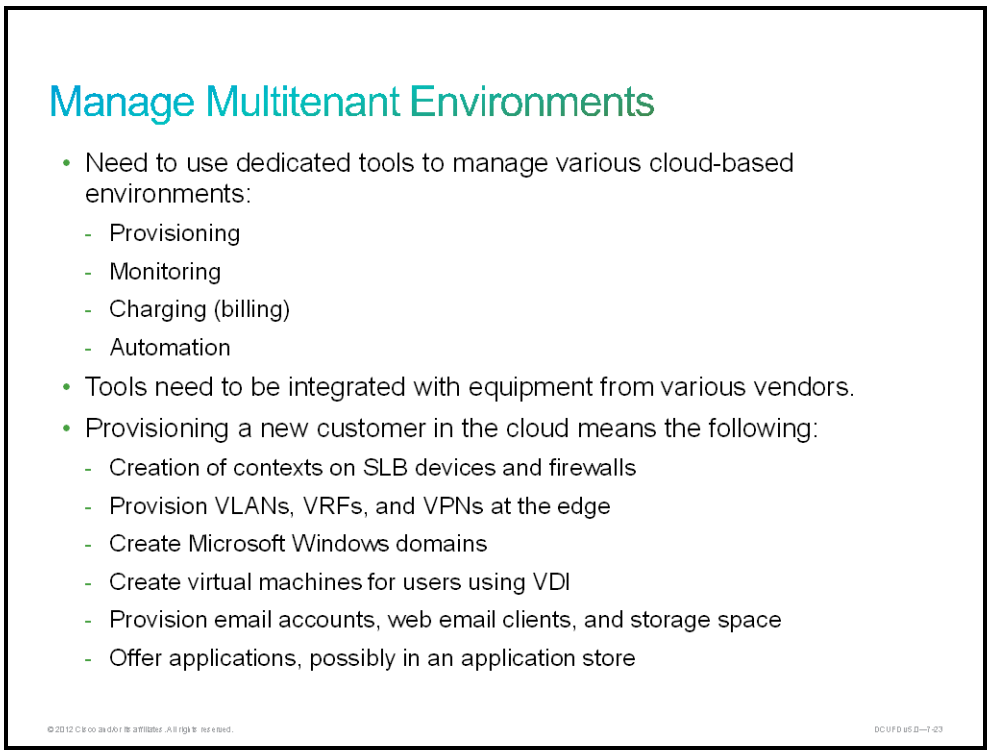
Carefully check which version of NetFlow records is supported by the NetFlow collector.

https://t.me/learningnets

certcollection.net

Manage Multitenant Environments

This topic describes how to secure management in multitenant environments.



When designing management for a multitenant data center, as found in various cloud-based solutions, you must emphasize management, monitoring, provisioning, and charging systems.

These tools must be integrated with multiple appliances, possibly from different vendors.

For example, the following is required if you want to add another customer to your virtual desktop infrastructure (VDI)-based cloud solution:

- Create contexts on server load-balancing (SLB) devices and firewalls
- Provision VLANs, VRFs, and VPNs at the edge
- Create Microsoft Windows domains
- Create virtual machines for users using VDI
- Provision email accounts, web email clients, and storage space
- Offer applications, possibly in an application store

Management of such systems is very complex and requires a lot of customization and integration work.

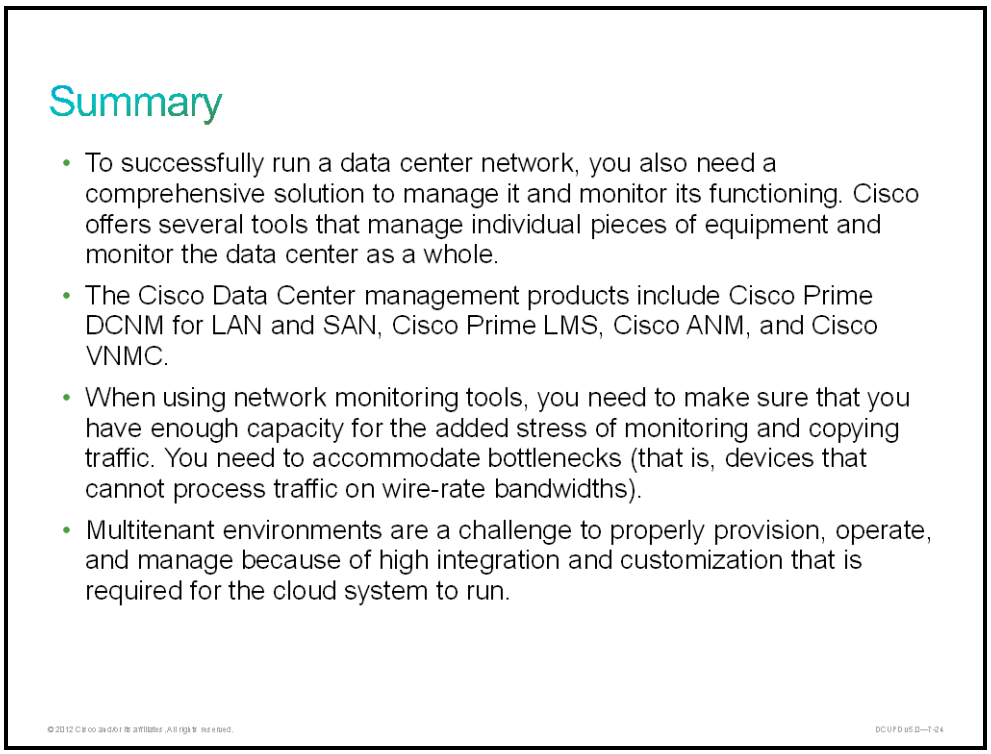
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

