DCUFI

Implementing Cisco Data Center Unified Fabric

Volume 2

Version 5.0

Student Guide

Text Part Number: 97-3212-01

https://t.me/learningnets

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

Table of Contents

Volume 2

Cisco Nexus Switch Advanced Feature Configuration 3-1

Overview 3-1
Module Objectives 3-1

Describing Cisco OTV 3-3

Overview 3-3
Objectives 3-3
Cisco OTV 3-4
Basic Cisco OTV Configuration 3-17
Advanced Cisco OTV Configuration 3-21
Summary 3-26

Configuring MPLS 3-27

Overview 3-27
Objectives 3-27
Multiprotocol Label Switching 3-28
Multiprotocol Label Switching VPNs 3-40
Summary 3-50

Configuring LISP 3-51

Overview 3-51
Objectives 3-51
Locator/ID Separation Protocol 3-52
Configuring LISP 3-60
Summary 3-64

Configuring QoS 3-65

Overview 3-65
Objectives 3-65
QoS on the Cisco Nexus Family of Switches 3-66
Modular QoS CLI 3-69
Marking 3-73
Mutation Mapping 3-83
Policing 3-86
Queuing and Scheduling 3-96
Monitoring QoS 3-107
Summary 3-108

Configuring Security Features 3-109

Overview 3-109
Objectives 3-109
DHCP Snooping 3-110
Dynamic ARP Inspection 3-113
IP Source Guard 3-115
Unicast Reverse Path Forwarding 3-117
Traffic Storm Control 3-120
Port Security 3-122
Control Plane Policing 3-124
Cisco TrustSec 3-128
Summary 3-132
Module Summary 3-133
References 3-134
Module Self-Check 3-135
Module Self-Check Answer Key 3-139

https://t.me/learningnets

Cisco Nexus Storage Features 4-1

Overview 4-1
Module Objectives 4-1

Describing Fibre Channel Protocol 4-3

Overview 4-3
Objectives 4-3
SCSI Protocol 4-4
Fibre Channel 4-6
Fibre Channel Login and Communication 4-10
Fibre Channel Addressing Schemes 4-16
VSANs and IVR 4-19
FSPF Protocol 4-30
Zoning 4-32
N-Port ID Virtualization 4-38
N-Port Virtualization 4-39
Summary 4-43

Describing FCoE Protocol 4-45

Overview 4-45
Objectives 4-45
FCoE Essentials 4-46
FCoE Architecture 4-51
FCoE Initialization Protocol 4-55
Supported FCoE Topologies 4-63
FCoE Hardware 4-72
Summary 4-75
References 4-75

Identifying DCB Enhancements 4-77

Overview 4-77
Objectives 4-77
Data Center Bridging 4-78
Priority Flow Control 4-81
Enhanced Transmission Selection 4-84
DCBX Protocol 4-85
Summary 4-88

Configuring FCoE 4-89

Overview 4-89
Objectives 4-89
FCoE Configuration 4-90
FCoE VLANs and Virtual Interfaces 4-96
FCoE with Enhanced vPC and Cisco Adapter FEX 4-106
Summary 4-112
References 4-112

Configuring SAN Switching Features 4-113

Overview 4-113
Objectives 4-113
FCoE on Cisco MDS 4-114
Fibre Channel Interfaces 4-116
Domain Parameters 4-121
VSAN Management 4-125
VSAN Trunking 4-130
SAN Port Channels 4-134
FLOGI and FCNS Databases 4-137
Cisco Fabric Services on Cisco MDS 4-139
Summary 4-147
References 4-147

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

Configuring NPV Mode **4-149**
Overview 4-149
Objectives 4-149
N-Port ID Virtualization 4-150
NPIV Configuration 4-152
NPV Mode 4-153
NPV Configuration 4-159
FCoE NPV Feature 4-162
FCoE NPV Configuration 4-169
Summary 4-174
References 4-174
Using SAN Management Tools **4-175**
Overview 4-175
Objectives 4-175
Cisco Prime DCNM for SAN 4-176
Cisco Device Manager 4-182
Summary 4-184
Module Summary 4-185
Module Self-Check 4-187
Module Self-Check Answer Key 4-194
Cisco Nexus Switch Management ***5-1***
Overview 5-1
Module Objectives 5-1
Using the CMP **5-3**
Overview 5-3
Objectives 5-3
Connectivity Management Processor 5-4
Configuring the CMP 5-6
Verifying the CMP 5-12
Upgrading the CMP 5-14
Using the CMP 5-16
Summary 5-18
Configuring User Management **5-19**
Overview 5-19
Objectives 5-19
User Management Features 5-20
Authentication, Authorization, and Accounting 5-21
Secure Shell 5-33
User Accounts and Roles 5-36
Summary 5-42
Describing System Management **5-43**
Overview 5-43
Objectives 5-43
System Management Features 5-45
Configuring Cisco Fabric Services 5-47
Configuring NTP and PTP 5-54
Configuring EEM 5-60
Configuring NetFlow 5-63
Configuring Smart Call Home 5-65
Configuring the Scheduler 5-70
Configuring SPAN and ERSPAN 5-73
Configuring SNMP 5-78
Using XML Interface 5-81
Implementing Cisco Prime DCNM for LAN 5-84
Integrating VMware vCenter Manager 5-91
Managing Licensing 5-96

https://t.me/learningnets

Upgrading the Firmware and EPLDs 5-99
Summary 5-104
Module Summary 5-105
References 5-106
Module Self-Check 5-107
Module Self-Check Answer Key 5-110

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

Module 3

Cisco Nexus Switch Advanced Feature Configuration

Overview

In this module, you will learn how to configure advanced features such as Cisco Overlay Transport Virtualization (OTV), Multiprotocol Label Switching (MPLS), Locator/ID Separation Protocol (LISP), security, and quality of service (QoS). You will learn how to configure Cisco OTV so that Layer 2 connectivity can be extended between sites, how to configure the MPLS feature so that the data center can take advantage of all the benefits that the appropriate technology provides, and be able to implement the LISP IP routing feature that enables the separation of the device identity and its location. Also, you will learn to choose and apply the relevant security features to support the application requirements and identify and configure QoS on the Cisco Nexus Family of switches.

Module Objectives

Upon completing this module, you will be able to identify and configure advanced features of the Cisco Nexus Family of switches. This ability includes being able to meet these objectives:

- Translate a given solution into an implementation plan for configuring both basic and advanced Cisco OTV features on the Cisco Nexus 7000 Series Switches
- Configure MPLS and MPLS VPN on the Cisco Nexus 7000 Series Switches
- Describe and configure LISP in the Cisco Data Center Network Architecture
- Configure QoS on the Cisco Nexus Family of switches based on a given requirement
- Evaluate the security features that are available on the Cisco Nexus Family of switches to identify which features should be implemented into a solution

https://t.me/learningnets

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

Lesson 1

Describing Cisco OTV

Overview

Cisco Overlay Transport Virtualization (OTV) is a “MAC-in-IP” technique that can be used to extend LANs over any transport. The transport can be Layer 2-based, Layer 3-based, IP-switched, label-switched, or anything else as long as it can carry IP packets. By using the principles of MAC routing, Cisco OTV provides an overlay that enables Layer 2 connectivity between separate Layer 2 domains. At the same time, it keeps these domains independent and preserves the fault-isolation, resiliency, and load-balancing benefits of an IP-based interconnection.

Objectives

Upon completing this lesson, you will be able to translate a given solution into an implementation plan for configuring both basic and advanced Cisco OTV features on the Cisco Nexus 7000 Series Switches. This ability includes being able to meet these objectives:

- Describe the terminology, benefits, and features of Cisco OTV in the Cisco Data Center Network Architecture
- Configure basic Cisco OTV features on the Cisco Nexus 7000 Series Switches
- Configure advanced Cisco OTV features on the Cisco Nexus 7000 Series Switches

https://t.me/learningnets

Cisco OTV

This topic describes the terminology, benefits, and features of Cisco OTV in the Cisco Data Center Network Architecture.

Distributed Data Centers

Goals when building the data center cloud include the following:

- Seamless workload mobility between multiple data centers
- Distributed applications closer to end users
- Pool and maximize global compute resources
- Ensure business continuity with workload mobility and distributed deployments

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

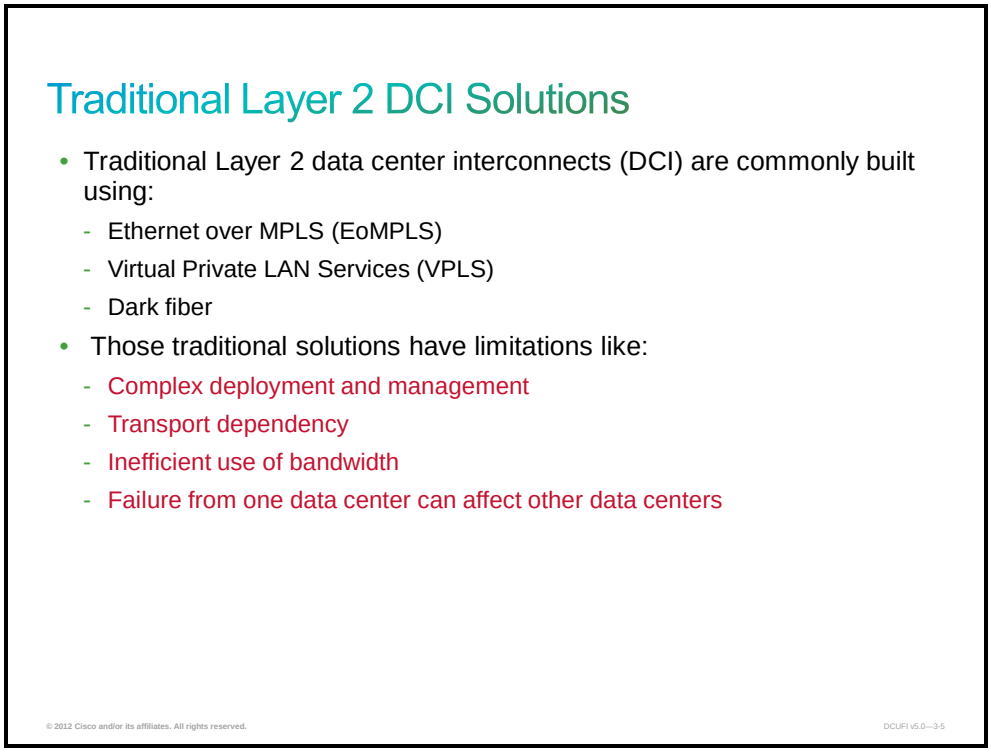
Businesses face the challenge of providing very high availability for applications while keeping operating expenses low. Applications must be available at any time and anywhere with optimal response times.

The deployment of geographically dispersed data centers allows IT architects to put in place effective disaster-avoidance and disaster-recovery mechanisms that increase the availability of the applications. Geographic dispersion also enables optimization of application response through improved facility placement and allows flexible mobility of workloads across data centers to avoid demand hotspots and fully utilize available capacity.

To enable all the benefits of geographically dispersed data centers, the network must extend Layer 2 connectivity across diverse locations. This connectivity must be provided without compromising the autonomy of data centers or the stability of the overall network.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets



The primary reasons to extend Layer 2 connectivity between data centers is that this connectivity is required to enable important technologies such as server clustering and workload mobility across data centers.

Traditionally, a number of different technologies have been used to provide Layer 2 data center interconnects (DCIs):

- **Ethernet over MPLS (EoMPLS):** EoMPLS can be used to provision point-to-point Layer 2 Ethernet connections between two sites using Multiprotocol Label Switching (MPLS) as the transport.
- **Virtual Private LAN Services (VPLS):** Similar to EoMPLS, VPLS uses an MPLS network as the underlying transport network. However, instead of point-to-point Ethernet pseudowires, VPLS delivers a virtual multiaccess Ethernet network.
- **Dark fiber:** In some cases, dark fiber may be available to build private optical connections between data centers using dense wavelength-division multiplexing (DWDM) or coarse wavelength-division multiplexing (CWDM).

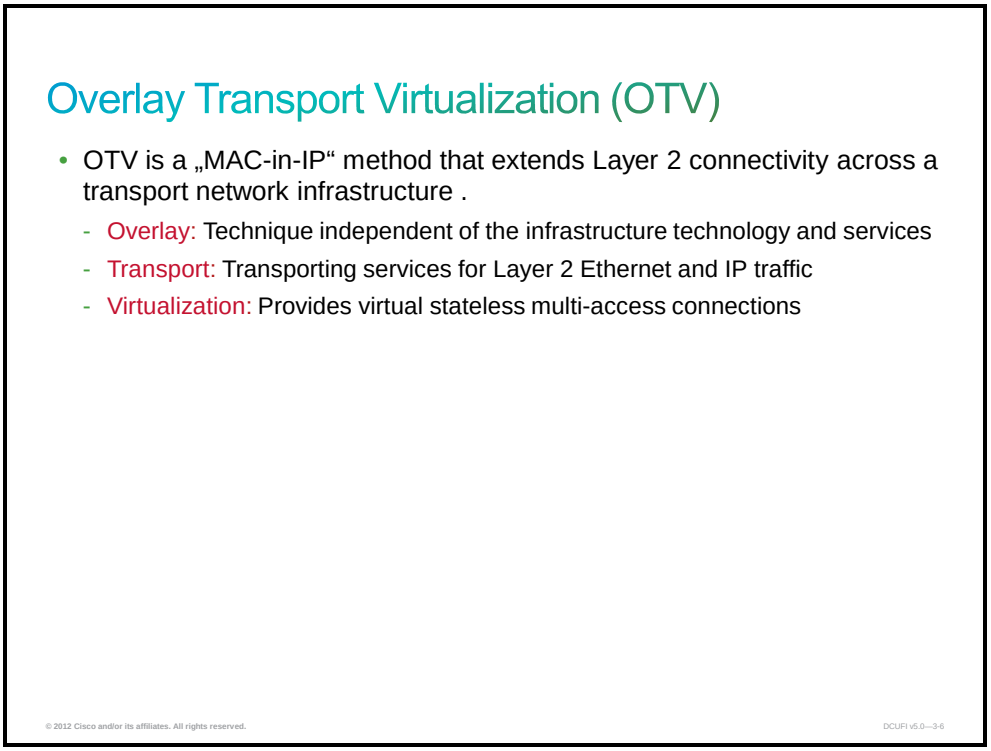
Although it is possible to build Layer 2 DCIs based on these technologies, they present a number of challenges:

- **Complex operations:** Traditional Layer 2 VPN technologies can provide extended Layer 2 connectivity across data centers. But these technologies usually involve a mix of complex protocols, distributed provisioning, and an operationally intensive hierarchical scaling model. The provisioning of many point-to-point connections or the complexity of the underlying transport technologies can add significantly to the operational cost of these types of DCIs. A simple overlay protocol with built-in capabilities and point-to-cloud provisioning is crucial to reducing the cost of providing this connectivity.
- **Transport dependent:** Traditional Layer 2 DCI technologies depend on the availability of a specific transport, such as dark fiber or an MPLS network. This dependency limits the options that are available to interconnect geographically dispersed data centers, because it requires the underlying transport to be available at all data center sites. A cost-effective

https://t.me/learningnets

solution for the interconnection of data centers must be transport agnostic and give the network architect the flexibility to choose any transport between data centers that is based on business and operational preferences.

- **Bandwidth management:** Traditional Layer 2 DCI technologies often do not allow the concurrent use of redundant connections because of the risk of Layer 2 loops. Balancing the load across all available paths while providing resilient connectivity between the data center and the transport network requires added intelligence above and beyond what is available in traditional Ethernet switching and Layer 2 VPN.
- **Failure containment:** The extension of Layer 2 domains across multiple data centers can cause the data centers to share failures that would normally be isolated when you interconnect data centers over an IP network. These failures propagate freely over the open Layer 2 flood domain. A solution that provides Layer 2 connectivity yet restricts the reach of the flood domain is needed to contain failures and preserve the resiliency that is achieved by using multiple data centers.



Cisco OTV overcomes the challenges that are inherent to traditional Layer 2 DCI technologies.

The name of the technology describes its important characteristics:

- **Overlay:** Cisco OTV provides an overlay VPN on top of an IP network. It is independent of the underlying infrastructure technologies and services. It does not impose any restrictions on the underlying infrastructure, as long as it is capable of transporting IP packets.
- **Transport:** Cisco OTV provides a Layer 2 transport across a Layer 3 network. It can take advantage of all the underlying capabilities of the underlying transport network, such as fast convergence, load balancing, and multicast replication.
- **Virtualization:** Cisco OTV provides a virtual multiaccess Layer 2 network that supports efficient transport of unicast, multicast, and broadcast traffic. Sites can be added to a Cisco OTV overlay without a need to provision additional point-to-point connections to the other sites. There are no virtual circuits, pseudowires, or other point-to-point connections to maintain between the sites. Packets are routed independently from site to site without a need to establish stateful connections between the sites.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

OTV Benefits over Traditional DCI

- The two benefits of OTV are:
  - Dynamic encapsulation
    - No pseudo-wire maintenance
    - Optimal multicast replication
    - Multipoint connectivity
    - Point-to-cloud model
  - Protocol learning
    - Preserve failure boundary
    - Built-in loop prevention
    - Automated multihoming
    - Site independence
- Not available at the Nexus 7000

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

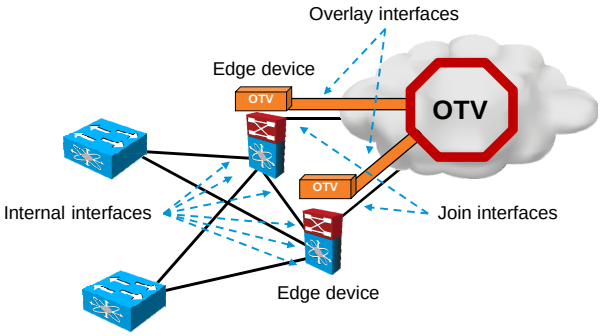
Cisco OTV is a “MAC-in-IP” encapsulation technique that rests on two technology pillars:

- **Dynamic encapsulation:** Cisco OTV uses a point-to-cloud model, which encapsulates Layer 2 frames in IP packets that can be independently routed to the destination site or sites. There is no need to provision point-to-point Layer 2 circuits between the sites. If the underlying IP network is multicast-enabled, then multicast frames can be encapsulated in IP multicast packets. The frames can be replicated and routed to the appropriate sites by the multicast IP transport network.
- **Protocol learning:** Cisco OTV uses MAC address routing that is based on the Intermediate System-to-Intermediate System (IS-IS) protocol to dynamically learn which MAC addresses are available in each of the sites. The use of a routing protocol eliminates the need to depend on Layer 2 flooding for MAC address learning. By eliminating unnecessary flooding, failure domains are bounded. Also, there are loop prevention mechanisms that are built into the routing protocol, which eliminates the need to extend Spanning Tree Protocol (STP) to break forwarding loops. Sites remain independent spanning-tree domains. Finally, the Cisco OTV neighbor discovery mechanism provides for automated multihoming of sites to the overlay network.

https://t.me/learningnets

OTV Components

- Edge device: encapsulation/decapsulation between L2 and OTV and all OTV functions
- Internal interfaces: connects to the VLANs that are to be extended
- Join interface: joins an overlay network
- Overlay interface: encapsulates the Layer 2 frames in IP packets



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

To understand the operation of Cisco OTV, it is important to establish some of the key terms.

Cisco OTV is an edge function. Layer 2 traffic is received from the switched network. For all VLANs that need to be extended to remote locations, the Ethernet frames are dynamically encapsulated into IP packets that are then sent across the transport infrastructure. A device that performs the Cisco OTV encapsulation and de-encapsulation functions between the Layer 2 network and the transport network is a *Cisco OTV edge device*.

The Cisco OTV edge device can either be in the core or aggregation layer of the data center. A site can have multiple edge devices to provide additional resiliency. This scenario is commonly referred to as multihoming.

By definition, a Cisco OTV edge device has interfaces that connect to the transport network and interfaces that connect to the Layer 2 switched network. The Layer 2 interfaces that receive the traffic from the VLANs that are to be extended across Cisco OTV are named the internal interfaces. These interfaces are regular Layer 2 interfaces, usually IEEE 802.1Q trunks. No Cisco OTV-specific configuration is required on the *internal interfaces*.

The *join interface* is used to source the Cisco OTV encapsulated traffic and send it to the Layer 3 domain of the data center network. The join interface is a Layer 3 entity. With the current release of the Cisco Nexus Operating System (NX-OS) Software, the join interface can only be defined as a routed point-to-point physical or logical interface.

A Cisco OTV overlay can only have a single join interface per edge device. Multiple overlays can share the same join interface.

The *overlay interface* is a logical multiaccess and multicast-capable interface that must be explicitly defined by the user. The entire Cisco OTV overlay configuration is applied on this logical interface. Every time the Cisco OTV edge device receives a Layer 2 frame that is destined for a remote data center site, the frame is logically forwarded to the overlay interface. This process causes the edge device to perform the dynamic Cisco OTV encapsulation on the Layer 2 frame and send it to the join interface toward the routed domain.

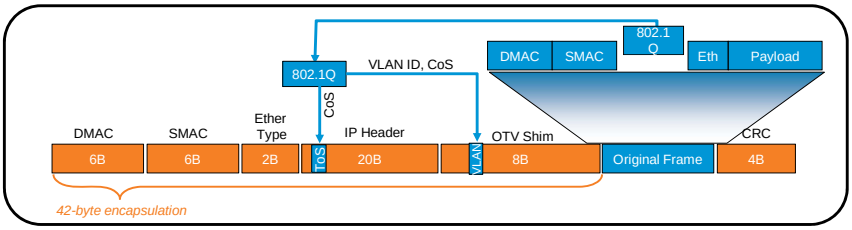
https://t.me/learningnets

OTV Packet Flow

- When the Layer 2 frame arrives at edge device, one of this two options is available:
  1. Frame is destined to a MAC address that is reachable through another internal interface

Forwarding

  2. Frame is destined to a MAC address that was learned over an overlay interface



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

When the Layer 2 frame arrives at edge device on an internal interface, Cisco OTV performs the MAC table lookup based on the destination address of the Layer 2 frame. If the frame is destined to a MAC address that is reachable through another internal interface, the frame is forwarded out on that internal interface. In this case, Cisco OTV performs no other actions and the processing of the frame is complete.

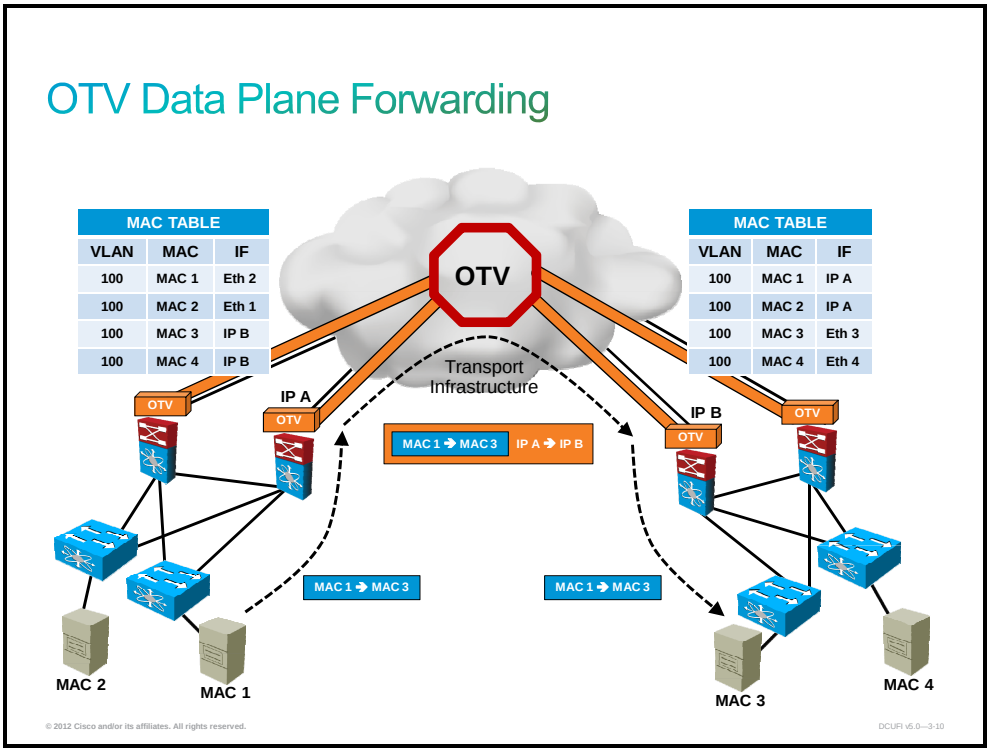
If the Layer 2 frame is destined to a MAC address that was learned over an overlay interface, Cisco OTV performs the following tasks:

Step 1 The cyclic redundancy check (CRC) and the 802.1Q fields are removed from the original Layer 2 frame.

Step 2 A Cisco OTV shim header is added, which includes the VLAN and class of service (CoS) values from the 802.1Q header and overlay ID information.

Step 3 An external IP header and outer Ethernet header are also added. The CoS value from the 802.1Q header is also mapped into the precedence bits of the type of service (ToS) byte in the IP header.

https://t.me/learningnets



After control plane adjacencies have been established between the Cisco OTV edge devices in different sites and MAC address reachability information has been exchanged, the MAC address tables in the Cisco OTV edge devices contain two different types of entries. In addition to the normal MAC address entries that point to a Layer 2 switch port, the MAC address table now also contains entries that point to the IP addresses of adjacent Cisco OTV neighbors.

To forward frames that are based on the MAC entries that are installed by Cisco OTV, the following procedure, which is illustrated in the figure, is followed:

Step 1 The Layer 2 frame is received at the Cisco OTV edge device. A traditional Layer 2 lookup is performed. However, this time, the information in the MAC address table for MAC address MAC 3 does not point to a local Ethernet interface. Rather, it points to the IP address of the remote Cisco OTV edge device that advertised the MAC reachability information for MAC 3.

Step 2 The Cisco OTV edge device encapsulates the original Layer 2 frame. The source IP address of the outer header is the IP address of its join interface, whereas the destination IP address is the IP address of the join interface of the remote edge device.

Step 3 The Cisco OTV encapsulated frame, which is a regular unicast IP packet, is carried across the transport infrastructure and delivered to the remote Cisco OTV edge device.

Step 4 The remote Cisco OTV edge device de-encapsulates the IP packet exposing the original Layer 2 frame.

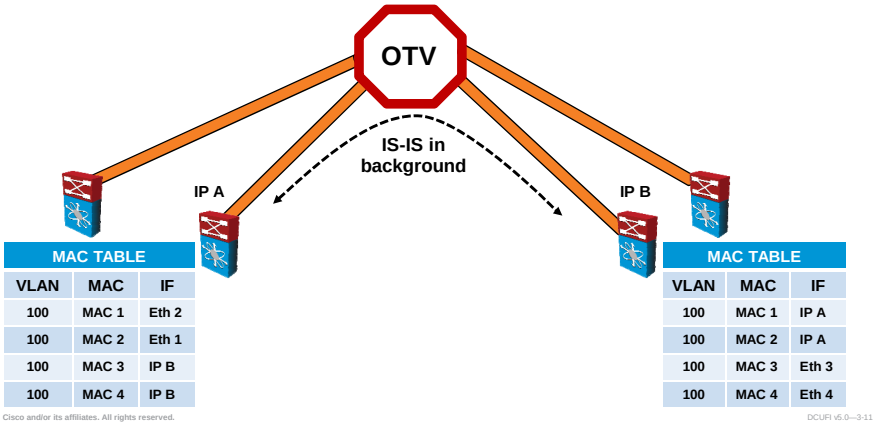
Step 5 The edge device performs a Layer 2 lookup on the original Ethernet frame. It discovers that it is reachable through a physical interface, which means it is a MAC address local to the site.

Step 6 The frame is delivered to the destination host that MAC 3 belongs to through regular Layer 2 Ethernet switching.

© 2012 Cisco Systems, Inc.
https://t.me/learningnets

Building the MAC Address Table

- OTV periodically advertises MAC address reachability
- MAC addresses are advertised in background and no specific configuration is required
- IS-IS runs in background – no need to configure or understand it



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

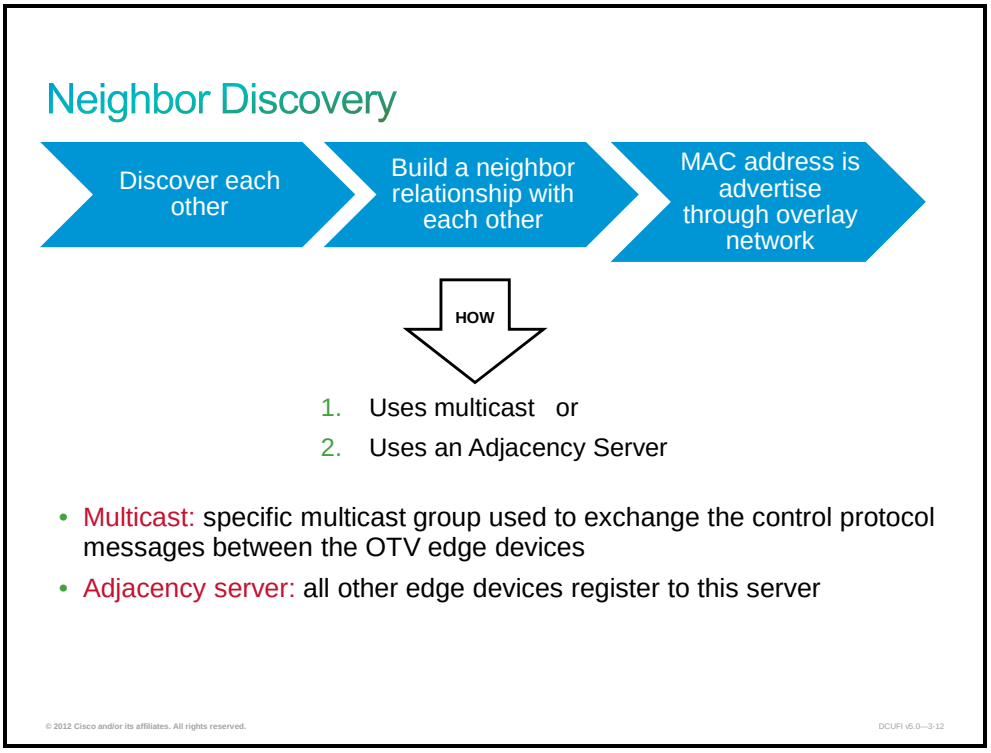
Before any encapsulation and de-encapsulation of frames across the overlay can be performed, it is necessary to exchange MAC address reachability information between the sites to create corresponding Cisco OTV MAC address table entries.

Cisco OTV does not depend on flooding to propagate MAC address reachability information. Instead, Cisco OTV uses a control plane protocol to distribute MAC address reachability information to remote Cisco OTV edge devices. This protocol runs as an overlay control plane between Cisco OTV edge devices. This means that there is no dependency with the routing protocol used in the Layer 3 domain of the data center or in the transport infrastructure.

The Cisco OTV control plane is transparently enabled in the background after creating the Cisco OTV overlay interface and does not require explicit configuration. Tuning parameters, like timers, for the Cisco OTV protocol is possible, but this is expected to be more of a corner case than a common requirement.

Note The routing protocol that is used to implement the Cisco OTV control plane is IS-IS. It was selected because it is a standards-based protocol, originally designed with the capability of carrying MAC address information in type, length, value (TLV) triplets. Despite the fact that IS-IS is used, the control plane protocol will be generically called “OTV protocol” to differentiate it from IS-IS being used as an IGP for IP version 4 (IPv4) or IP version 6 (IPv6). It is not necessary to have a working knowledge of IS-IS configuration to implement Cisco OTV. However, some background in IS-IS can be helpful when you are troubleshooting Cisco OTV.

https://t.me/learningnets



To advertise MAC address reachability, the Cisco OTV edge devices need to discover all the other edge devices on the overlay, and establish adjacencies with each other.

The neighbor discovery and adjacency creation process can be achieved in two ways, depending on the nature of the transport network interconnecting the various sites:

- If the transport is multicast-enabled, a specific multicast group can be used to exchange the control protocol messages between the Cisco OTV edge devices.
- If the transport is not multicast-enabled, an alternative deployment model, which will be available in a future Cisco NX-OS Software release, can be used. In this model, a Cisco OTV edge device is configured as an adjacency server to which all other edge devices register. The adjacency server communicates the list of devices belonging to a given overlay to all other Cisco OTV edge devices.

To take advantage of all Cisco OTV benefits, the use of a multicast enabled transport is recommended when you are extending VLANs between more than two data centers. Not only does this simplify the neighbor discovery and adjacency building process, but it also allows for more efficient forwarding of multicast traffic between the sites.

© 2012 Cisco Systems, Inc.
https://t.me/learningnets

Neighbor Discovery (Cont.)

- When the transport network is multicast enabled, OTV uses a multicast group for neighbor discovery
  - Edge devices join the multicast group using IGMP
  - PIM configuration is not required on edge devices
  - OTV hellos and updates are encapsulated in the multicast group
  - Core multicast replication is used to send updates to all OTV neighbors
- To be able to unicast replicate, each OTV node must know a list of neighbors to replicate to.
  - With NX-OS 5.2(1) software release, OTV can be deployed with unicast-only transport
  - Choose which OTV edge will be the Adjacency Server and configure it
  - OTV edge devices must first to “register” to Adjacency Server
  - When an OTV edge device is configured with adjacency server address they are added to the unicast-replication-list (URL).

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

Assuming that the transport network is multicast-enabled, all Cisco OTV edge devices can be configured to join a specific Any Source Multicast (ASM) group on which they simultaneously act as a receiver and source. The specific ASM multicast group that is to be used for Cisco OTV control is not prescribed and any unused ASM group can be used. If the transport is not under your own control, but provided by a service provider, you will have to coordinate the use of this ASM group with the provider.

The Cisco OTV edge devices do not act as multicast routers, but as multicast endpoints. Protocol Independent Multicast (PIM) does not need to be enabled on the Cisco OTV join interface.

Edge devices use Internet Group Management Protocol (IGMP) to join the ASM control group on the join interface to be able to receive the Cisco OTV control protocol packets from the remote edge devices. Edge devices encapsulate their Cisco OTV protocol hellos and updates through the dynamic Cisco OTV encapsulation process, using the ASM control group as the destination IP address. The multicast routing protocols that are deployed in the transport network replicate the packets and deliver them to all the remote edge devices.

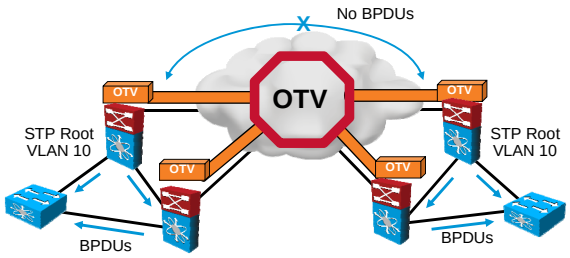
With Cisco NX-OS Software Release 5.2(1), Cisco OTV can be deployed with unicast-only transport. To communicate with all the remote Cisco OTV devices, each Cisco OTV node needs to know a list of neighbors to replicate the control packets to. Rather than statically configuring in each Cisco OTV node, the list of all neighbors is maintained on one (or more) Cisco OTV edge device to perform a specific role that is named Adjacency Server. Every Cisco OTV device that has to join a specific Cisco OTV logical overlay needs first to register itself with the adjacency server by sending Cisco OTV hello messages to it. All other Cisco OTV neighbor addresses are discovered dynamically through the adjacency server.

The reception of the hello messages from all the Cisco OTV edge devices helps the adjacency server to build up the list of all the Cisco OTV devices that should be part of the same overlay (named unicast-replication-list). This list is periodically sent in unicast fashion to all the listed Cisco OTV devices, so that they can dynamically be aware of all the Cisco OTV neighbors that are in the network.

https://t.me/learningnets

Spanning Tree and OTV

- OTV is site transparent for STP:
  - Each site maintains its own STP topology
  - An OTV edge device only sends and receives BPDUs on internal interfaces
  - This mechanism is built into OTV and requires no additional configuration



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

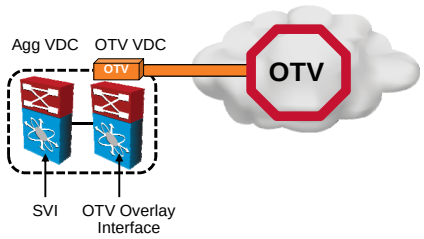
Cisco OTV by default does not transmit STP bridge protocol data units (BPDUs) across the overlay. This function is a native Cisco OTV function that does not require the use of any explicit configuration, such as BPDU filtering. This method allows every site to remain an independent spanning-tree domain: Spanning-tree root configuration, parameters, and the spanning-tree protocol flavor can be decided on a per-site basis.

The separation of spanning-tree domains fundamentally limits the fate sharing between data center sites. A spanning-tree problem in the control plane of a given site would not produce any effect on the remote data centers.

https://t.me/learningnets

OTV and SVI

- OTV and a Layer 3 VLAN interface cannot be configured for the same VLAN
- Possible solutions:
  - Configure OTV in a separate VDC (recommended)
  - Move the SVI to another device (increases complexity)



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

A VLAN can either have Layer 3 connectivity through a VLAN interface (switch virtual interface [SVI]) or the VLAN can be extended over Cisco OTV. If you have a design scenario that requires the VLAN to be extended over Cisco OTV to a remote site and have Layer 3 connectivity through a VLAN interface, you must configure Cisco OTV in a separate virtual device context (VDC) from the VDC that contains the VLAN interfaces.

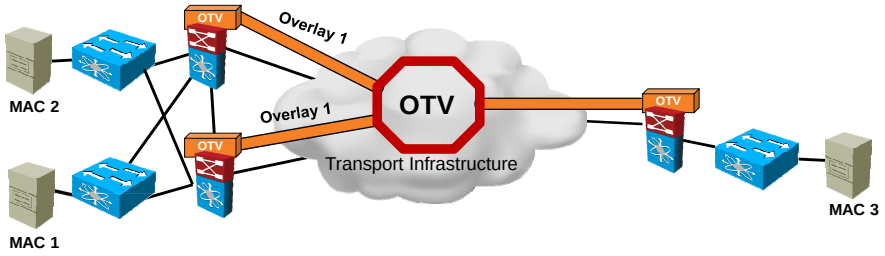
The Agg VDC contains all the configuration and physical links for the aggregation layer of a data center. The Agg VDC also includes the VLAN interfaces (SVIs) for any VLANs that need Layer 3 connectivity. The Agg VDC is connected to the Cisco OTV VDC through a loopback cable over a trunk port. This trunk port carries any VLAN traffic that needs to be extended over the overlay network.

The Cisco OTV VDC also includes a trunk port that accepts this VLAN traffic. All Cisco OTV configuration exists in the Cisco OTV VDC. The overlay interface has an extended VLAN range that includes VLANs from the Agg VDC that have Layer 3 connectivity through VLAN interfaces. These extended VLANs are isolated in a separate VDC from the VLAN interfaces in the Agg VDC. The Agg VDC decides on whether a Layer 2 frame is forwarded to the local VLAN interface to Layer 3 or whether the Layer 2 frame is sent over the trunk port to the Cisco OTV VDC and encapsulated for the overlay network.

https://t.me/learningnets

Multihomed Sites and Load Balancing

- Multihome site – more then one edge device is present in overlay network
- The site VLAN is a local VLAN and should not be extended across the overlay
- OTV elects authoritative edge device (AED) to be forwarder for a subset of the extended VLANs
- The AED is responsible for advertising the MAC addresses and forwarding traffic to and from the overlay for its set of VLANs
- Load-balancing is achieved through designation of a subset of all VLANs for each edge device



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

One critical function that is built into the Cisco OTV protocol is multihoming, where two or more Cisco OTV edge devices provide LAN extension services to a given site. This redundant node deployment, combined with the fact that STP BPDUs are not sent across the Cisco OTV overlay, could potentially lead to the creation of a bridging loop between sites. To prevent this loop, Cisco OTV has a built-in mechanism to help ensure that only one of the edge devices forwards traffic for a given VLAN. The edge device that has the active forwarding role for the VLAN is called an authoritative edge device (AED) for that VLAN.

The AED has two main tasks:

- Forwarding Layer 2 unicast, multicast, and broadcast traffic between the site and the overlay and vice versa
- Advertising MAC address reachability information to the remote edge devices

The AED role is negotiated, on a per-VLAN basis, between all the Cisco OTV edge devices belonging to the same site. To decide which device should be elected as an AED for a given site, the Cisco OTV edge devices establish an internal Cisco OTV control protocol peering.

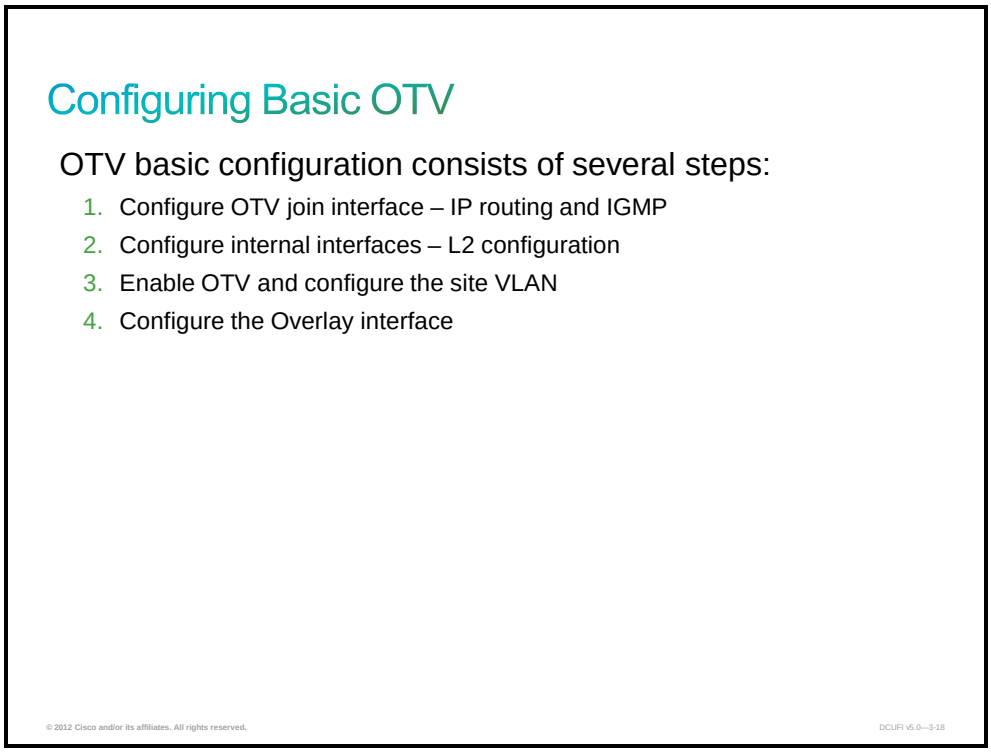
The internal adjacency is established on a dedicated VLAN, named the site VLAN. The site VLAN should be carried on multiple Layer 2 paths internal to a given site, to increase the resiliency of this internal adjacency. The internal adjacency is used to negotiate the AED role.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Basic Cisco OTV Configuration

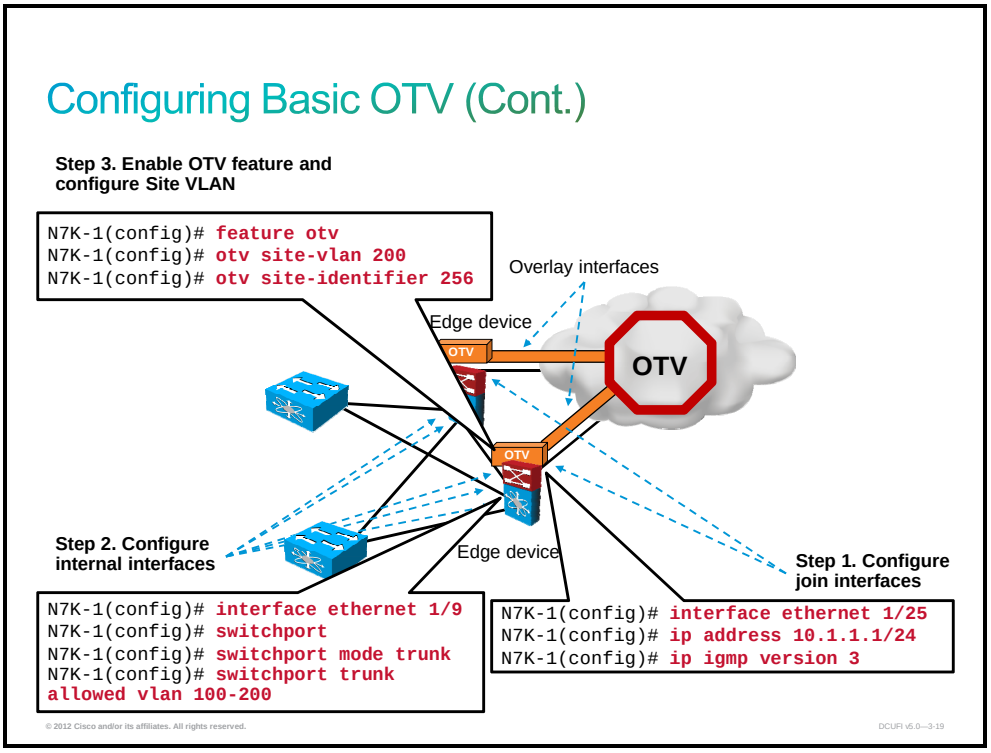
This topic explains how to configure basic Cisco OTV features on the Cisco Nexus 7000 Series Switches.



Configuring Cisco OTV consists of a series of simple tasks:

- The Cisco OTV join interface needs to be configured. This includes configuring IP routing for the transport network and configuring IGMP. There is no Cisco OTV-specific configuration on the join interface.
- The internal interfaces need to be configured. This task consists of straightforward Layer 2 configuration. There is no Cisco OTV-specific configuration on the internal interfaces.
- Cisco OTV needs to be enabled and the Cisco OTV site VLAN needs to be configured.
- The overlay interface that connects the edge devices to the overlay network needs to be configured. This is where most of the Cisco OTV-specific configuration is performed.

https://t.me/learningnets



The Cisco OTV join interface is a routed interface that connects to the Layer 3 transport network. In the current Cisco NX-OS Software release, this link has to be a point-to-point link. It can be a routed port, a routed port channel, or a subinterface of a routed port or port-channel. The join interface cannot be a loopback or switched virtual interface (SVI).

Internet Group Management Protocol version 3 (IGMPv3) needs to be enabled on the join interface because Source Specific Multicast (SSM) groups in the transport network are used to encapsulate the multicast traffic that needs to be transported across the overlay. There is no need to enable the PIM multicast routing protocol on the interface. The Cisco OTV edge device acts as a multicast endpoint, not as a multicast router.

For the IP routing in the transport network, any IP routing protocol can be used. Before configuring Cisco OTV, verify that you have complete IP connectivity between all Cisco OTV edge devices.

The internal interfaces are regular Layer 2 interfaces. There is no Cisco OTV-specific configuration on these interfaces. One point to consider is that the Cisco OTV site VLAN needs to be extended between the Cisco OTV peers within the site, but not across the overlay. Therefore, you should ensure that the Cisco OTV site VLAN is allowed on the internal interfaces. The example in the figure shows how to configure Cisco OTV internal interfaces. The allowed VLAN list should include both the site VLAN and the list of VLANs that are to be extended across the overlay.

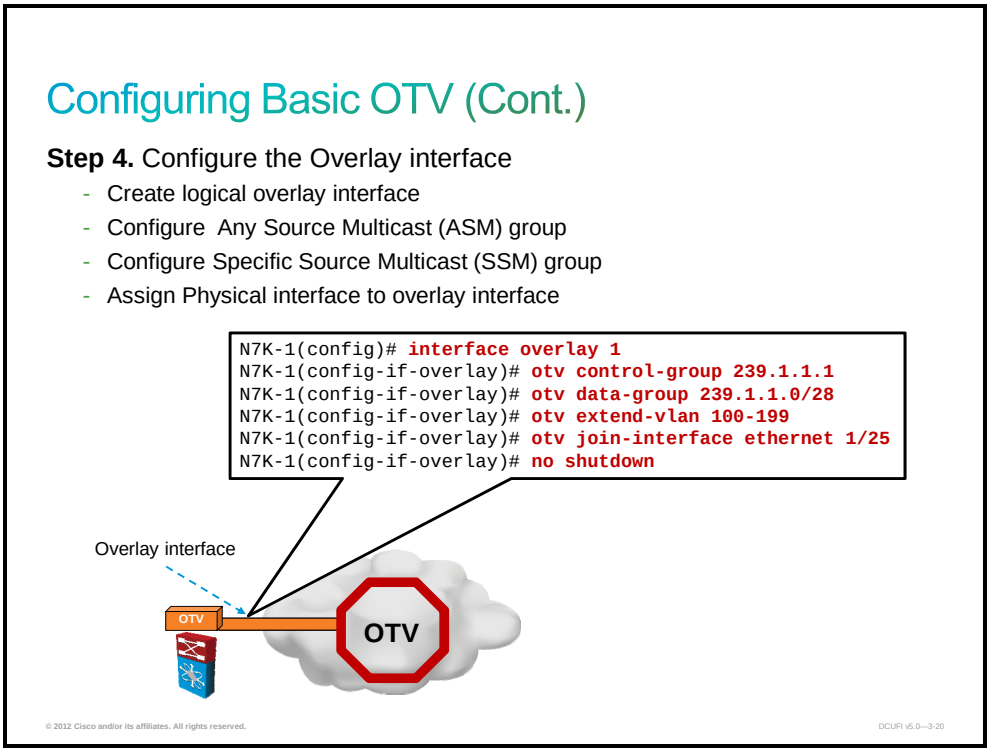
The Cisco OTV feature requires the Transport Services license to be installed on the Cisco Nexus 7000 Series switch. Next, the site VLAN needs to be configured. The default site VLAN is VLAN 1. It is important that the site VLAN is carried on the internal interfaces between the Cisco OTV edge devices within the site. The site VLAN should not be extended across the Cisco OTV overlay. The site VLAN is used by the different Cisco OTV edge devices within the site to communicate and should be configured consistently between the edge devices. Dual site adjacency uses both site VLAN and site identifier to determine if there are other edge devices on the local site and if those edge devices can forward traffic.

© 2012 Cisco Systems, Inc.
https://t.me/learningnets

The following commands are used for configuring the Cisco OTV feature:

- **feature otv:** This command enables the Cisco OTV feature. This feature requires the Transport Services license.
- **otv site-vlan** *vlan-id***:** This command defines the Cisco OTV site VLAN that is used between Cisco OTV edge devices in the same site. The default site VLAN is VLAN 1.
- **otv site-identifier** *id***:** This command configures the site identifier. The same site identifier should be configured on all local Cisco OTV edge devices. The site identifier should be unique across different sites.

https://t.me/learningnets



Next, the Cisco OTV overlay interface must be created and the following parameters must be assigned to it:

- **Join interface:** This interface is used to join the overlay. The join interface must connect to the transport network. There can only be one join interface per overlay. For added resiliency, a port-channel interface can be used as the join interface.
- **Multicast control group:** This ASM multicast group is used to communicate with remote Cisco OTV edge devices. It should be the same on all edge devices that are participating in the overlay.
- **Multicast data groups:** These SSM multicast groups are used to encapsulate multicast data from the site and transport it across the overlay to the remote edge devices.
- **Extended VLANs:** These VLANs are the VLANs that are to be extended across the overlay.

The example in the figure shows how to configure a Cisco OTV overlay interface. The following commands are used:

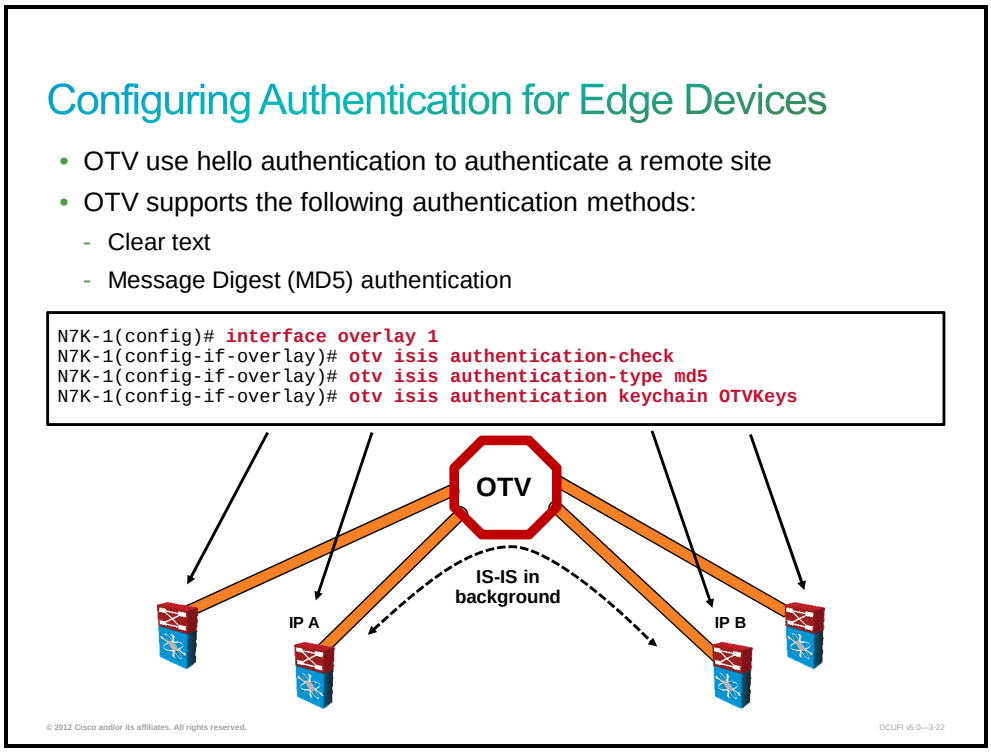
- **interface overlay** *number***:** This command creates a Cisco OTV overlay interface.
- **otv join-interface** {**ethernet** *slot/port* | **port-channel** *channel-number*}**:** This command associates the Cisco OTV join interface with the overlay interface. The Cisco OTV overlay interface will not be activated until you configure a join interface.
- **otv control-group** *multicast-addr***:** This command defines the ASM multicast group that is used for Cisco OTV control plane communication on the overlay.
- **otv data-group** *multicast-addr/mask***:** This command defines a range of SSM multicast groups that will be used to carry multicast data traffic between the sites.
- **otv extend-vlan** *vlan-range***:** This command specifies the VLAN or range of VLANs that are to be extended across the overlay network. This range should not include the site VLAN.
- **no shutdown**: This command enables the overlay interface.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Advanced Cisco OTV Configuration

This topic explains how to configure advanced Cisco OTV features on the Cisco Nexus7000 Switches.



Authentication for the Cisco OTV control-plane protocol hello messages can be configured. Cisco OTV use hello authentication to authenticate a remote site before Cisco OTV creates an adjacency to that remote site. Each overlay network uses a unique authentication key. An edge device only creates an adjacency with a remote site that shares the same authentication key and authentication method.

Cisco OTV supports the following authentication methods:

- Cleartext
- Message Digest 5 (MD5) authentication

The following commands are used:

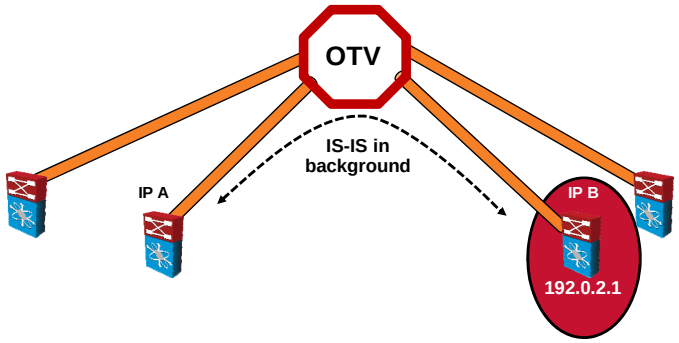
- **interface overlay** *number***:** This command creates a Cisco OTV overlay interface.
- **otv isis authentication-check:** This command enables authentication of hello messages between Cisco OTV edge devices. The default is enabled.
- **otv isis authentication-type** {**cleartext** | **md5**}**:** This command is used to configure the authentication method.
- **otv isis authentication keychain** *keychain-name***:** This command is used to configure the authentication keychain for edge device authentication. The keychain name can be any case-sensitive alphanumeric string up to 16 characters.

https://t.me/learningnets

Configuring OTV Adjacency Servers

- Local edge device or remote edge device can be configured to act as an adjacency server

```
N7K-1(config)# interface overlay 1
N7K-1(config-if-overlay)# otv adjacency-server unicast-only
N7K-1(config-if-overlay)# otv use-adjacency-server 192.0.2.1 unicast-only
```



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

Local edge device or remote device can act as an adjacency server. The example in the figure shows how to configure a Cisco OTV adjacency server.

The following commands are used:

- **interface overlay** *number***:** This command creates a Cisco OTV overlay interface.
- **otv adjacency-server unicast-only:** (Optional) This command is used to configure the local edge device to act as an adjacency server.
- **otv use-adjacency-server** *primary-ip-address* [*secondary-ip-address*] **unicast-only**: This command is used to configure the local edge device to use a remote adjacency server. The IP address format is in dotted decimal notation. The *secondary-ip-address* argument is the IP address of the backup adjacency server, if you have configured a backup adjacency server.

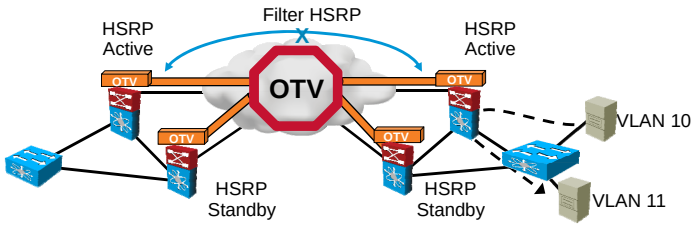
© 2012 Cisco Systems, Inc.

https://t.me/learningnets

FHRP Filtering and VLAN Extension

- Extended VLANs commonly use a first-hop redundancy protocol (FHRP), such as HSRP or VRRP
- Only one router is active for an FHRP group
- All traffic is forwarded to the virtual IP and MAC of the active router
- Result: Suboptimal routing

- The FHRP suboptimal routing problem is solved using the FHRP traffic flitering
- One active FHRP router per site
  - A VLAN access list filters the FHRP control packets
  - OTV MAC route filter stops the announcement of the FHRP virtual MAC address



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

Data center networks commonly use a First Hop Redundancy Protocol (FHRP), such as the Hot Standby Router Protocol (HSRP), Virtual Router Redundancy Protocol (VRRP), or Gateway Load Balancing Protocol (GLBP). The combination of Cisco OTV and FHRPs could lead to suboptimal routing for VLANs that are extended across the overlay. When a VLAN is extended across multiple sites, only a single router will be elected as the active gateway and all traffic from that VLAN will go through that router.

As can be seen in the figure, this could cause traffic between two hosts that are in the same site, but on different VLANs, to be forwarded across the overlay to the other data center to be routed and then forwarded back across the overlay to the destination host. Clearly, this behavior is undesirable and should be prevented.

The desired behavior for the extended VLANs is to elect an FHRP active router per site, to help ensure that traffic is routed locally when possible. To achieve the desired behavior, the FHRP traffic should be filtered from the overlay. If the routers in the different sites do not see the hellos of each other, they will elect an active and standby router in each site. The active router in the site will be responsible for forwarding the traffic for the virtual IP and MAC address in that site.

To filter the FHRP traffic, two steps need to be taken. A VLAN access control list (VACL) must be implemented on the Cisco OTV edge device to block the FHRP hellos and prevent them from being forwarded on the overlay. Second, the virtual FHRP MAC address should not be announced to other sites, as it normally would. This result can be achieved by implementing a Cisco OTV MAC route filter.

https://t.me/learningnets

HSRP Filtering Example

- This example shows the configuration of a VLAN access list that filters HSRP traffic:

```
N7K-1(config)# ip access-list HSRP
N7K-1(config-acl)# permit udp any 224.0.0.2/32 eq 1985

N7K-1(config)# ip access-list ANY-IP
N7K-1(config-acl)# permit ip any any

N7K-1(config)# vlan access-map FILTER-HSRP 10
N7K-1(config-access-map)# match ip address HSRP
N7K-1(config-access-map)# action drop

N7K-1(config)# vlan access-map FILTER-HSRP 20
N7K-1(config-access-map)# match ip address ANY-IP
N7K-1(config-access-map)# action forward

N7K-1-pod5(config)# vlan filter FILTER-HSRP vlan-list 100-199
```

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

The example in the figure shows how to filter HSRP packets on a Cisco OTV edge device. A VACL is implemented that matches HSRP traffic by its destination multicast address and UDP port number. The VACL drops these packets, while it forwards any other IP traffic. The VACL is applied to the VLANs 100–199, which represent the range of VLANs that are extended across the overlay.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

HSRP Filtering Example (Cont.)

- This example shows the configuration of a OTV route filter that filters the announcement of the HSRP virtual MAC address:

```
N7K-1(config)# mac-list NOT-HSRP-VMAC deny 0000.0c07.ac00 0000.0000.00ff
N7K-1(config)# mac-list NOT-HSRP-VMAC permit 0000.0000.0000 ffff.ffff.ffff

N7K-1(config)# route-map NO-HSRP-ANNOUNCE permit 10
N7K-1(config-route-map)# match mac-list NOT-HSRP-VMAC

N7K-1(config)# otv-isis default
N7K-1(config-router)# vpn Overlay1
N7K-1(config-router-vrf)# redistribute filter route-map NO-HSRP-ANNOUNCE
```

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

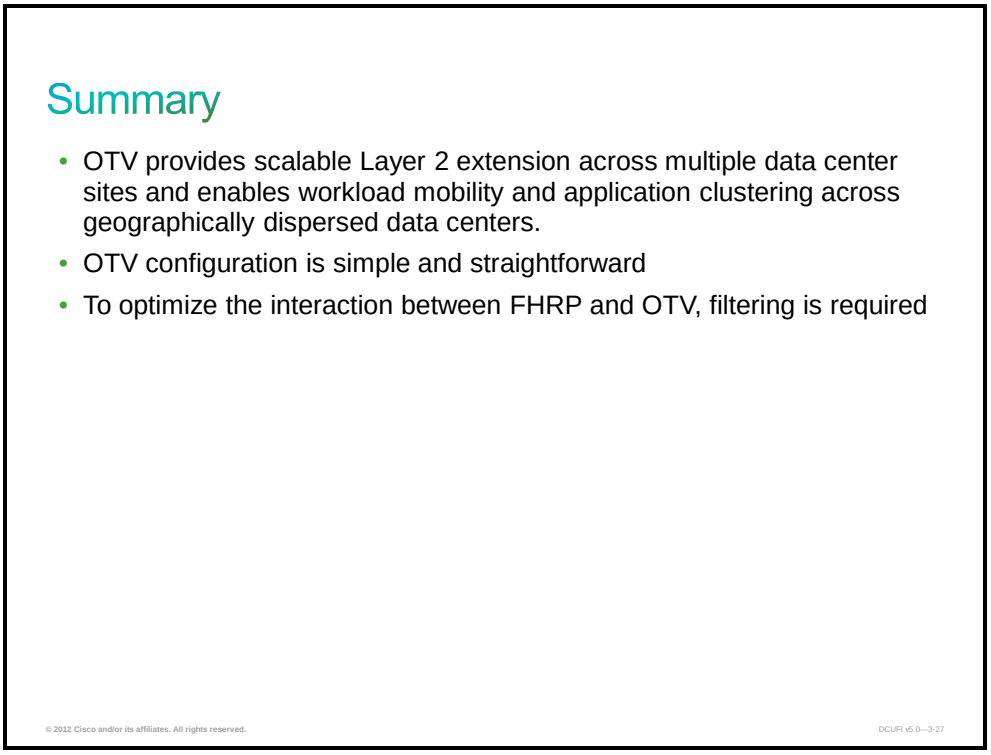
The example in the figure shows how to configure an OTV route filter to filter the announcement of the HSRP virtual MAC addresses. First, a MAC address access list is created that permits all MAC addresses except the HSRP virtual MAC address block. Next, this access list is tied to a route map because all route filtering in the Cisco NX-OS Software is performed through route maps. This route map filters out the HSRP MAC addresses that are defined in the access list. Finally, the route map is applied to the Cisco OTV IS-IS control protocol for the overlay.

The end result of the configuration is that HSRP remains strictly separated per site and each site has its own active HSRP router, which performs the inter-VLAN routing function.

https://t.me/learningnets

Summary

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



© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Lesson 2

Configuring MPLS

Overview

Multiprotocol Label Switching (MPLS) is a switching mechanism that is often found in service provider environments. MPLS uses traditional IP routing and supports several services that are required in next-generation IP networks.

The lesson discusses the basic concepts and terminology regarding MPLS switching, and describes the functions of MPLS labels and forwarding structures that are used by MPLS routers. You will learn about business needs that next-generation data centers must support, and what benefits MPLS technology provides in the data center environment. The lesson also describes how to configure MPLS on Cisco Nexus 7000 Series Switches and how to verify MPLS operations on this platform.

Objectives

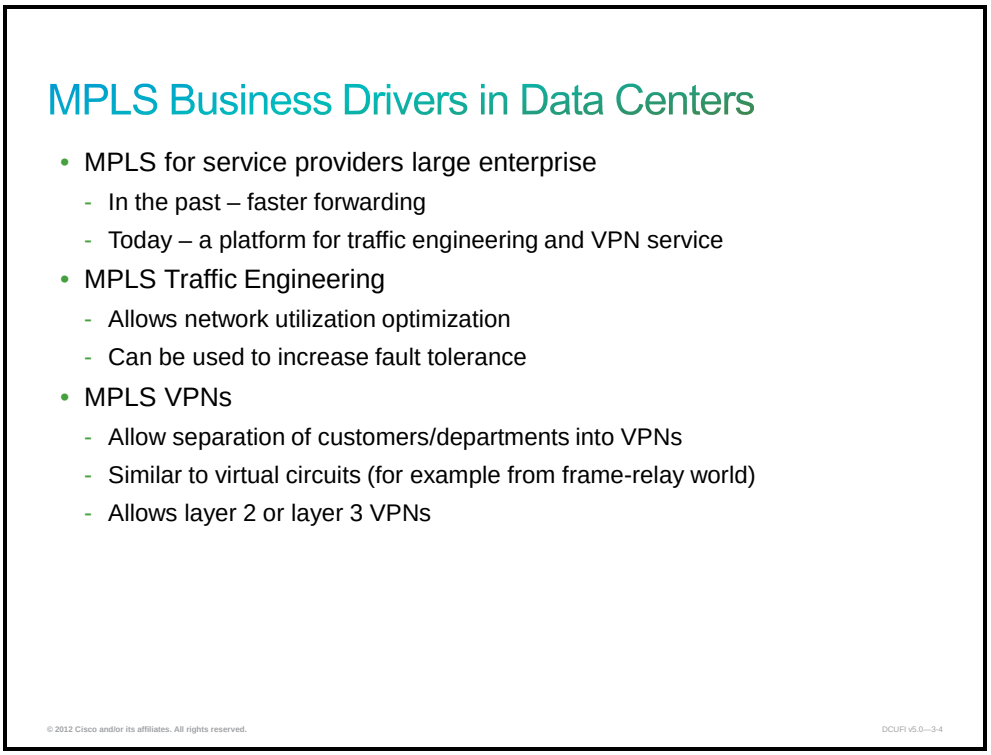
Upon completing this lesson, you will be able to configure both MPLS and MPLS VPN on the Cisco Nexus 7000 Series Switches. This ability includes being able to meet these objectives:

- Describe the terminology, operation, and configuration of MPLS in the Cisco Data Center Network Architecture
- Describe the benefits, operation, and configuration of MPLS VPNs in the Cisco Data Center Network Architecture

https://t.me/learningnets

Multiprotocol Label Switching

This topic describes the terminology, operation, and configuration of MPLS in the Cisco Data Center Network Architecture.



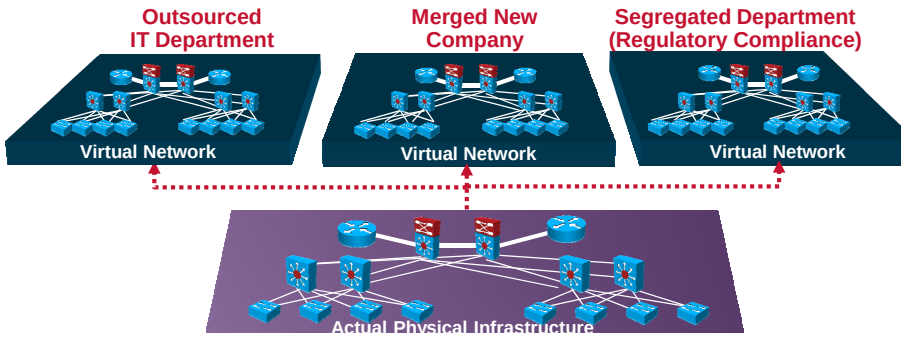
In an MPLS environment, only the routers on the edge of the MPLS domain perform routing lookup. All other routers forward packets that are based on labels. In modern routers, MPLS label switching is not any faster than IP routing and MPLS is not used because of switching performance. What really makes MPLS useful in large enterprise and service provider networks is that it enhances Border Gateway Protocol (BGP) routing and provides different services and applications, such as Layer 2 and 3 VPNs, quality of service (QoS) and traffic engineering (TE). TE is a powerful tool that is used for defining explicit routes through the network, and it uses Resource Reservation Protocol (RSVP) for reserving resources through the network. It is primarily used to optimize network resources and as tool to increase fault tolerance of the network. MPLS VPNs are used to separate customers or departments into virtual networks, historically similar to access with virtual circuits such as Frame Relay networks.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Network Virtualization

- Giving one physical network the ability to support multiple virtual networks
- End-user / customer perspective - connected to a dedicated network
- Maintains Hierarchy and Virtualizes devices, data paths and services
- Allows for better utilization of network resources and security



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

Network virtualization design provides solutions for modern business problems and copes with the highest demands. For example, the modern corporate network needs to allow external visitors Internet access, while preventing any possibility of unauthorized access to internal corporate resources and services. This access control can be achieved through network virtualization, by dedicating a logical virtual network to the external visitor access. There are also more complex scenarios, where enterprise IT departments act as service providers who offer access to the enterprise network to many customers, while simultaneously creating logical isolation between them. The architecture for network virtualization encompasses three functional areas:

- **Access control:** Authenticating, authorizing, and gaining access to network and server resources
- **Path isolation:** Transport of traffic over isolated paths to access VLANs and resources to service edge
- **Service edge:** Providing isolated application environment and access to isolated and shared services

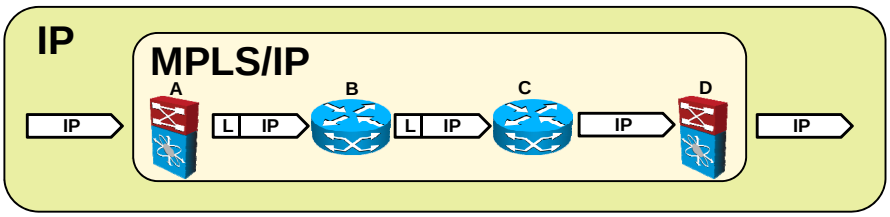
These are the critical factors that compel businesses to consider network virtualization:

- **Cost reduction and consolidation:** Design that uses a single physical network to support multiple users and virtual networks
- **Simpler Operation, Administration, and Maintenance (OAM):** Reducing the amount of network devices that have to be managed and monitored
- **End-to-end compliance and security:** Maintaining segmentation of the network for different departments over a single device, campus, WAN, or data center
- **High availability:** Taking advantage of virtualization through clustering devices that appear as one (increased uptime)
- **Service provisioning:** Rolling out new services faster over common secure infrastructure
- **Data center interconnect (DCI):** Interconnecting Layer 2 and Layer 3 networks across data centers

https://t.me/learningnets

Introduction to Multiprotocol Label Switching

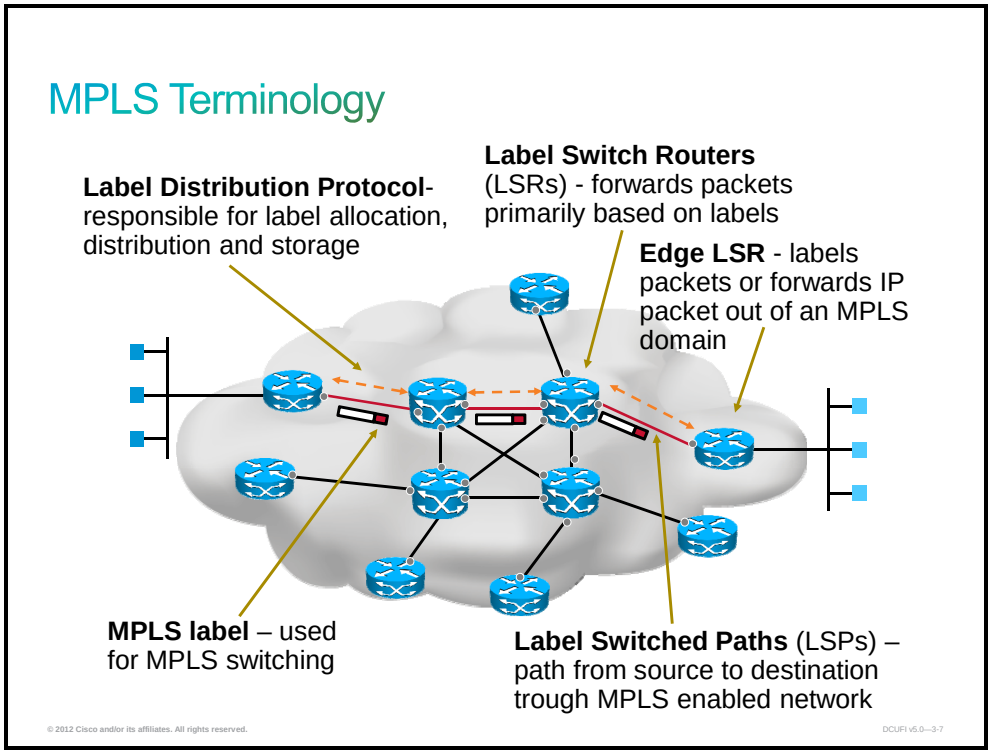
- Technology that enhances IP routing and CEF switching in core networks
- Switching mechanism where packets are switched based on labels:
  - Labels usually correspond to destination IP networks
- Additional header, called MPLS label, is inserted and used for MPLS switching
- LDP is a protocol that is used between MPLS-enabled routers to exchange labels



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

MPLS is a technology that is primarily used in service provider core networks, but in recent years it is also often found in data center environments. It enhances IP routing by implementing switching mechanisms where packets are switched based on labels. MPLS alleviates classic IP routing using Cisco Express Forwarding by introducing an additional header into a packet. This additional header is called an MPLS label. MPLS switches packets that are based on labels lookup instead of IP address lookup. For switching, MPLS uses a 32-bit label header that is inserted between Layer 2 and Layer 3.

https://t.me/learningnets



The MPLS label is used for MPLS switching. This *label* is inserted between Layer 2 and Layer 3 and can be used regardless of the Layer 2 protocol.

In an MPLS domain, there are two types of routers:

- **Label switch router (LSR):** A device that forwards packets that are primarily based on labels
- **Edge LSR:** A device that primarily labels packets or forwards IP packets out of an MPLS domain

LSRs and edge LSRs are usually capable of doing both label switching and IP routing. Their names are based on their positions in an MPLS domain. Routers that have all interfaces that are enabled for MPLS are called LSRs because they mostly forward labeled packets (except for the penultimate LSR).

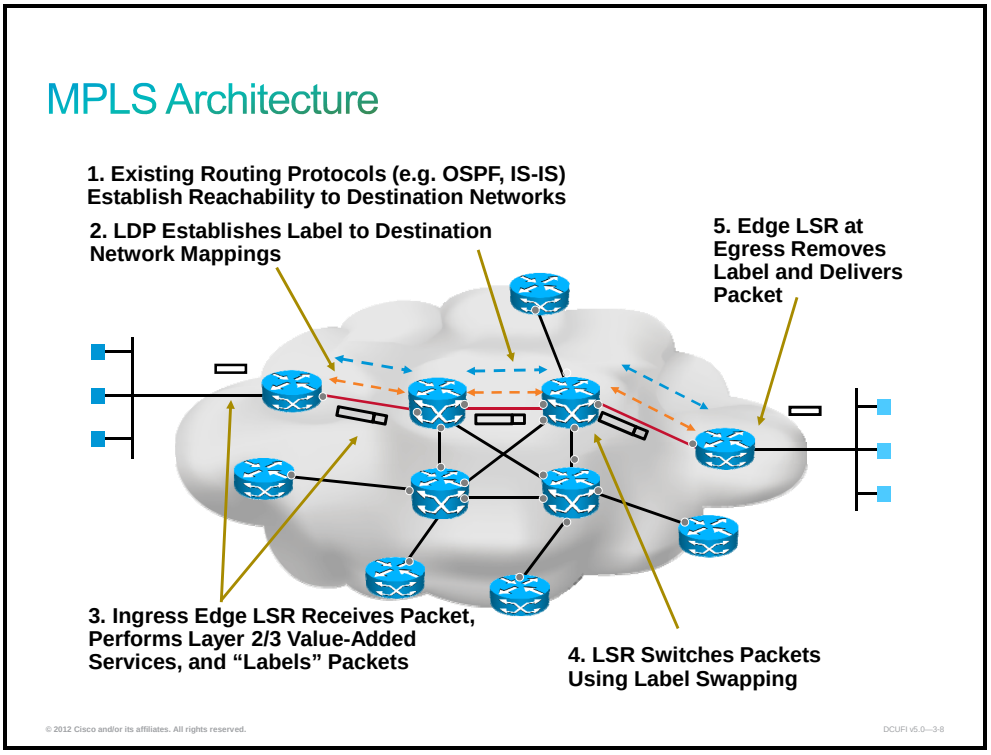
Label Distribution Protocol (LDP), which runs in the router control plane, is responsible for label allocation, distribution, and storage.

A sequence of labels used to reach a destination is called label-switched path (LSP). LSPs are unidirectional, that means that the return traffic uses different LSPs.

The forwarding information base (FIB) table, which consists of destination networks, next hops, outgoing interfaces, and pointers to Layer 2 addresses, is populated by a routing table and Address Resolution Protocol (ARP) cache. The routing table is in turn populated by a routing protocol. Additionally, the MPLS label is added to destination networks if an outgoing interface is enabled for MPLS and a label has been received from the next hop router. LDP is responsible for adding a label to the FIB table.

The label forwarding information base (LFIB) contains incoming (locally assigned) and outgoing (received from next hop) labels. LDP is responsible for exchanging labels and storing them into the LFIB.

https://t.me/learningnets



MPLS architecture can be summarized into the following five steps:

Step 1 Existing interior gateway protocols (IGPs) (for example, Open Shortest Path First [OSPF] and Intermediate System-to-Intermediate System [IS-IS]) establish topology routing information between nodes and establish reachability to the destination networks based on the IP routing table (routing information base [RIB]) and the IP forwarding table (forwarding information base [FIB])

Step 2 LDP uses IGP routing information and establishes label-to-destination network mappings (label binding information that is stored in the label information base [LIB])—such a state as when the routing protocol and LDP have populated all the tables is called steady-state. After the steady-state, all packets are label-switched, except on the ingress and egress router. Although it takes longer for LDP to exchange labels (compared with a routing protocol), a network can use the FIB table in the meantime; therefore, there is no routing downtime while LDP exchanges labels between adjacent LSRs. LDP label mapping exchange (between MPLS nodes) takes place after IGP has converged, because LDP depends on IGP convergence.

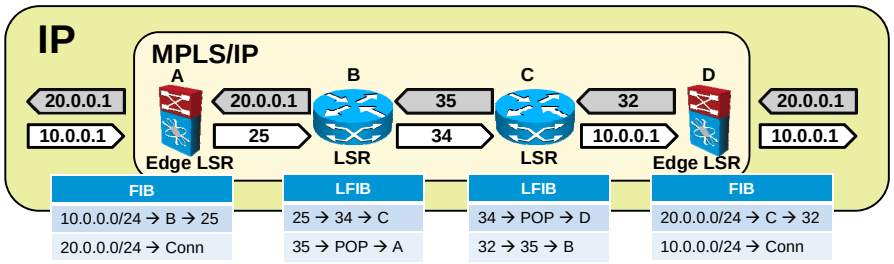
Step 3 The ingress edge LSR receives the packet, performs Layer 2 and Layer 3 value-added services, and "labels" packets. It receives the IP packet, and FIB lookup is performed.

Step 4 LSR switches packets using a label-swapping technique. When a packet carrying a label arrives at a LSR, the LSR uses the label as the index in its LIB. For an incoming label, LIB carries a matching entry with the corresponding outgoing label, interface, and link-level encapsulation information to forward the packet. Based on the information in the LIB, the LSR swaps the incoming label with the outgoing label and transmits the packet on the outgoing interface with the appropriate data link layer encapsulation.

Step 5 Egress LSR removes the label and delivers the packet using the FIB table.

© 2012 Cisco Systems, Inc.
https://t.me/learningnets

MPLS Forwarding Structures



- FIB is used to forward unlabeled IP packets or to label packets if a next-hop label is available.
- LFIB is used to forward labeled packets. Received label is swapped by next-hop label.

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

The data plane on a router is responsible for forwarding packets, which are based on decisions that are made by routing protocols (which run in the router control plane). The data plane on an MPLS-enabled router consists of two forwarding structures:

- **FIB:** When a router is enabled for Cisco Express Forwarding, FIB is used to forward IP packets that are based on decisions that are done by routing protocols. FIB is populated from a routing table and in general includes destination networks, next hops, outgoing interfaces, and pointers to Layer 2 addresses. FIB on an MPLS-enabled router also contains an outgoing label, if an outgoing interface is enabled for MPLS. FIB lookup is done when an IP packet is received. Based on the result, the router can send out an IP packet or a label can be imposed.
- **LFIB:** LFIB is used when a labeled packet is received. LFIB in general contains incoming and outgoing labels, the outgoing interface, and the next hop router. When an LFIB lookup is done, the result can be either to swap a label and send a labeled packet, or the result can be to remove a label and send an IP packet.

These combinations of forwarding packets are possible:

- A received IP packet (FIB) is forwarded based on the IP destination address and sent as an IP packet.
- A received IP packet (FIB) is forwarded based on the IP destination address and sent as a labeled packet.
- A received labeled packet (LFIB) is forwarded based on the label; the label is changed (swapped) and the labeled packet is sent.

A received labeled packet (LFIB) is forwarded based on the label; the label is removed and the IP packet is sent.

https://t.me/learningnets

Label Distribution Protocol (LDP)

- Forwarding structures used by MPLS have to be populated.
- FIB is populated by:
  - Routing table, which is populated by a routing protocol
  - MPLS label is added to the FIB by LDP
- LFIB is populated by:
  - LDP
- LDP is responsible for advertisement and redistribution of MPLS labels between MPLS routers.

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

Forwarding structures that are used by MPLS have to be populated with labels. LDP, which runs in the router control plane, is responsible for label allocation, distribution, and storage.

The FIB table, which consists of destination networks, next hops, outgoing interfaces, and pointers to Layer 2 addresses, is populated by a routing table and ARP cache. The routing table is in turn populated by a routing protocol. Additionally, the MPLS label is added to destination networks, if an outgoing interface is enabled for MPLS and a label has been received from the next-hop router. LDP is responsible for adding a label to the FIB table.

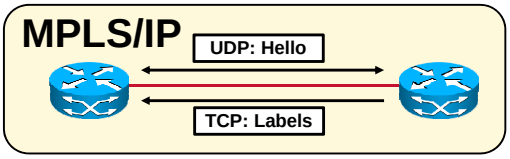
The LFIB contains incoming (locally assigned) and outgoing (received from the next hop) labels. The LDP is responsible for exchanging labels and storing them into an LFIB.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

LDP Session

- Adjacent routers establish LDP session:
  - An MPLS-enabled routers first discover neighbors using hello packets sent to 224.0.0.2 (FF02::2) using UDP on port 656
  - An MPLS-enabled neighbor will respond to hello packets by establishing a TCP session on port 656 to a peer's router ID
- After the LDP session is established, labels can be exchanged.



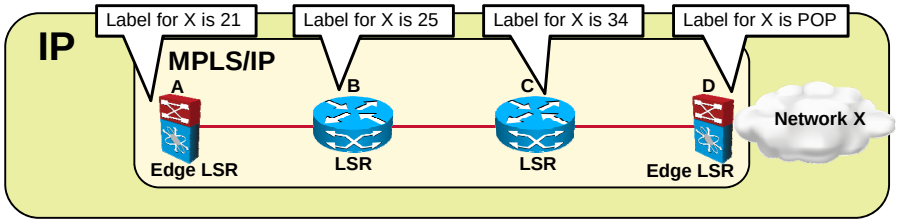
© 2012 Cisco and/or its affiliates. All rights reserved.

Before labels can be exchanged, MPLS-enabled routers must establish adjacencies. This is done in two steps:

- **LDP discovery:** MPLS routers first discover neighbors using hello messages that are sent to all the routers on the subnet as UDP packets with a multicast destination address of 224.0.0.2 (FF02::2 on IPv6) and destination port number of 646.
- **LDP adjacency:** A neighboring MPLS router that received hello packets responds by opening a TCP session with the same destination port number 646, and the two routers begin to establish an LDP session through unicast TCP.

https://t.me/learningnets

Label Allocation



- Each device generates a label for each network in a routing table:
  - Labels have local significance
  - Label allocation is asynchronous
- For path discovery and loop avoidance, LDP relies on routing protocols.
- Networks originating on the outside of the MPLS domain are not assigned any label on the edge LSR – instead the POP label is advertised.

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

First, each MPLS-enabled router has to locally allocate a label for each network that is known to a router. Labels are locally significant (a label for the same network has a different value on different routers), and allocation of labels is asynchronous (the routers assign labels independently of each other).

LDP is not responsible for finding a shortest, loop-free, path to destinations. Instead, LDP relies on routing protocols to find the best path to destinations. If, however, a loop does occur, a Time to Live (TTL) field in the MPLS label prevents a packet from looping indefinitely.

On the edge LSR, networks originating on the outside of the MPLS domain are not assigned a label. Instead, the Post Office Protocol (POP) (implicit null) label is advertised, which instructs the penultimate router to remove a label.

In the example, all routers except the router D assign a label for network X. Router D assigns an implicit null label for that network.

https://t.me/learningnets

Hardware and Software Support

- All MPLS functionality supported on all M-series I/O modules in all Nexus 7000 chassis
- F1 cards will NOT support MPLS natively
- NX-OS 5.2(1) onwards - 8GB Supervisor RAM required if more than 2 VDCs are used with MPLS
- MPLS license is required

I/O Module	Photo	MPLS support
N7K-M132XP-12 N7K-M132XP-12L		✓
N7K-M148GT-11 N7K-M148GT-11L		✓
N7K-M148GS-11 N7K-M148GS-11L		✓
N7K-M108X2-12L		✓

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

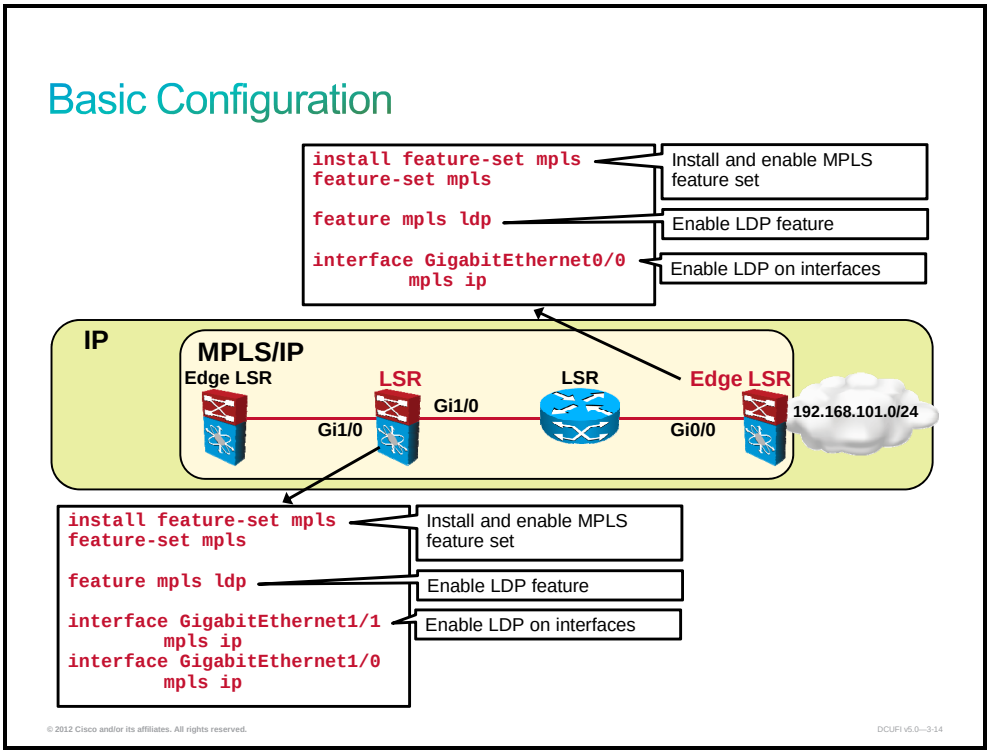
All MPLS functionalities are supported on all Cisco Nexus 7000 M-1 and M-2 Series I/O modules in all Cisco Nexus 7000 Series chassis.

Cisco Nexus 7000 F1-Series cards will NOT support MPLS natively, but they can take advantage of proxy routing with M-1 Series modules for MPLS forwarding.

To enable MPLS functionality on your Cisco Nexus 7000 Series Switches, software on the chassis needs to be upgraded to Cisco Nexus Operating System (NX-OS) Software Release 5.2(1).

The MPLS Services Package (MPLS-PKG), feature license, part number N7K-MPLS1K9 is required for MPLS functionality.

https://t.me/learningnets



MPLS functionalities are grouped into feature sets. A feature set is a collection of components that performs a specific set of functions. The administrator must install and enable the MPLS feature set before configuring the components that belong to particular feature set.

Note MPLS on Cisco NX-OS requires a MPLS license. The administrator of the network must install the MPLS license to be able to enable any features that belong to the MPLS feature set.

MPLS LDP enables LSRs to request, distribute, and release label prefix binding information to other LSR routers in a network. LDP enables LSRs to discover potential peers and to establish LDP sessions with those peers for exchanging label binding information.

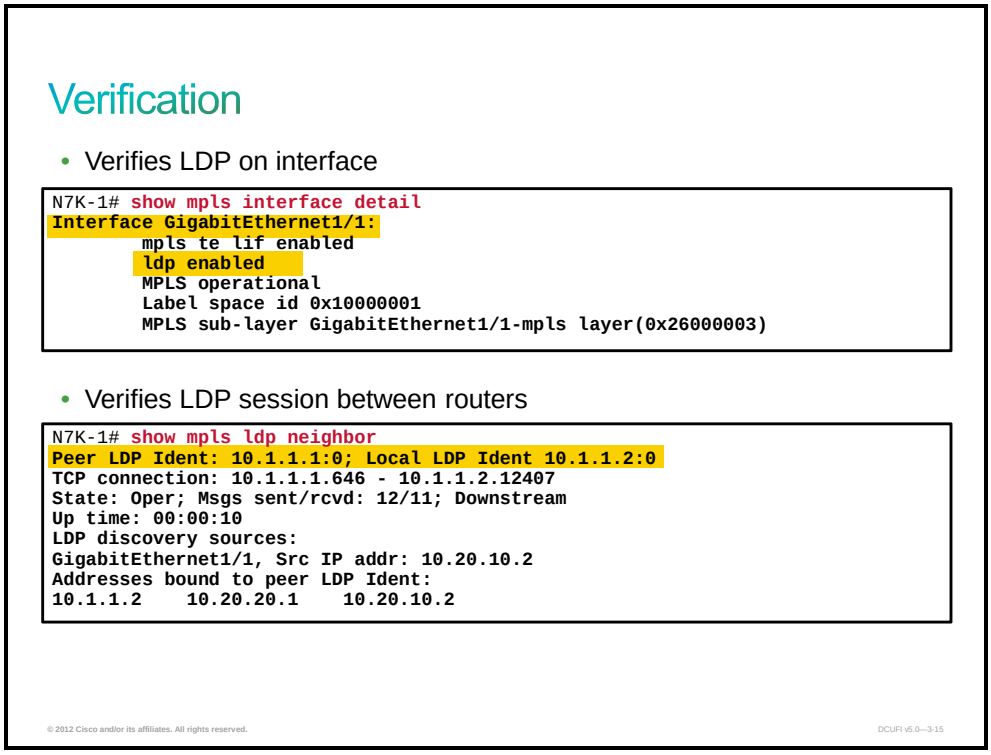
The following commands in global configuration mode are used to install the MPLS feature that is set and enable MPLS and LDP feature sets on the device:

- **install feature-set mpls**
- **feature-set mpls**
- **feature mpls ldp**

To enable LDP on the Cisco NX-OS switch, first enter the interface configuration mode for a desired interface. Then enable LDP using the **mpls ip** command. In the example, MPLS is enabled on GigabitEthernet0/0 interface at Edge LSR and on GigabitEthernet0/1 and GigabitEthernet1/1 on LSR in the core.

Note For a complete command reference, refer to the *Cisco Nexus 7000 Series Command Reference Master Index* at http:/www.cisco.com/en/US/docs/switches/datacenter/sw/6_x/nx-os/master/index/6_xMasterIndex.html.

© 2012 Cisco Systems, Inc.
https://t.me/learningnets



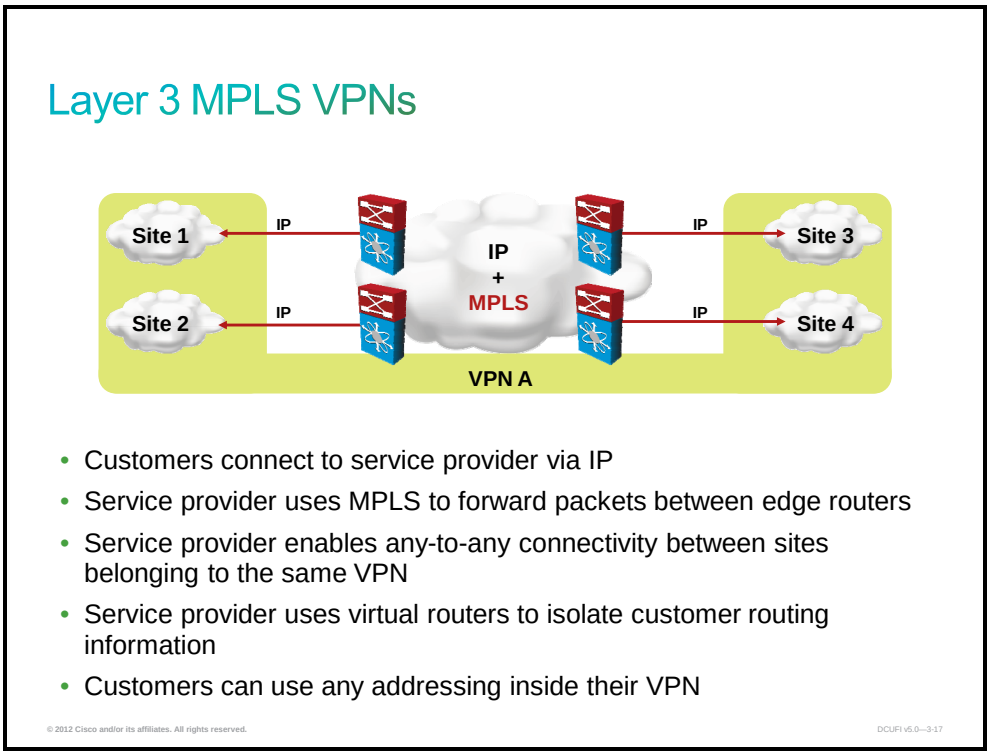
There are various **show** commands that are available to verify MPLS operations. The following example shows how to verify that interface Gigabit Ethernet 1/1 has been configured to use LDP.

To verify LDP neighbors and the state of LDP sessions, use the **show mpls ldp neighbor**. The output displays the LDP identifier, the router ID of a peer (IP address of the highest configured loopback interface, or the highest IP address of an active interface, if loopback is not configured), and the TCP connection.

https://t.me/learningnets

Multiprotocol Label Switching VPNs

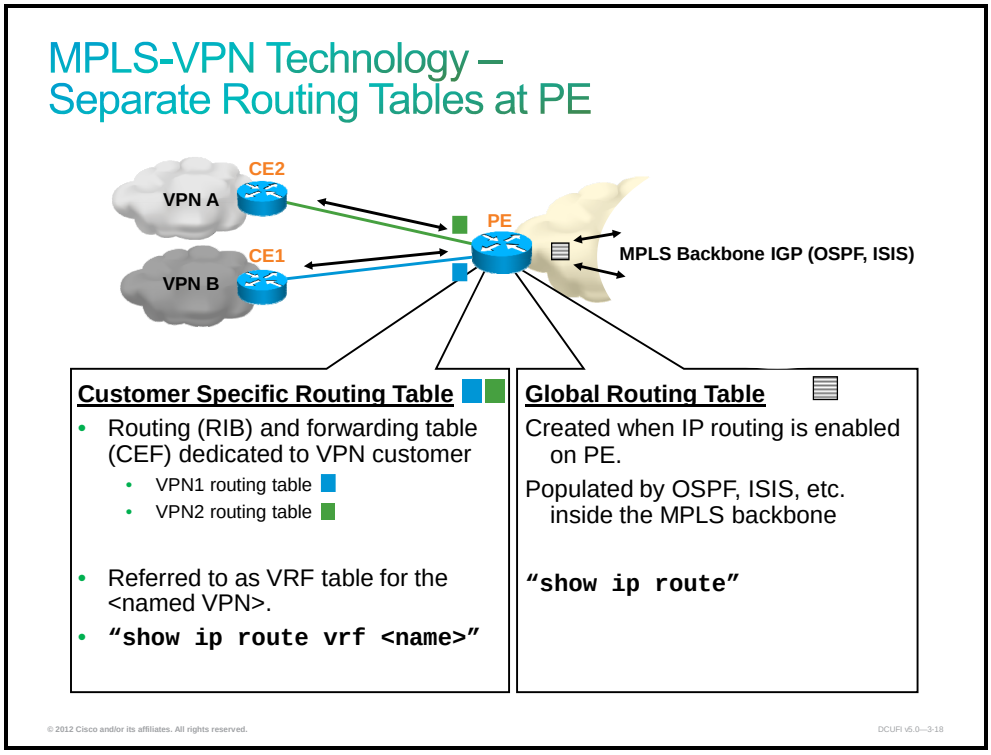
This topic describes how to understand the benefits, operation, and configuration of MPLS VPNs in the Data Center Network Architecture.



The main characteristic of Layer 3 MPLS VPNs is that customers connect to a service provider via IP. They need to establish IP routing (static or dynamic) to exchange routing information between customer sites that belong to the same VPN. Because different customers might use the same private IP address ranges, the service provider cannot perform normal IP forwarding. Instead, MPLS must be used, to ensure isolation in the data plane between packets belonging to different customers, yet potentially having the same IP addresses. Virtual routers (virtual routing and forwarding instances [VRFs]) are used on service provider routers to isolate customer routing information. MPLS seamlessly provides any-to-any connectivity between sites that belong to the same VPN.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets



Traditional Cisco software can support several different routing protocols. In some cases, even several completely isolated copies of the same routing protocol are supported. For example, several OSPF processes can be used.

It is important to understand that for several important routing protocols, such as Routing Information Protocol (RIP), Enhanced Interior Gateway Routing Protocol (EIGRP), IS-IS, or BGP, Cisco IOS Software supports only a single copy of the protocol running in the router. These protocols cannot be used directly between provider edge (PE) and customer edge (CE) routers in VPN environments because each VPN (or, more precisely, each VRF) needs a separate, isolated copy of the routing protocol to prevent undesired route leakage between VPNs. Furthermore, VPNs can use overlapping IP address spaces (for example, each VPN could use the subnetworks of network 10.1.1.0/24), which would also lead to routing confusions if all VPNs shared the same copy of the routing protocol.

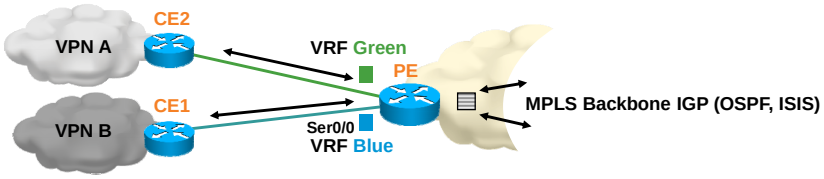
"Routing contexts" were introduced in Cisco IOS Software to support the need for separate isolated copies of VPN routing protocols. Routing contexts can be implemented as separate routing processes (in OSPF), similar to traditional Cisco IOS Software implementation, or as separate isolated "instances" of the same routing protocol.

If the routing contexts are implemented as instances of the same routing protocol, each instance contains its own independent routing protocol parameters. Examples would include networks over which the routing protocol is run, timers, authentication parameters, passive interfaces, and neighbors. This independence allows the network designer maximum flexibility in implementing routing protocols between PE and CE routers.

https://t.me/learningnets

Virtual Routing and Forwarding Instance (VRF)

- Representation of VPN customer inside the MPLS network
- Each VPN is associated with at least one VRF
- VRF configured on each PE and associated with PE-CE interface(s)
- VRF-aware routing protocol (static, RIP, BGP, EIGRP, ISIS, OSPF)
- No changes needed at CE



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

A VRF is a routing and forwarding instance that you can use for a single VPN site or for many sites that are connected to the same PE router if, and only if, these sites share exactly the same connectivity requirements.

The major data structure that is associated with MPLS VPN implementation on Cisco IOS platforms is the VRF table. It contains routes that should be available to a particular set of sites. VPN interfaces are assigned to VRF in the way that multiple interfaces can be assigned to one VRF, but interface can belong to only one VRF. VRF table data structure encompasses an IP routing table that is identical in function to the following:

- The global IP routing table in Cisco IOS Software
- A Cisco Express Forwarding table that is identical in function to the global Cisco Express Forwarding table FIB
- Specifications for routing protocols running inside the VRF instance

Other MPLS VPN attributes that are associated with a VRF table are as follows:

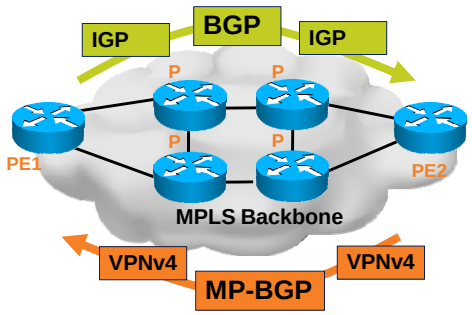
- The route distinguisher (RD), which is prepended (for example, RD + IP address) to all routes exported from the VRF into the global VPNv4 (also called VPN IPv4) BGP table
- A set of export route targets (RTs) which is attached to any route exported from the VRF
- A set of import RTs, which is used to select VPNv4 routes that are to be imported into the VRF

Commonly, the customer sites are connected to a number of PE devices. The example in the figure shows only on a single PE router, which contains two VRFs—one for each customer. Each customer is connected to the PE router. Each of VPNs belonging to these customers (VPN A and VPN B) is represented by VRF. VRF is configured on each PE router and interface that connects the CE and PE router that is assigned to that PE router.

https://t.me/learningnets

How MPLS VPNs Work

- Networks are learned via an IGP from a customer or via BGP from other MPLS backbone routers.
- Labels/VPNv4 address prefixes are propagated via MP-BGP. Two labels are used:
  - The top label points to the egress router.
  - The second label identifies egress VRF



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

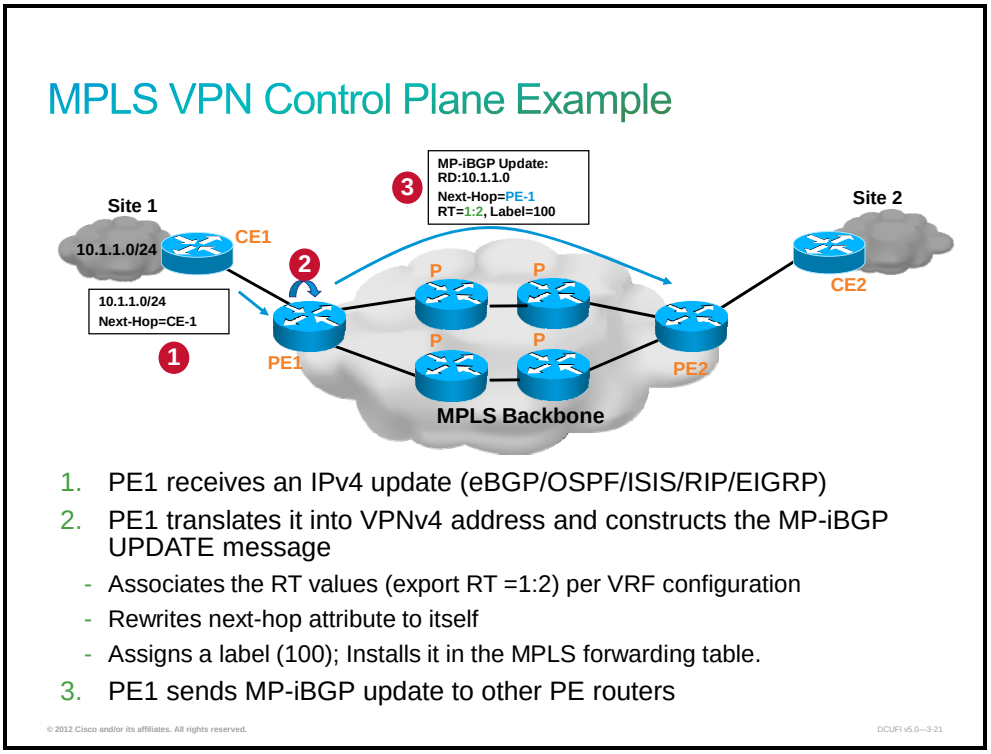
In the MPLS network, the PE router provides isolation between customer networks. However, this information needs to be carried over the MPLS backbone network. Because every customer has its own address space, which can overlap between customers, the PE router must learn those routes via the IGP routing protocol if they are not statically configured. This routing information must be carried over the MPLS VPN network to other PE routers. The PE installs the backbone routes (IGP) in the *global routing table* and the VPN routes in the *VRF routing table* or tables.

The RD is a 64-bit identifier that is prepended to the customer network IGP-learned prefixes. The RD, together with the 32-bit customer network prefixes, forms a 96-bit long address, which is called the VPNv4 address. A unique RD must be configured per VRF instance on the PE router.

VPNv4 prefixes are exchanged between PE routers using the Multiprotocol BGP (MP-BGP) protocol. The MP-BGP is also responsible for label assignment, specifically two labels:

- Top label or outer label that points to the egress router
- Inner label that identifies the egress VRF

https://t.me/learningnets



The MPLS VPN control plane function can be summarized in several steps as shown in the example in the figure:

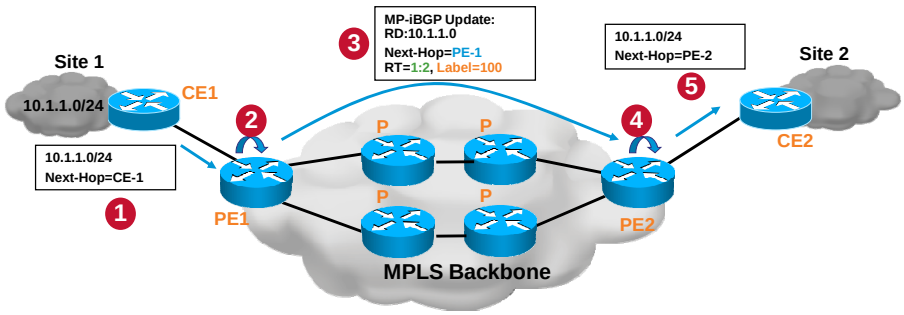
Step 1 The PE1 router receives an IPv4 update via the routing protocol that is configured on the customer side.

Step 2 The PE router creates a VPNv4 address by associating the configured RD with the prefix received in Step 1. It also associates the export RT value per the configuration on the VRF that is assigned to the customer at Site 1. The RT identifies which VRFs keep which routes. This VPNv4 address makes the customer IPv4 address unique inside the service provider MPLS network. It then rewrites the next hop attribute to itself and assigns a label to be inserted into the MPLS forwarding table.

Step 3 PE1 sends the Multiprotocol Internal Border Gateway Protocol (MP-IBGP) to other PE routers.

© 2012 Cisco Systems, Inc.
https://t.me/learningnets

MPLS VPN Control Plane Example (Cont.)



4. PE2 receives and checks whether the RT=1:2 is locally configured as ‘import RT’ within any VRF, if yes, then
   - PE2 translates VPNv4 prefix back in IPv4 prefix
   - Updates the VRF CEF Table for 10.1.1.0/24 with label=100
5. PE2 advertises this IPv4 prefix to CE2 (using whatever routing protocol)

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

Step 4 PE2 receives the MP-IBGP update and checks whether the RT that is received is locally configured as “import RT” within any VRF. If it finds the match, PE2 translates the VPNv4 prefix back into the IPv4 prefix. It then updates the VRF Cisco Express Forwarding table for the received customer prefix and the label.

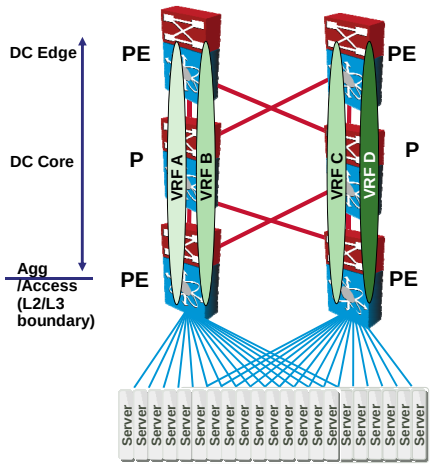
Step 5 The PE2 router then advertises the IPv4 prefix to CE2 using the routing protocol that is configured between these two.

https://t.me/learningnets

MPLS L3 VPN in Data Centers

Modern Data Centres require

- Scalable segmentation for IPv4 & IPv6 traffic
- Simpler security management
- Dedicated paths between Data Centre locations
- consolidation and access to shared services
- Predictable QoS provisioning trough the network

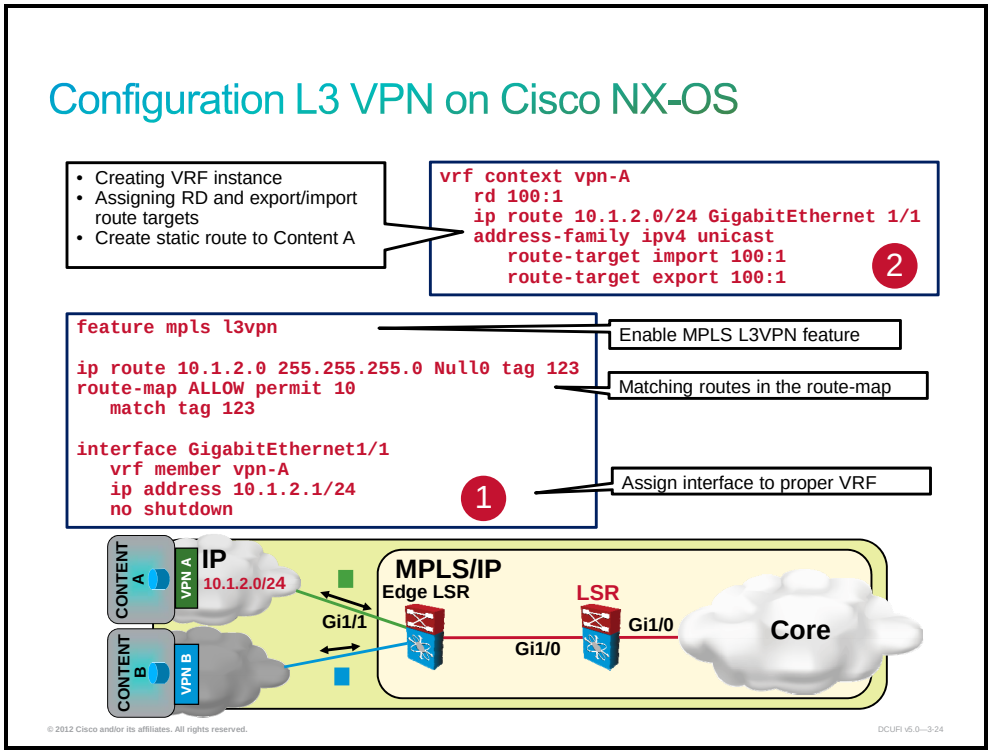


© 2012 Cisco and/or its affiliates. All rights reserved.

MPLS VPN technology can offer several modern requirements that are needed in the data center of today:

- It provides scalable segmentation for both IP version 4 (IPv4) and IP version 6 (IPv6) traffic, allowing very large data centers to scale almost without limits. Traditional VLAN separation is limited by the number of VLANs that can be theoretically created.
- It simplifies management of customer traffic, due to simpler security separation. It also increases enterprise outsourcing flexibility and provides opportunity for service providers.
- It overrides problems with active/active data centers that are spread across the network that needs dedicated traffic path between locations.
- It offers consolidation and access to shared services such as IP telephony or video services by using multicast technology and separation of user groups by VPN.
- It provides QoS provisioning using MPLS transport network and traffic engineering methods.

https://t.me/learningnets



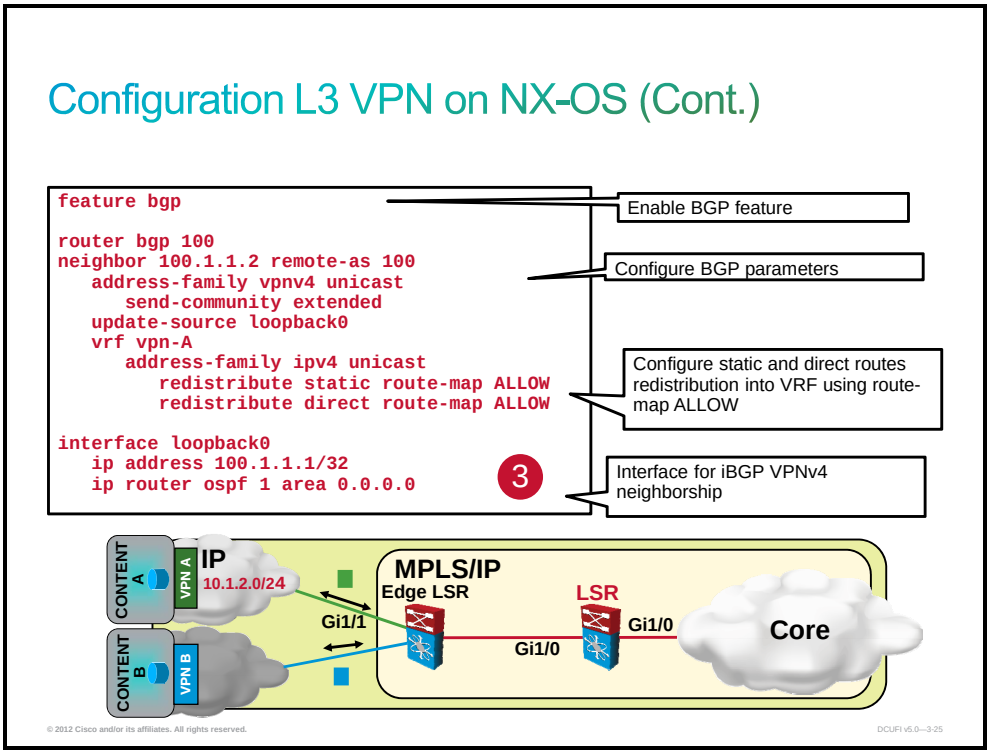
The example in the figure shows how to configure a Cisco Nexus 7000 Series switch to act as an MPLS Layer 3 VPN PE router. Before configuring MPLS Layer VPN on Cisco Nexus NX-OS, the administrator must ensure that the switch is in the correct VDC. (If it is not, the administrator must use the **switchto vdc** command.)

There are several routing protocols that the customer can use to advertise their network prefixes. In the example that is provided in the figure, the Cisco Nexus 7000 switch acts as the PE router in the MPLS network. It connects to the two contents, Content A and Content B, using two different physical interfaces, so that the Gigabit Ethernet 1/1 interface is connected to the server that offers Content A and that interface is assigned to the configured VRF, VPN A.

One of the first mandatory things is to enable MPLS L3VPN feature as show in the example. Prefixes are redistributed into BGP using a route map that allows traffic that is tagged with the tag *123*.

In the VRF context configuration section, the RD must be chosen for the creation of VPNv4 advertisements. In this section, the static route to Content A is also configured, and the import and export route targets are defined in the address family VRF configuration.

https://t.me/learningnets



Propagation of the IGP must be done using the BGP protocol. The BGP feature must be activated before any BGP configuration. After the BGP parameters are configured, redistribution of routes into the VRF is configured using the **redistribute** command. This way, the MP-IBGP protocol is triggered to redistribute routes that are listed in this configuration section.

The configuration of the interface that is used for IBGP VPNv4 neighbor relationship is also shown in the example.

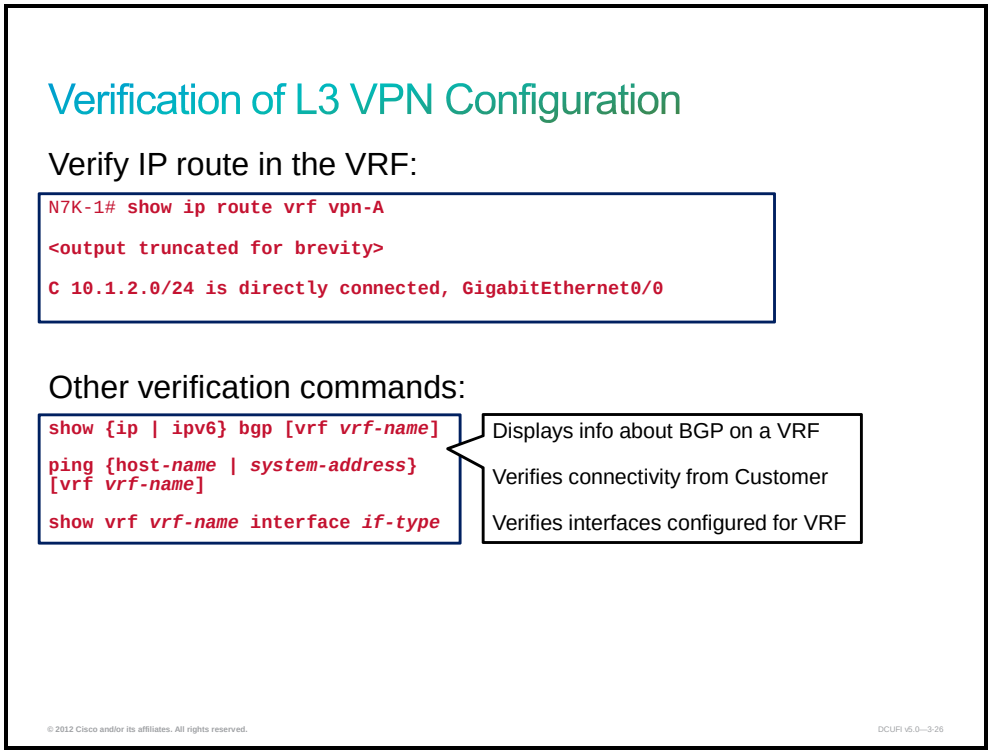
The following commands are used:

- **feature mpls l3vpn:** Enables the MPLS Layer 3 VPN feature.
- **vrf context** *vrf-name***:** Defines the VPN routing instance by assigning a VRF name.
- **rd** *route-distinguisher*: Configures the RD. The *route-distinguisher* argument adds an 8-byte value to an IPv4 prefix to create a VPN IPv4 prefix configuration mode.
- **address-family {ipv4 | ipv6}:** Specifies the IPv4 address family type and enters the address family configuration mode.
- **route-target {import |export}***route-target-ext-community***:** Specifies an RT extended community for a VRF as follows:
  - The **import** keyword imports routing information from the target VPN extended community.
  - The **export** keyword exports routing information to the target VPN extended community.
  - The *route-target-ext-community* argument adds the RT extended community attributes to the VRF's list of import or export RT extended communities.
- **vrf member** *vrf-name*: Associates a VRF with the specified interface or subinterface. The vrf-name argument is the name that is assigned to a VRF.

© 2012 Cisco Systems, Inc.
https://t.me/learningnets

- **{ip | ipv6} route** *prefix nexthop*: Defines static route parameters for every PE-to-CE session. The prefix and next hop are as follows:
  - IPv4—in dotted decimal notation
  - IPv6—in hexadecimal format
- **feature bgp**: Enables the BGP feature.
- **redistribute static route-map** *map-name*: Redistributes static routes into BGP.
- **redistribute direct route-map** *map-name*: Redistributes directly connected routes into BGP.

Note For a complete command reference, refer to the *Cisco Nexus 7000 Series Command Reference Master Index* at http:/www.cisco.com/en/US/docs/switches/datacenter/sw/6_x/nx-os/master/index/6_xMasterIndex.html



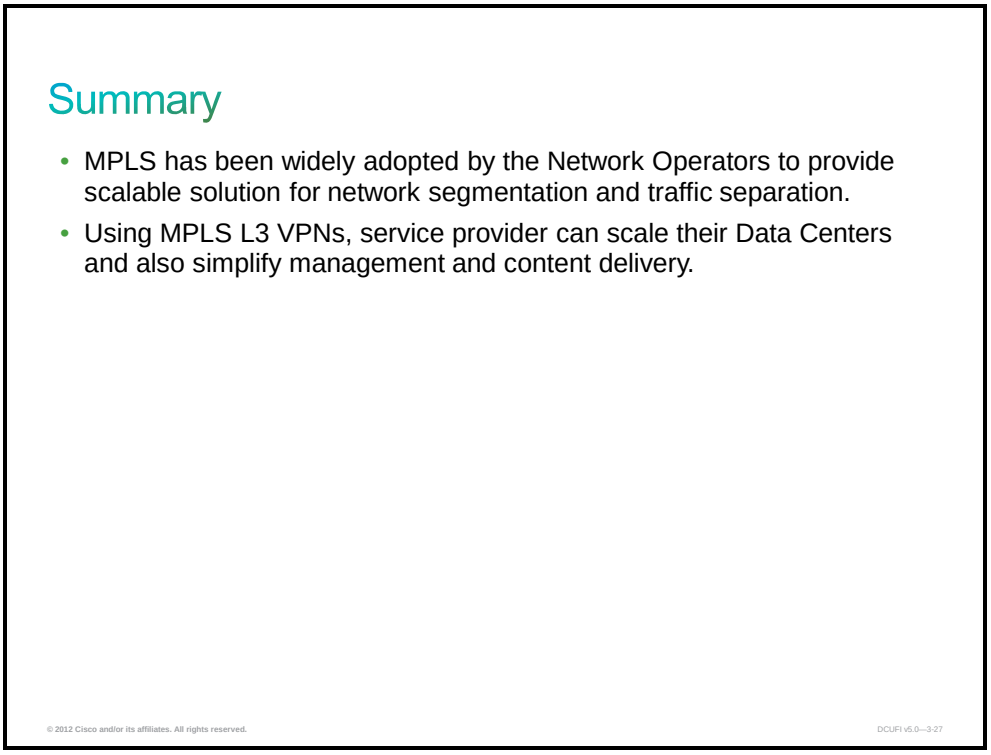
Layer 3 VPN configuration can be verified using several CLI commands that are shown in the figure.

One of the main goals is to verify that customer routes are accessible through the MPLS network and routes are contained in the customer VRF routing table.

https://t.me/learningnets

Summary

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



- MPLS enhances IP routing by implementing a switching mechanism where packets are switched based on 32-bit labels.
- LDP is a protocol that is used between MPLS-enabled routers to exchange labels.
- Basic MPLS configuration is done using a few commands.
- Modern data centers require a scalable solution for isolation of customer content.
- The service provider uses virtual routing contexts (VRFs) to isolate customer routing information.
- Customer routing information is redistributed using the MP-BGP protocol that is configured on a PE device.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Lesson 3

Configuring LISP

Overview

Locator/ID Separation Protocol (LISP) is a next-generation IP routing feature that enables separation of the identity of a device and its location. Creation of separate namespaces for endpoint identification and the location of the endpoint creates space for greater scalability of the Internet routing system and improved efficiency of multihomed locations. This lesson explains LISP operation and gives examples of LISP configuration.

Objectives

Upon completing this lesson, you will be able to describe and configure LISP in the Cisco Data Center Network Architecture. This ability includes being able to meet these objectives:

- Describe the terminology, components, and operation of LISP in the Cisco Data Center Network Architecture
- Configure LISP in the Data Center Network Architecture

https://t.me/learningnets

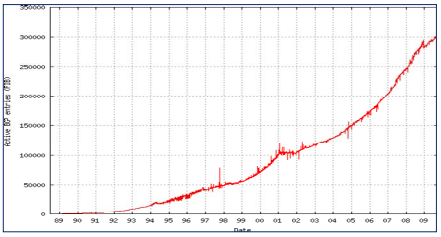
Locator/ID Separation Protocol

This topic explains the terminology, components, and operation of LISP in the Cisco Data Center Network Architecture.

Locator/ID Separation Protocol

LISP, originally conceived to address internet scaling

- What causes scaling issues?
  - IP addresses denote both location and identity today.
  - Overloaded IP address semantics makes efficient routing impossible.
- Why are scaling issues bad?
  - Routers require lots of expensive memory.
  - It is expensive for network builders and operators.



"… routing scalability is the most important problem facing the Internet today and must be solved … "

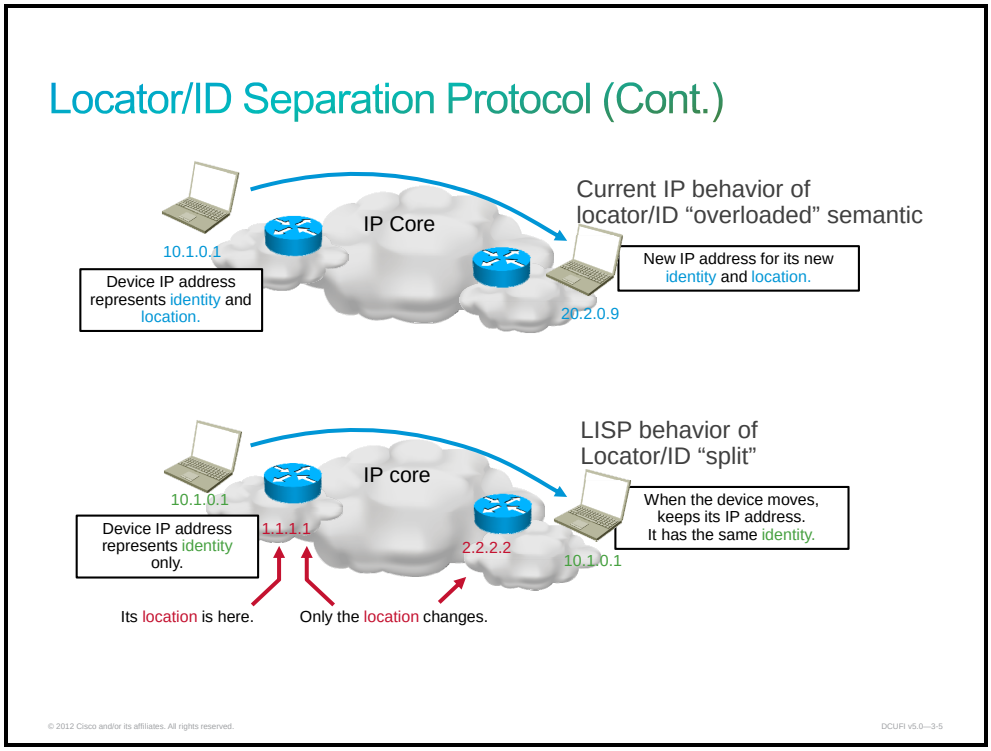
Internet Architecture Board (IAB)
October 2006 Workshop (written as RFC 4984)

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

In current Internet architecture, the IP address of an endpoint denotes both its location and identity. LISP is a protocol that enables separation of endpoint identification and its location. Current Internet architecture combines two functions into one namespace: the IP address. LISP enables separation of the two functions into two distinctive namespaces: hosts and locators. The basic idea of the separation solves problems of scalability or increasing the default-free zone (DFZ) as a consequence of multihoming, and also adds some features and functionalities that an enterprise or service provider may consider in their networks.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

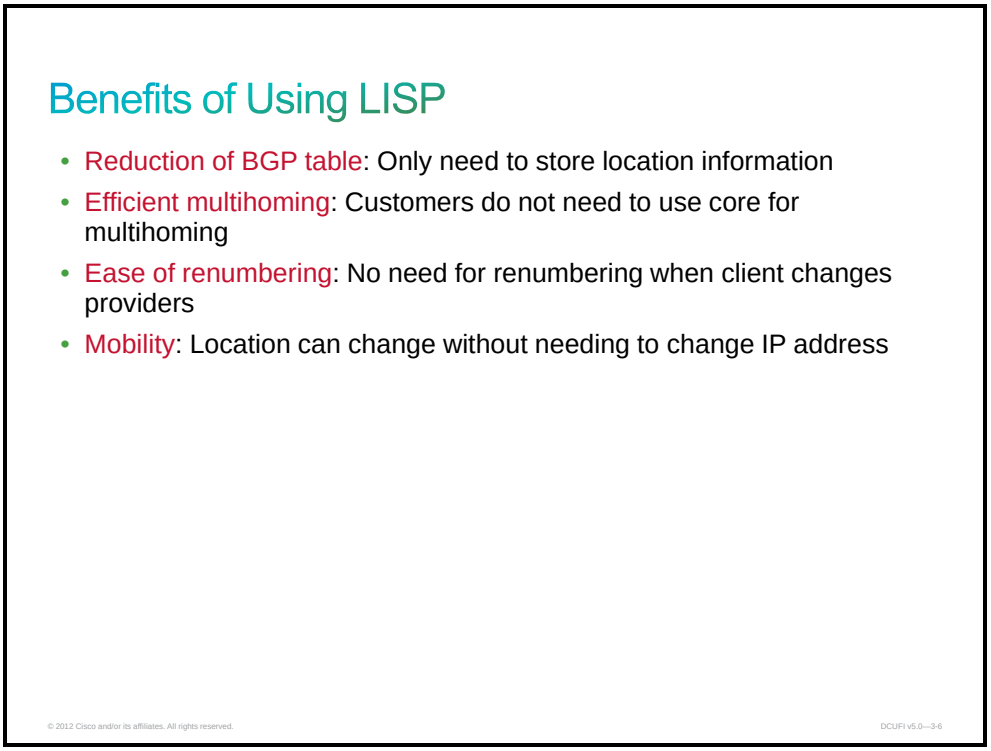


During the history of the Internet, it has become obvious that use of a single IP address space for both endpoint location and identification creates routing scaling problems. One of the very visible results of this behavior is the rapid growth of DFZ as a consequence of multihoming, traffic engineering, and nonaggregatable address locations.

In the first part of the diagram shown in the figure, without LISP, changing location from one place to another also changes the client identifier. With LISP, only the location identifier is changed and the client keeps its original IP version 4 (IPv4) or IP version 6 (IPv6) address. Splitting the device identity, which is known as an endpoint identifier (EID), and its location, which is known as its routing locator (RLOC), into two different namespaces allows you to create separate databases for EID and RLOC functions.

The EID addresses that hosts use for tracking sockets and connections, and sending and receiving packets, do not change during operations. Routers forward packets that are based on IP destination addresses. Because client-generated packets are encapsulated using a LISP header and outer IP header, outer addresses represent the RLOC identifier.

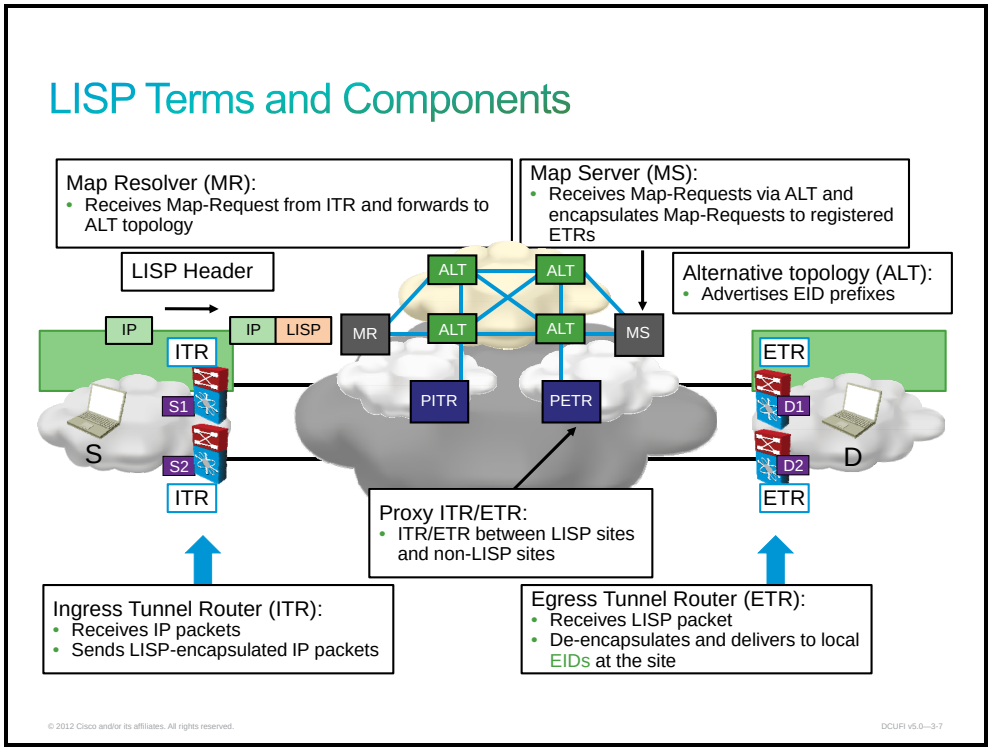
https://t.me/learningnets



Numerous benefits could be achieved using the LISP approach. Here are some of the benefits:

- **Reduction of Border Gateway Protocol (BGP) routing table:** Using a separate address space for identification and location allows a significant reduction of aggregation information in the routing tables.
- **Efficient multihoming:** Customers that are connected to different service providers can control their own policies of packet flow into their sites without using extra routing resources in the core.
- **Ease of renumbering:** When a client changes providers, the client does not need to be renumbered, because the host EID is in the separate address space of routing information.
- **Mobility without address change:** With LISP, it is possible to move to a different point in the network and keep the original IP address or acquire a new address based on the new network location, whether the new location is at the same physical location but at a different service provider, or in a different physical location.

© 2012 Cisco Systems, Inc.
https://t.me/learningnets



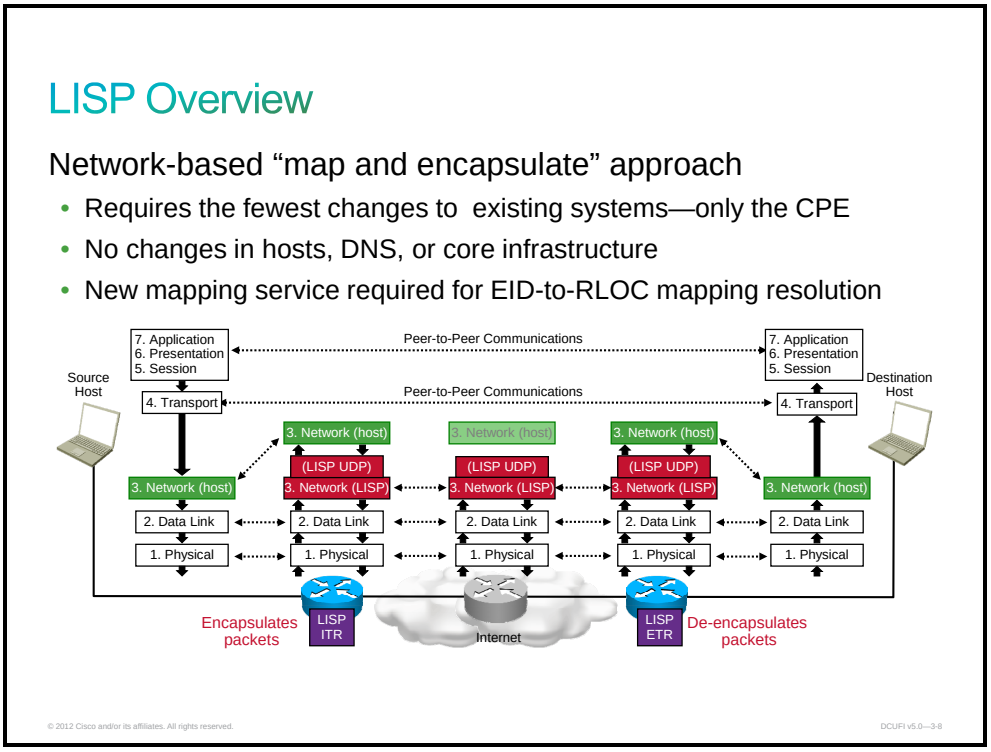
The main components and their functions are briefly described here:

- **LISP header**: Referred to as an outer IPv4 or IPv6 header, a UDP header and a LISP 8-byte header that follows the UDP header.
- **Ingress Tunnel Router (ITR)**: A router that accepts IP packets and encapsulates them into the LISP header and performs an EID-to-RLOC mapping lookup. The router prepends the outer LISP header that is based on the response of the EID-to-RLOC mapping lookup. The outer LISP header contains a globally routable RLOC in the source address field, and the mapping lookup response is contained in the destination address field. The destination address in the LISP header does not mean that it must be an ETR address; it can also be a proxy device that has more information about EID-to-RLOC mappings that are closer to the destination EID.
- **Egress Tunnel Router (ETR)**: A router that accepts an LISP packet that has a destination address that matches its own RLOC. This router strips the outer LISP header and forwards the packet that is based on the inner IP header.
- **Proxy ITR and proxy ETR (PITR and PETR)**: Performs the same function as ITR and ETR but is used between LISP sites and non-LISP sites.
- **Map Resolver (MR)**: Receives Map-Request messages from the ITR and forwards them through the alternative topology, or ALT topology. Map-Request messages are sent from an ITR when it needs an EID mapping, to test an RLOC for reachability.
- **Map Server (MS)**: LISP ETR registers associated EID prefixes to the MS. It also injects routes for registered LISP sites into the ALT topology. ALT receives Map-Request messages via ALT topology and encapsulates Map-Request messages to registered ETRs. Map-Reply messages are sent from an ETR in response to a valid map request to provide the EID and RLOC mapping and site ingress policy for the requested EID. This information is sent to MS.
- **Alternative topology (ALT)**: A device that is deployed to build out an overlay network that provides a mechanism for managing EID prefix aggregation. It advertises EID prefixes in an alternate BGP topology over Generic Routing Encapsulation (GRE), including to the MR.

https://t.me/learningnets

The following terms are also used in LISP:

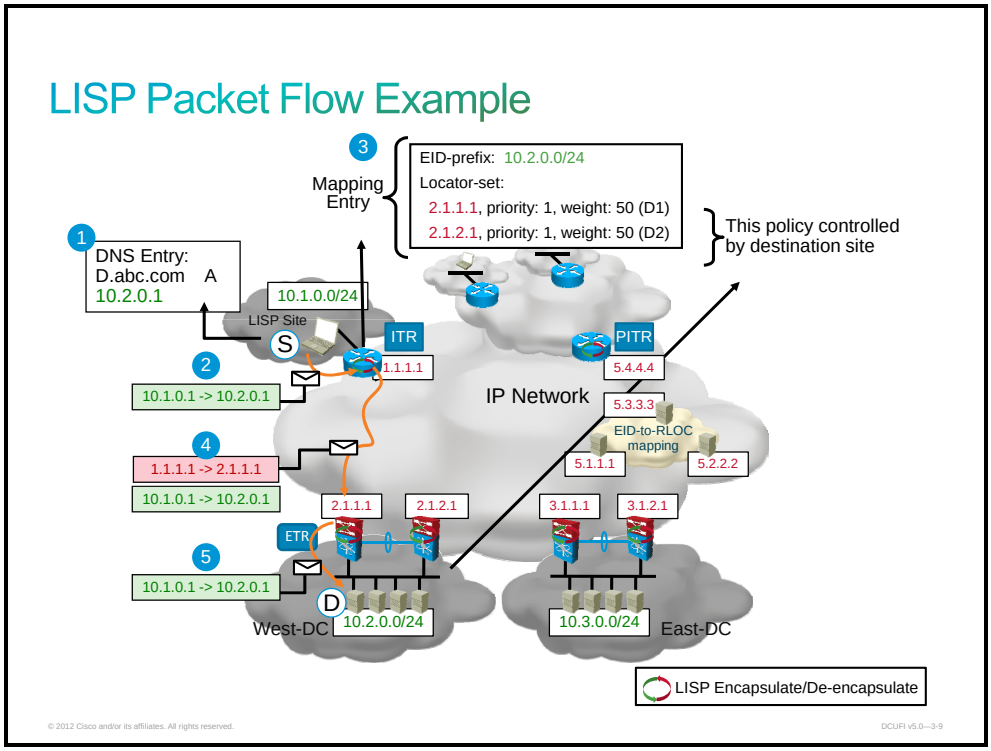
- **Provider-independent (PI) address**: An address block that is assigned from the address pool where address blocks are not associated with any particular location in the network, typically from the service provider. It is not topologically aggregatable.
- **Provider-assigned (PA) addresses**: A block of IP addresses that are assigned to a site by each service provider to which the site is connected. It is typically aggregated in the routing system, and LISP uses only aggregatable addresses for the RLOC identifiers.
- **Routing Locator (RLOC)**: The IPv4 or IPv6 address of the customer router that de-encapsulates the LISP packet that is leaving the LISP network. It is considered as a location address of ETR to which the packet is destined.
- **Endpoint Identifier (EID)**: IPv4 or IPv6 addresses that are used in the source or destination address fields in the inner LISP header of the packet.



As mentioned earlier, LISP packets are encapsulated through the network. For LISP operations only customer premises equipment (CPE) devices perform the mapping and encapsulation process, and no changes are needed for the hosts, Domain Name System (DNS) servers, or in the core infrastructure. As with DNS, which requires DNS service for proper operation, LISP also requires EID-to-RLOC mapping resolution service for proper operation.

At the ITR, EID-to-RLOC mapping is performed to add new LISP headers to the packet. The destination IPv4 or IPv6 address is the RLOC address of the ETR router that is obtained via the mapping request process. The packet is routed through the LISP network, and it arrives at the ETR router. When the ETR router receives the LISP packet that has the destination address that matches the local RLOC address, it then de-encapsulates the packet and forwards it to the destination that is found in the de-encapsulated inner IP packet.

© 2012 Cisco Systems, Inc.
https://t.me/learningnets



This figure gives a brief overview of the LISP packet flows from the source to the destination. The following steps are involved:

Step 1 The source host is sending a packet to D.abc.com to resolve the DNS name to the IP address, the same as it would do if it was not using LISP.

Step 2 The packet is forwarded to an ITR, which is configured as a default router in this scenario. The destination address in the packet is the EID address of the DNS server that is located at the remote site. The source address is the EID address of the host that performs the DNS request.

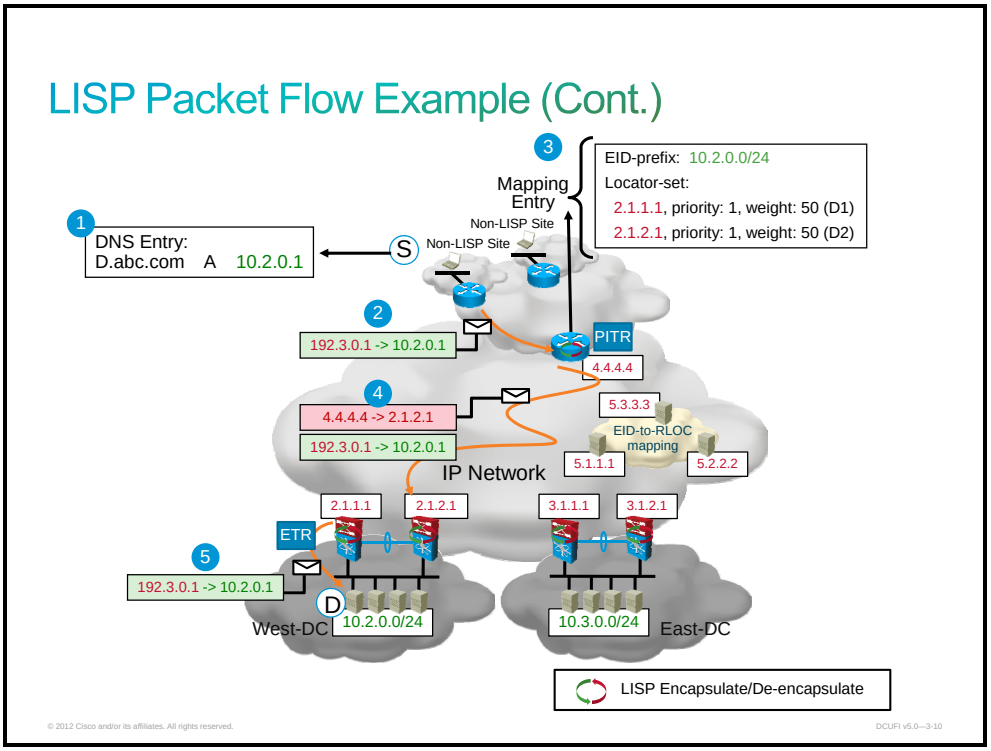
Step 3 The ITR maps the EID destination to the RLOC of the ETR at the destination site. Mapping is accomplished using the local cache, or using map request and map reply mechanisms. The ITR prepends the LISP header to the packet. The LISP header contains the source IPv4 or IPv6 address of one of the ITR RLOC IP addresses. Until the map request and map reply process is completed, the destination EID address is used as the destination IP address in the LISP header. Subsequent packets have a destination RLOC address in the LISP header.

Step 4 The LISP packet with the source IP address of the ITR RLOC and the destination address of the ETR RLOC arrives on the ETR device.

Step 5 When the ETR router receives the LISP packet that has the destination address that matches the local RLOC address, it then de-encapsulates the packet and forwards it to the destination that is found in the inner IP packet. If the ETR receives a packet that has the destination EID in the LISP header, the EID-to-RLOC Map-Reply message is originated by the ETR and is addressed to the source RLOC (ITR).

In this example, two RLOCs are received with the same weight of 50. This means that load balancing will be used to send packets to the destination using RLOCs at the destination site.

https://t.me/learningnets



In the example that is seen in the figure, the DNS request is sent from the non-LISP-enabled site. The following steps are involved in this scenario:

Step 1 The host from the non-LISP site performs a DNS lookup and sends the packet to the default gateway.

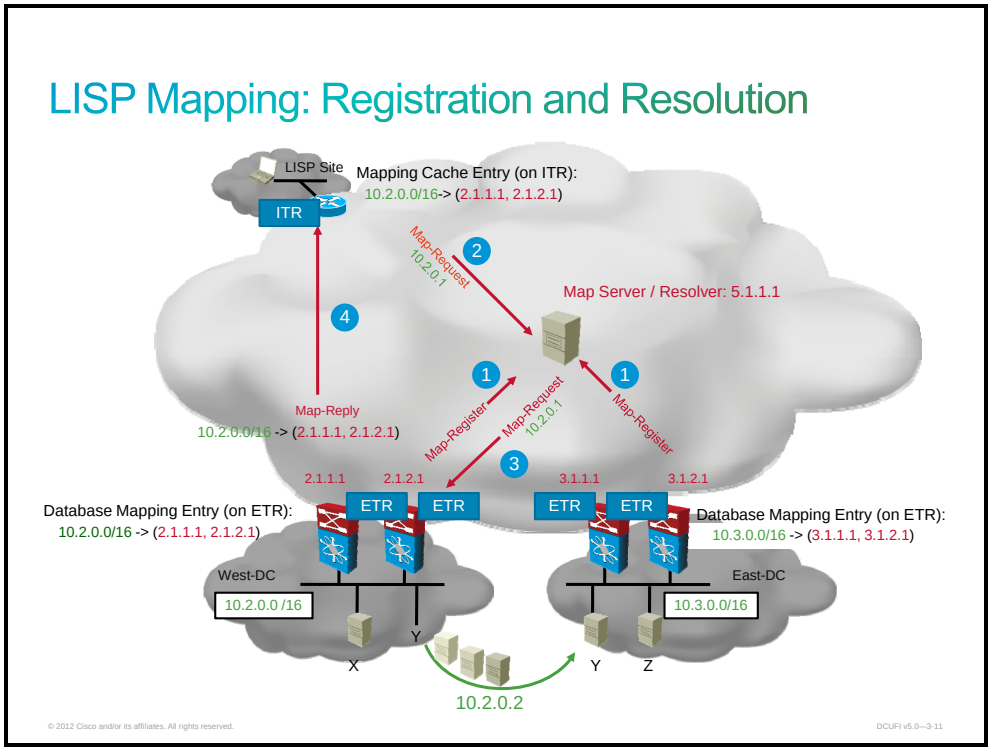
Step 2 Using the routing process, the packet reaches the PITIR router with the source IP address of the source host and the destination IP address as destination EID.

Step 3 At the PITR router, mapping lookup is performed. For the same EID, multiple RLOCs can be configured with different weights for load balancing.

Step 4 Packets are encapsulated with the LISP header with the source RLOC of the PITR router and the destination address of the ETR RLOC.

When the ETR router receives the LISP packet that has the destination address that matches the local RLOC address, it then de-encapsulates the packet and forwards it to the destination that is found in the inner IP packet.

© 2012 Cisco Systems, Inc.
https://t.me/learningnets



Registration and resolution of EID prefixes is performed using Map Servers and Map Resolvers. The steps to establish the EID-to-RLOC database and the EID mapping resolution can be summarized as follows:

Step 1 The ETR devices register the EID prefixes to the configured MS in the form of EID-to-RLOC prefixes. One EID can be mapped to multiple RLOCs because EID prefixes can be achieved using multihomed addresses (RLOCs). There is no synchronization protocol between multiple Map Servers, and each EID must register to each MS independently.

Step 2 The ITR router sends Map-Request messages to get the RLOC address for the EID in the Map-Request message to the MS that is found in its configuration.

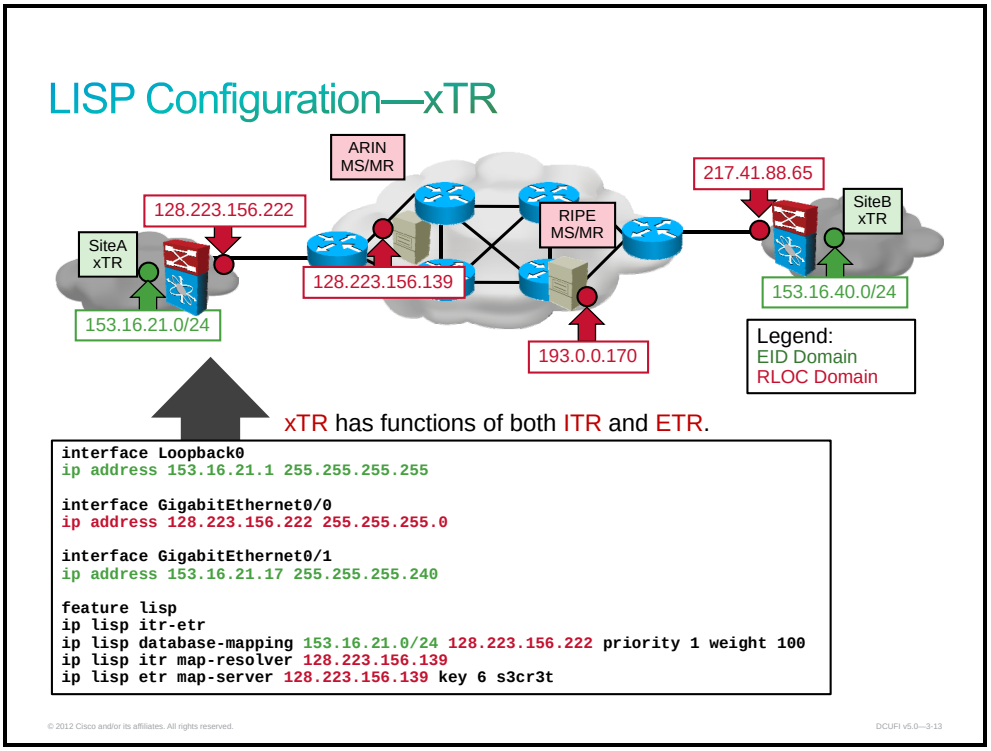
Step 3 The MS performs the lookup in its own database to see which RLOC address is mapped to the requested EID. When the entry is found, the MS forwards the Map-Request to the RLOC that is found in the database of the registered EID-to-RLOC mappings.

Step 4 When a Map-Request message is received on the ETR, it creates a Map-Reply message that contains EID-to-RLOC mappings that are associated to the destination EID site.

https://t.me/learningnets

Configuring LISP

This topic describes how to configure LISP in the Cisco Data Center Network Architecture.



CPE devices commonly implement both ITR and ETR functions. This type of CPE device is referred to as an xTR.

Before you configure specific configuration commands, you need to enable LISP functionality on the Cisco Nexus Operating System (NX-OS) using the **feature lisp** global configuration command. Also ensure that you are in the correct VDC.

In the example that is shown on the slide, Loopback0 and Fastethernet 0/1 are assigned an IP address that represents EID prefixes, and can be combined as the 153.16.21.0/24 EID prefix block. Since the Cisco Nexus 7000 Series switch will work as an ITR or ETR device, you must enable these two functionalities using the **ip lisp itr-etr** command. You can also use separate commands to enable only one of the functionalities: **ip lisp itr** for ITR and **ip lisp etr** for the ETR functionality.

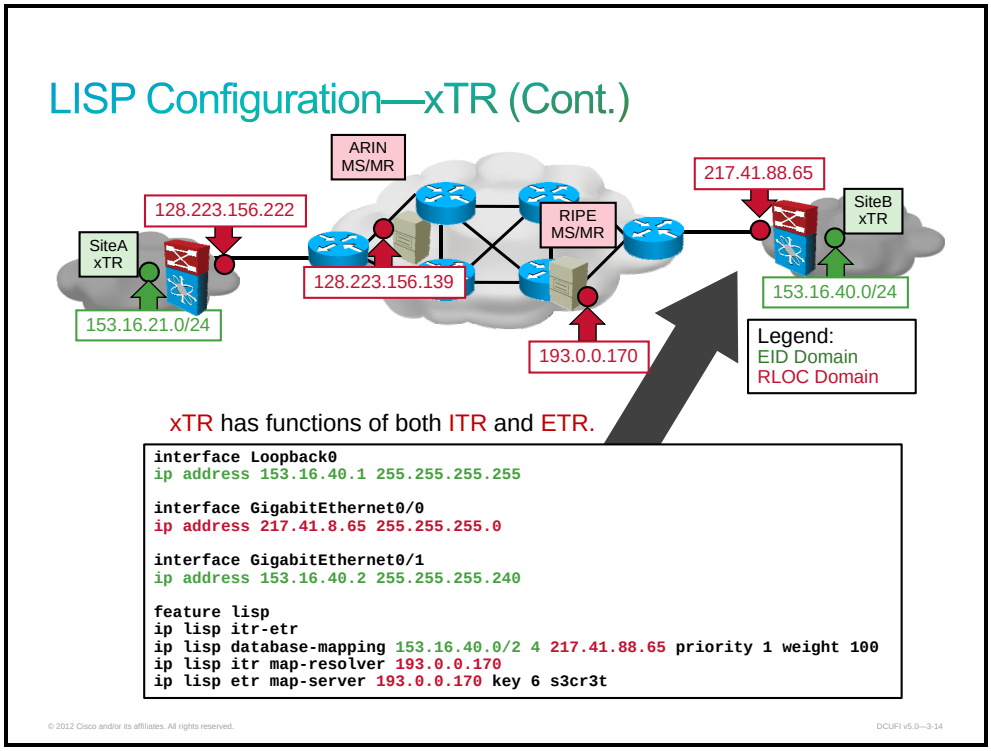
The EID-to-RLOC mapping relationship and associated traffic policy for all IPv4 or IPv6 EID prefixes for LISP SiteA is configured with **ip lisp database mapping**, where 153.16.21/0 represents the EID prefix, and 128.223.156.222 is actually the RLOC. One site can have multiple locators that are associated with the same prefix (in the case of multihoming), and using same command, you can configure all of the locators for that particular EID prefix. When a site has multiple EID prefixes, you need to configure all of the locators for each of the EID prefixes. If the site has multiple ETR devices, ensure that all options among the ETRs are consistent.

To configure the locator address of the MR to which this router sends Map-Request messages use the **ip lisp itr map-resolver** command in global configuration mode. The **ip lisp etr map-server** command configures the locator address of the LISP MS to which this device, acting as an IPv4 or IPv6 LISP ETR, registers.

For complete list of LISP configuration commands, see the *Cisco Nexus 7000 Series NX-OS LISP Command Reference* at http://www.cisco.com/en/US/docs/switches/datacenter/sw/6_x/nx-os/lisp/command/reference/preface.html.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

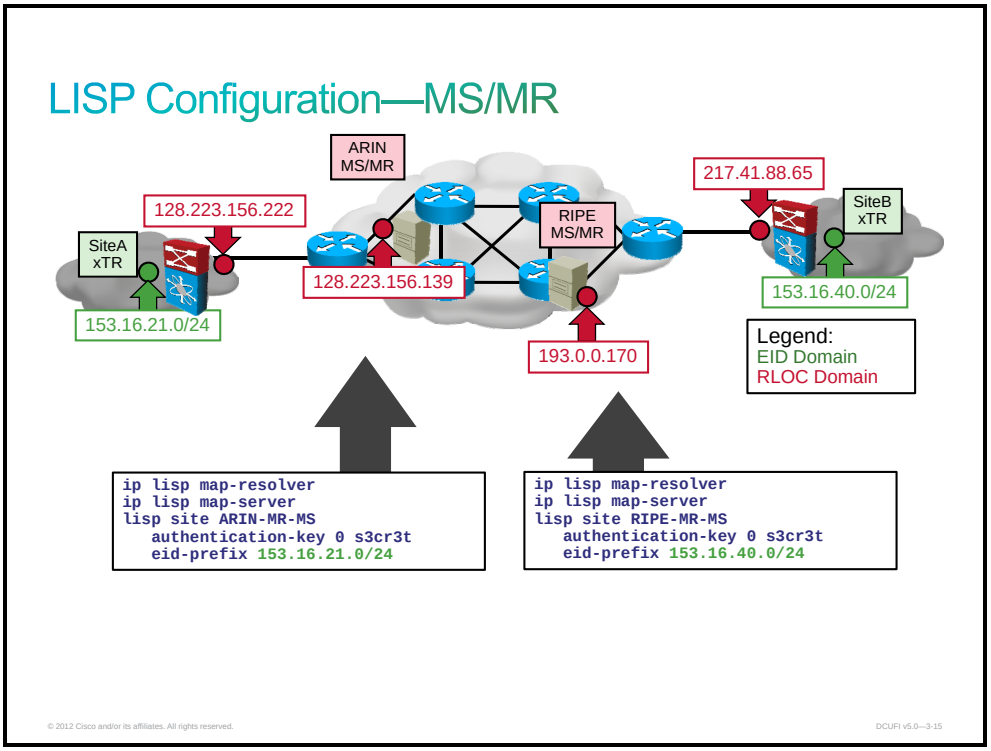


The configuration of SiteB is very similar to the configuration of SiteA. Also, Loopback0 and Fastethernet 0/1 are assigned IP addresses that represent EID prefixes, and in this case, can be combined as the 153.16.40.0/24 EID prefix block.

Every device that acts as an ETR must register to the MS that has predefined EID prefixes that can be registered. In this case, SiteB uses RIPE MS/MR for registration and resolution.

All other configuration is straightforward, and explained for SiteA.

https://t.me/learningnets

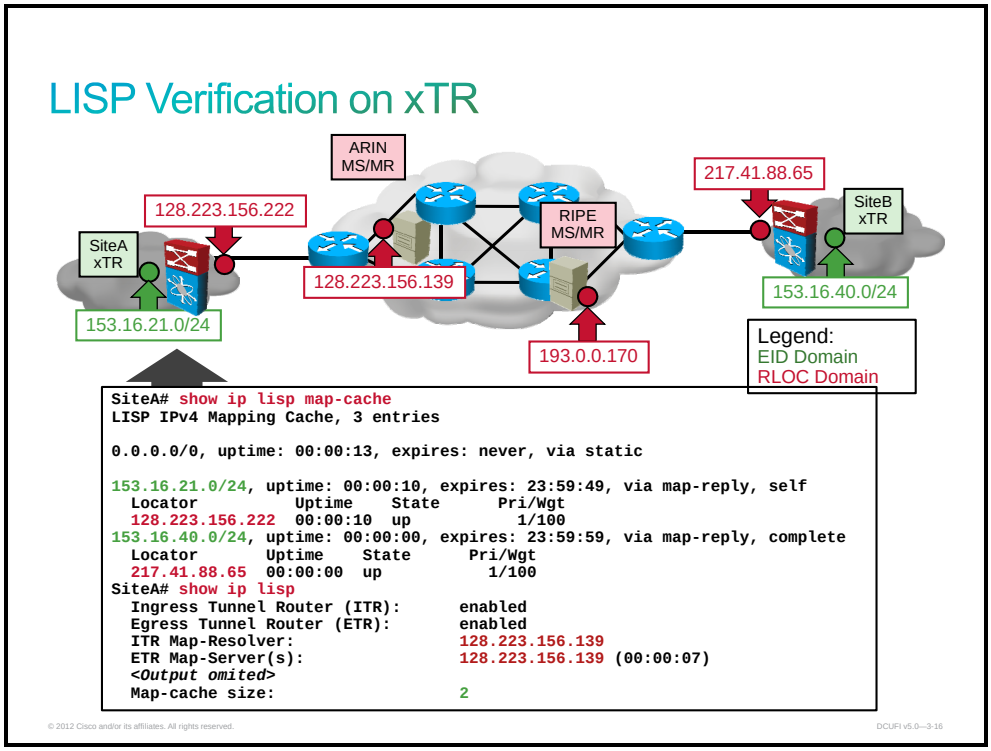


First of all, you need to enable LISP MS and MR functionalities using the **ip lisp map server** and **ip lisp map resolver** configuration commands respectively. For the MR functionality, this command is the only mandatory configuration command that is needed.

With the **lisp site** command, you are creating the LISP site name. Enter LISP configuration mode for defining site-specific parameters such as EID prefixes and keys. Using the **eid-prefix** command, authoritative EID prefixes for the RIPE site are configured. The authentication key must match on all ETRs that are registering to this site. In other words, the MS must be configured with EID prefixes that match the EID prefixes that are configured on the ETRs, and a key must match the one that is configured on all of the ETR that are registered to this MS.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets



When the ITR router completes a Map-Resolution request, the given EID-to-RLOC mapping is stored in the map cache. To see this map cache, use the **show ip lisp map-cache** command. This command also shows static mappings that are configured, in addition to dynamic mappings that are gathered through the Map-Resolution process. In the figure green and red indicated EID and RLOC respectively in the output.

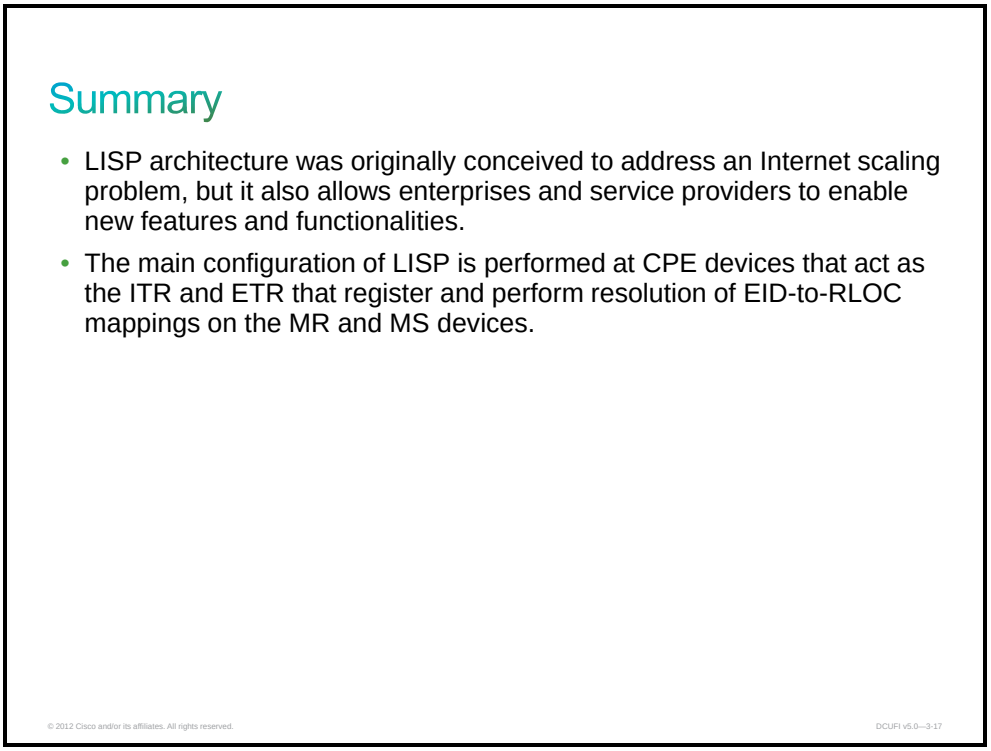
The **show ip lisp** command, displays all configured IPv4 or IPv6 LISP configuration parameters, including the function of the device (ITR, ETR or both), MS and MR addresses, size of map cache, and so on

For a complete list of the LISP verification commands on Cisco Nexus NX-OS Software, visit http://www.cisco.com and see the *Cisco Nexus 7000 Series NX-OS LISP Command Reference* at http://www.cisco.com/en/US/docs/switches/datacenter/sw/6_x/nx-os/lisp/command/reference/lisp_cr.html.

https://t.me/learningnets

Summary

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



© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Lesson 4

Configuring QoS

Overview

Converged IP networks must provide secure, predictable, measurable, and sometimes guaranteed services. Quality of service (QoS) provides network administrators and architects with a set of techniques that are used to manage network resources.

The moment an IP packet enters the network, it is classified and usually marked with its class identification. From that point on, the packet is treated by various QoS mechanisms according to the packet classification. Depending upon the mechanisms it encounters, the packet could be expedited, delayed, or even dropped.

This lesson describes QoS mechanisms, and how to properly apply QoS techniques in the Cisco Nexus Operating System (NX-OS) Software.

Objectives

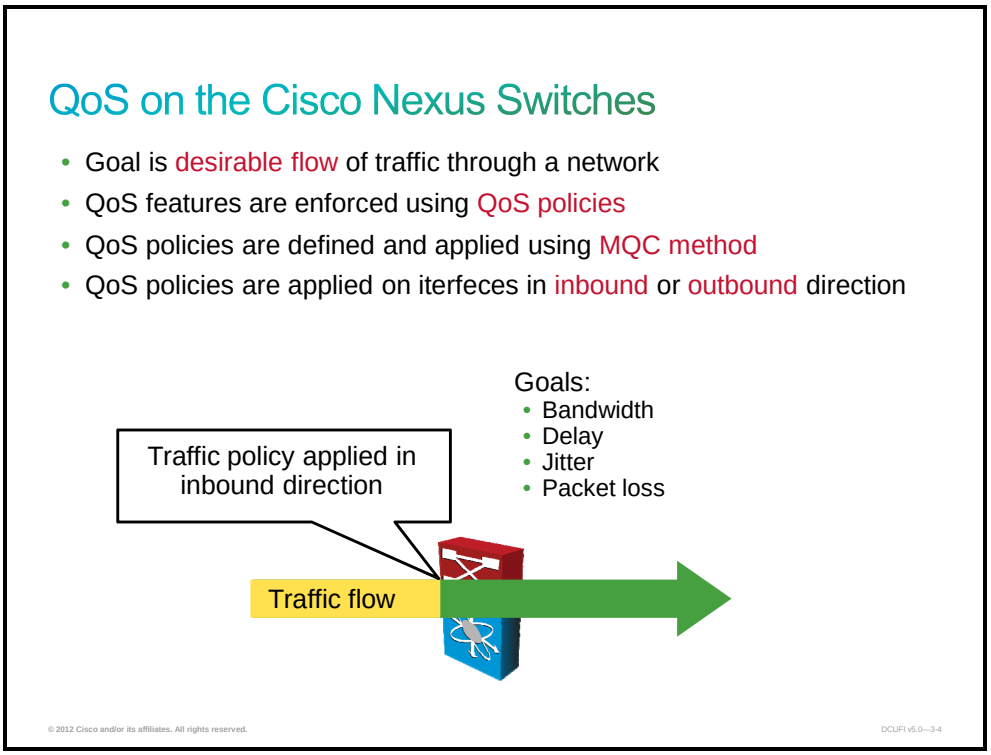
Upon completing this lesson, you will be able to configure QoS on the Cisco Nexus Family of switches based on a given requirement. This ability includes being able to meet these objectives:

- Describe the QoS features that are available on the Cisco Nexus Family of switches as well as the default QoS behavior
- Use the MQC for configuring the QoS features
- Configure QoS marking on the Cisco Nexus Family of switches
- Configure mutation mapping for QoS on the Cisco Nexus Family of switches
- Configure policing on the Cisco Nexus Family of switches
- Configure scheduling and queuing features on the Cisco Nexus Family of switches
- Monitor QoS statistics on the Cisco Nexus Family of switches

https://t.me/learningnets

QoS on the Cisco Nexus Family of Switches

This topic describes the QoS features that are available on the Cisco Nexus Family of switches as well as the default QoS behavior.



QoS features are used to provide the most desirable flow of traffic through a network. With QoS techniques, administrators can classify network traffic, police and prioritize the traffic flow, and provide mechanisms that can avoid occurrences of congestion.

Networks must provide these capabilities:

- Enough bandwidth for critical applications. In some cases, the network infrastructure must guarantee this bandwidth as well.
- Low and constant delay for delay sensitive applications. Important delay-sensitive packets must be sent first or before best effort traffic.
- Low or no packet loss. Some applications can adapt to packet loss by slowing down the transmit rate, some applications do not care, and some of them require no packet loss.

In storage applications such as the Small Computer Systems Interface (SCSI), I/O read and writes suffer from the loss of a single data frame. In the case of a lost Fibre Channel frame, the whole SCSI read or write Fibre Channel exchange will be restarted.

QoS provides classification of network traffic, policing and prioritization of traffic flows, and congestion avoidance that is based on fields in the packets that are flowing through the system. QoS on the Cisco Nexus 7000 Family of switches adheres to the Cisco Modular QoS CLI (MQC) configuration model for the creation of traffic classes and policies.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

QoS features in NX-OS software

Main categories of QoS features in NX-OS:

- Classification: Each class-oriented QoS mechanism has to support some type of classification
- Marking: Used to mark packets based on classification, metering, or both.
- Mutation: Changing of packet header QoS fields to all incoming or outgoing packets
- Policing: Used to enforce a rate limit by dropping or marking down packets
- Queuing and Scheduling: Queuing and congestion mechanisms allows to control the bandwidth allocated to traffic classes

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

Classification is a fundamental part of QoS mechanisms. Classification is the identifying and splitting of traffic into different classes. In a QoS-enabled network, all traffic is classified at the input interface of every QoS-aware device. It is essentially the process of partitioning of traffic that is based on different characteristics (source or destination IP addresses, VLAN, QoS fields in an IP or Ethernet header, and so on).When you define the traffic class, the next step is to choose match/ or match not criteria.

Marking, also known as coloring, involves marking each packet as a member of a network class so that devices throughout the rest of the network can quickly recognize the packet class. Marking is performed as close to the network edge as possible and is typically done using an MQC method that is described in the following topic.

Mutation is technique that is used to change the packet header QoS fields in every incoming or outgoing packet that goes through the particular interface. You can apply mutation mapping to the packet fields, class of service (CoS), differentiated services code point (DSCP), IP precedence, or to the internal field discard class.

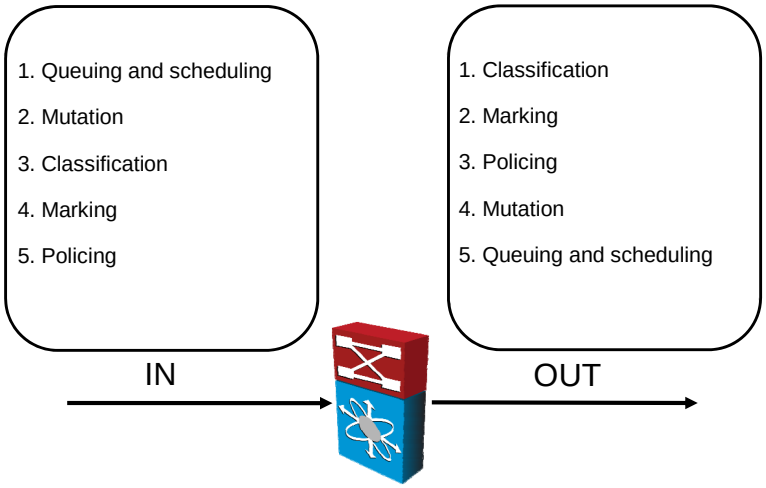
Traffic policing feature limits the input or output transmission rate of a class of traffic that is based on user-defined criteria, and it can mark packets by setting values such as IP precedence, QoS group, or DSCP value. Policing mechanisms can be set to first drop traffic classes that have lower QoS priority markings.

Queuing and scheduling processes allow you to control the bandwidth that is allocated to traffic classes. *Congestion management* (queuing algorithms) uses the marking on each packet to determine in which queue to place packets. Different queues are given different treatment by the queuing algorithm that is based on the class of packets in the queue. Generally, queues with high-priority packets receive preferential treatment. *Congestion avoidance* techniques monitor network traffic flows to anticipate and avoid congestion at common network and internetwork bottlenecks before problems occur. They are typically implemented on output interfaces wherever a high-speed link or set of links feeds into a lower-speed link.

https://t.me/learningnets

Sequencing of QoS actions on Nexus Switch

- Cisco NX-OS device processes the QoS policies based on whether they are applied to ingress or egress direction



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

The Cisco NX-OS switch processes the QoS policies that are defined on interfaces based on whether they are applied to the ingress or egress direction.

The sequence of QoS actions in the ingress direction of traffic is as follows:

Step 1 Queuing and scheduling

Step 2 Mutation

Step 3 Classification

Step 4 Marking

Step 5 Policing

The sequencing of QoS actions in the egress direction of traffic is as follows:

Step 1 Classification

Step 2 Marking

Step 3 Policing

Step 4 Mutation

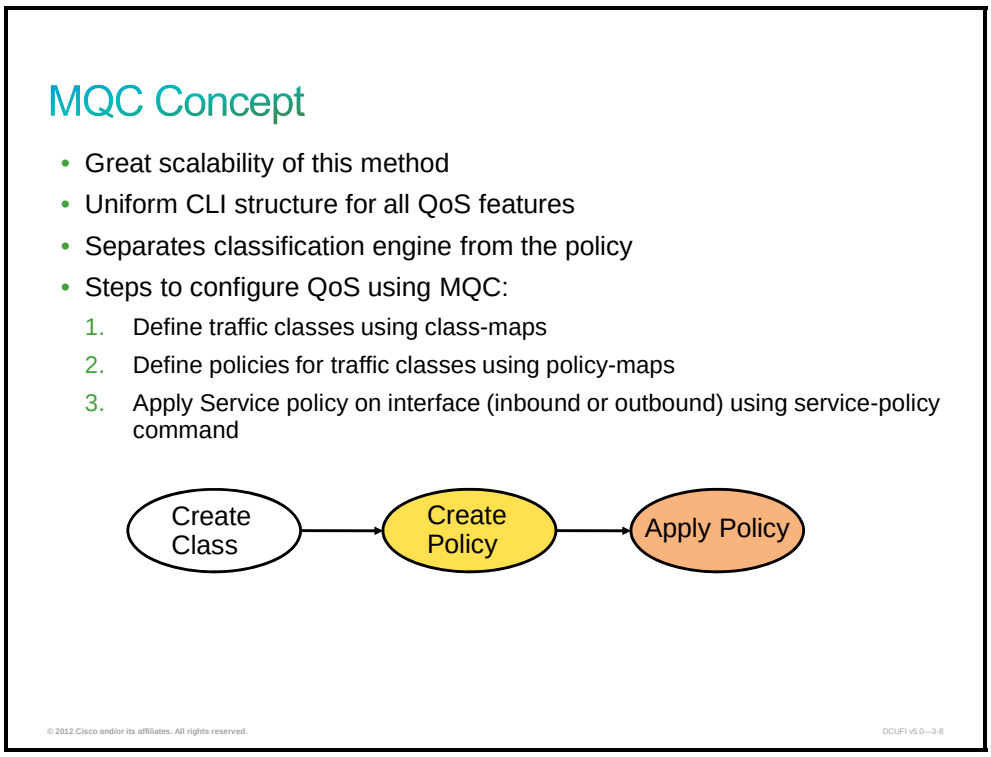
Step 5 Queuing and scheduling

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Modular QoS CLI

This topic explains how to use the MQC for configuring QoS features.



The MQC was introduced to allow any supported classification to be used with any QoS mechanism.

The separation of classification from the QoS mechanism allows new Cisco software versions to introduce new QoS mechanisms and reuse all available classification options. Alternatively, old QoS mechanisms can benefit from new classification options.

Another important benefit of the MQC is the reusability of configurations. MQC allows the same QoS policy to be applied to multiple interfaces. The MQC, therefore, is a consolidation of all the QoS mechanisms that have so far only been available as standalone mechanisms.

Implementing QoS by using the MQC consists of three steps:

Step 1 Configure classification by using the **class-map** command.

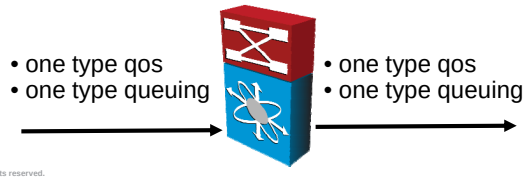
Step 2 Configure traffic policy by associating the traffic class with one or more QoS features using the **policy-map** command.

Step 3 Attach the traffic policy to inbound or outbound traffic on interfaces by using the **service-policy** command.

https://t.me/learningnets

MQC Command and Object Types

- MQC on Nexus devices provide three command types to define traffic classes and policies:
  - class-map – class of traffic based on match criteria
  - table-map – define one set of packet field values to another set of field values
  - policy-map – defines set of policies to be applied on class-maps
- Class-map and Policy-map object types:
  - network qos - defines CoS properties across switches and VDCs
  - qos – used for marking, mutations, ingress port trust state and policing
  - queuing – MQC objects used marking, queuing and shaping
- One policy map of each type can be applied in one direction:



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

MQC on Cisco NX-OS Software provides three command types that are used to define traffic classes and policies:

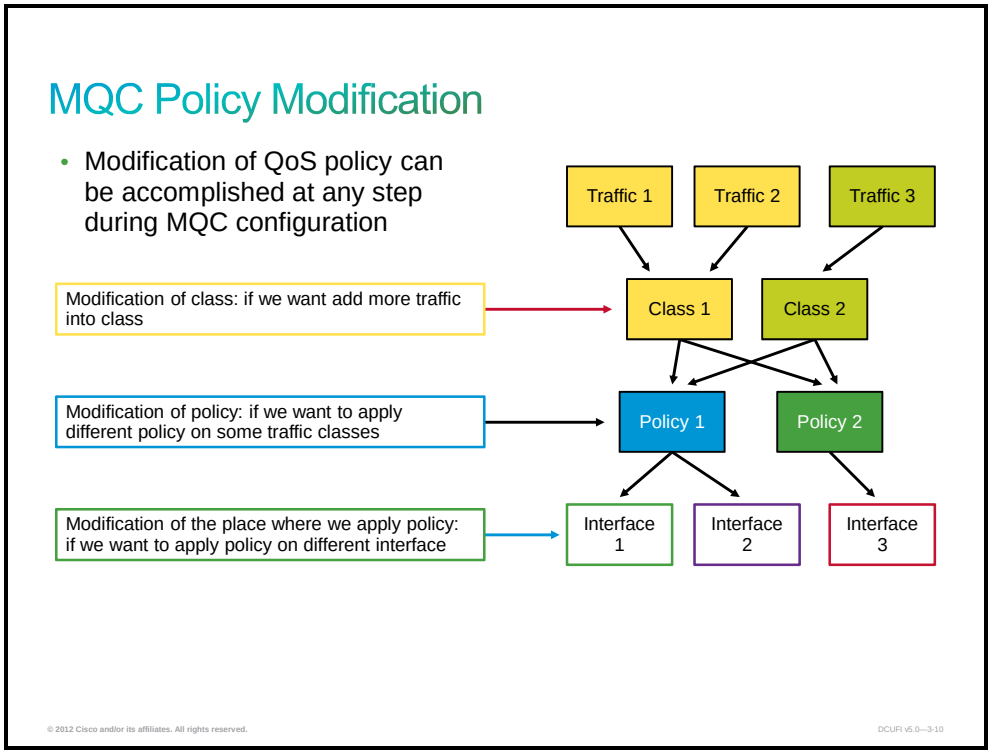
- **class-map:** Used for creation of traffic classes using match criteria.
- **table-map:** Used to define mapping rules to set one type of packet field values to another set of packet field values. The table maps are referenced in policy maps, and cannot be used in queuing policies.
- **policy-map:** Defines a set of rules that are applied on individual classes that are defined under a particular policy map. Policies can be attached to ports, port channels, VLANs, subinterfaces or tunnels.

In addition these policy map and class map types can be created:

- **network qos:** Defines CoS properties across switches and virtual device contexts (VDCs)
- **qos:** Used for marking, mutations, ingress port trust state, and policing (This type is also the default policy map and class map type.)
- **queuing:** This type of MQC object is used for queuing and scheduling

QoS policy applied on network interfaces can be queuing policy or QoS policy. When a packet is passing through the interface that has a QoS policy applied, it is matched sequentially to a class of traffic starting from the first traffic class definition. When a match is found, the policy actions for that class are applied to the packet. If there is no match against all configured classes, the default class-default traffic class is matched. This class can be used to perform mutations.

https://t.me/learningnets



Modification of the class: QoS traffic class can be modified in different ways. You can create a new class and edit existing classes by adding more traffic into it or by adding conditional matching statements on existing classes, or remove classes that are no longer being used by any policy.

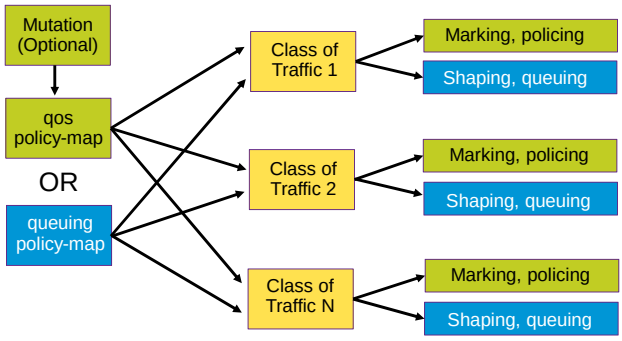
Modification of policy: Policy can be modified in many ways. You can apply a different policy on the traffic class, apply a child policy, or simply add new per-hop behavior (PHB) to the existing traffic class. A change of policy is immediately reflected on the traffic that is passing through the interface.

Modification of attachment point or direction: The same policy can be applied on multiple interfaces. You can disable or enable policy on an interface by entering one command.

https://t.me/learningnets

Type QoS and Queuing Policies

- Two types of policy maps: qos and queuing
- Mutiation map can be applied on qos policy map
- Individual actions are applied on traffic class basis



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

In the Cisco NX-OS Software, there are two types of policy maps: qos and queuing.

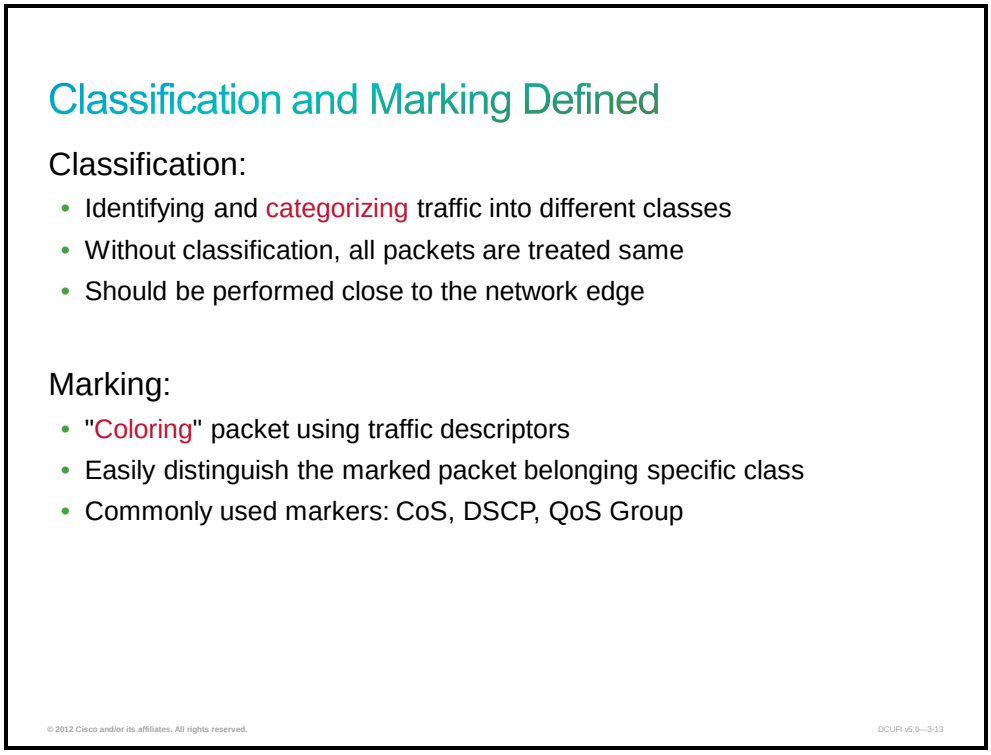
The type qos policy can be used to mark packets, apply mutations, set the ingress port to trust state, or even police packets. Type qos policies and actions on traffic classes in this type of policy are shown on the diagram in the figure using green color. On these types of policies, mutation can be also applied.

The type queuing policies can be used for marking, queuing, and shaping. Type queuing policies and actions on traffic classes in this type of policies are shown on the diagram in blue color.

https://t.me/learningnets

Marking

This topic describes how to configure QoS Marking on the Cisco Nexus Family of switches.

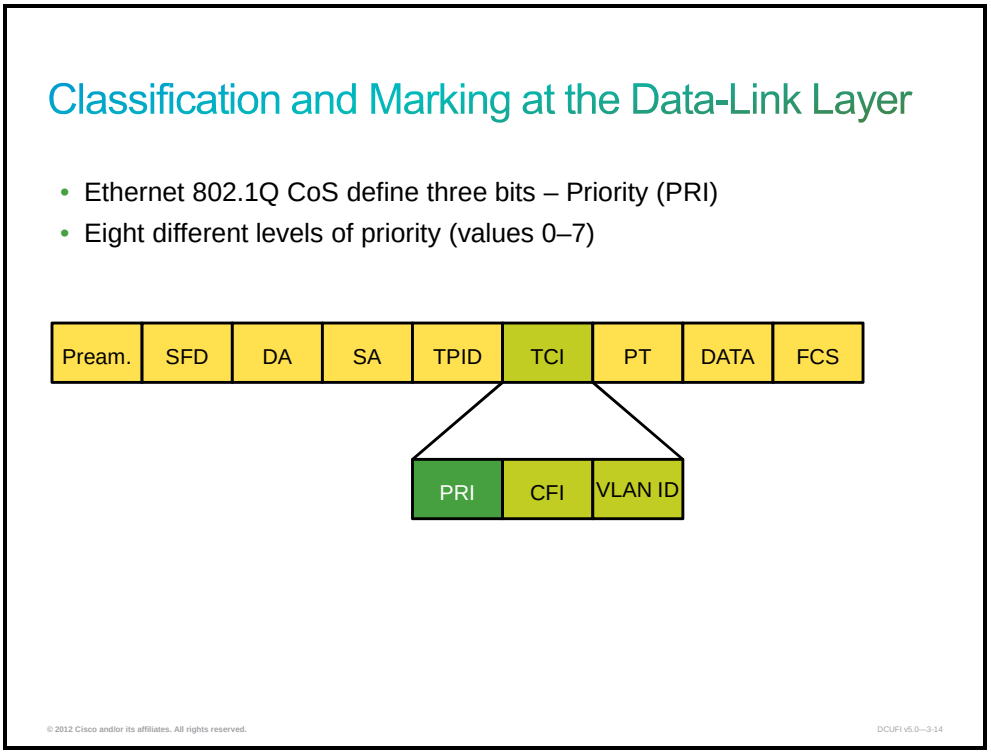


Before any QoS applications or mechanisms can be applied, traffic must be identified and sorted into different classes. QoS is applied to these different traffic classes. Network devices use classification to identify traffic as belonging to a specific class. After network traffic is sorted, marking can be used to color (tag) individual packets so that other network devices can apply QoS features uniformly to those packets as they travel through the network.

Marking allows network devices to classify a packet or frame based on a specific traffic descriptor. Typically, the traffic descriptors that are used include CoS, DSCP, or IP precedence bits. Marking can be used to set information in the Layer 2 or Layer 3 packet headers.

Marking a packet or frame with its classification allows network devices to easily distinguish the marked packet or frame as belonging to a specific class. After the packets or frames are identified as belonging to a specific class, QoS mechanisms can be uniformly applied to ensure compliance with administrative QoS policies.

https://t.me/learningnets



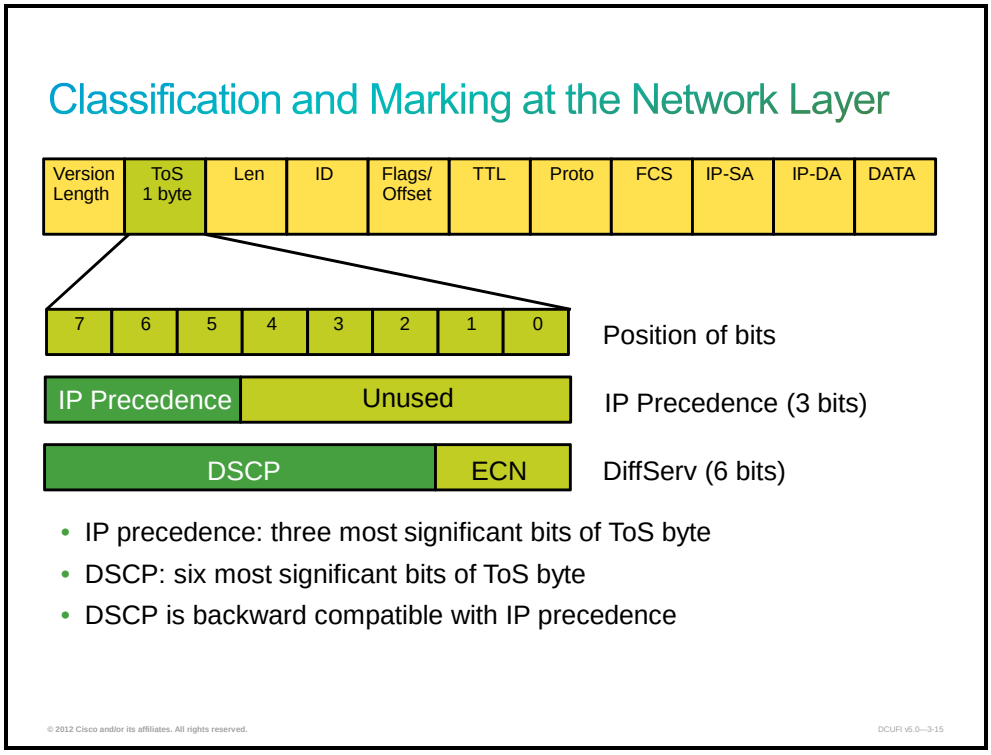
The 802.1Q standard is an IEEE specification for implementing VLANs in Layer 2 switched networks. The 802.1Q specification defines two 2-byte fields, Tag Protocol Identifier (TPID) and Tag Control Information (TCI), which are inserted within an Ethernet frame following the source address field. The TPID field is currently fixed and assigned the value 0x8100. The TCI field is composed of these three fields, of which the following field is of interest when you implement QoS at Layer 2.

User priority bits (3 bits): These bits can be used to mark packets as belonging to a specific CoS. The CoS marking uses the 3 IEEE 802.1p user priority bits and allows a Layer 2 Ethernet frame to be marked with eight different levels of priority (values 0–7). These 3 bits allow for eight levels of classification, allowing a direct correspondence with the IP version 4 (IPv4) (IP precedence) type of service (ToS) values. The 802.1p specification defines these standard definitions for each CoS:

- CoS 7 (111): network
- CoS 6 (110): Internet
- CoS 5 (101): critical
- CoS 4 (100): flash-override
- CoS 3 (011): flash
- CoS 2 (010): immediate
- CoS 1 (001): priority
- CoS 0 (000): routine

One disadvantage of using CoS markings is that frames will lose their CoS markings when they are transiting a non-802.1Q or non-802.1p link. This problem includes any type of non-Ethernet WAN link. Therefore, a more permanent marking should be used for network transit, such as Layer 3 IP DSCP marking. This marking is typically accomplished by translating a CoS marking into another marker or simply using a different marking mechanism.

© 2012 Cisco Systems, Inc.
https://t.me/learningnets



At the network layer, IP packets are typically classified based on source or destination IP address, or the contents of the ToS byte. Data link layer media often changes as a packet travels from its source to its destination. Because a CoS field does not exist in a standard Ethernet frame, CoS markings at the data link layer are not preserved as packets traverse nontrunked or non-Ethernet networks. Using marking at the network layer (Layer 3) provides a more permanent marker that is preserved from source to destination.

IP Precedence

Originally, only the first 3 bits of the ToS byte were used for marking, referred to as IP precedence. However, newer standards have made the use of IP precedence obsolete in favor of using the first 6 bits of the ToS byte for marking, referred to as DSCP.

The header of an IPv4 packet contains the ToS byte. IP precedence uses three precedence bits in the ToS field of the IPv4 header to specify CoS for each packet. IP precedence values