Summary

This topic summarizes the primary points that were discussed in this lesson.



https://t.me/learningnets

certcollection.net

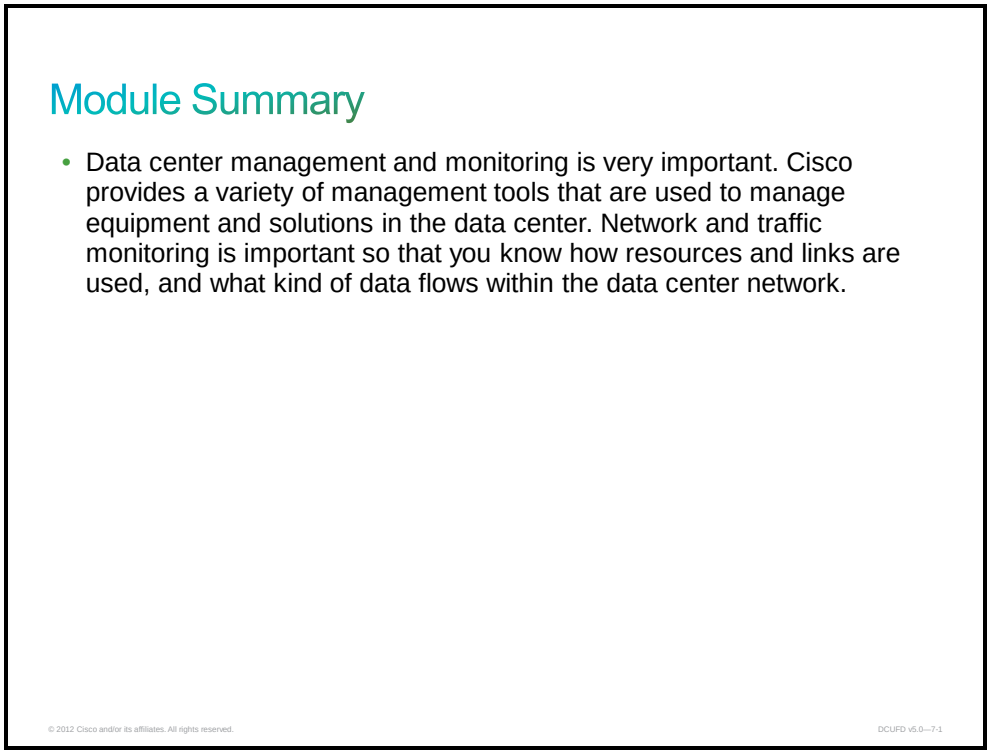
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

Module Summary

This topic summarizes the primary points that were discussed in this module.



Managing and monitoring a data center network is a complex task, especially in large data centers. To successfully manage a data center, you need management software and solutions that are specific to the equipment. Monitoring of resources, link utilization, and network traffic types helps to achieve maximum stability of a data center network.

https://t.me/learningnets

certcollection.net

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollection.net

Module Self-Check

Use these questions to review what you learned in this module. The correct answers and solutions are found in the Module Self-Check Answer Key.

Q1) What are the two primary reasons to use network management tools? (Choose two.) (Source: Designing Data Center Management Solutions)

A) network selection
B) network monitoring
C) network configuration
D) network abstraction

Q2) What are two Cisco network management products? (Choose two.) (Source: Designing Data Center Management Solutions)

A) Cisco Virtual Switch Module
B) Cisco Network Analysis Module
C) Cisco DCNM for SAN
D) Cisco Application Networking Manager
E) Cisco Virtual Services Appliance
F) Cisco NetFlow Collector

Q3) Which protocol is used for traffic reporting? (Source: Designing Data Center Management Solutions)

A) SNMP
B) XML
C) NetFlow
D) FlowMask

Q4) What are the three priorities that the management software should provide when managing multitenant data centers? (Choose three.) (Source: Designing Data Center Management Solutions)

A) design templates
B) provisioning
C) billing
D) monitoring
E) user access control

https://t.me/learningnets

certcollection.net

Module Self-Check Answer Key

Q1) B, C

Q2) C, D

Q3) C

Q4) B, C, D

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets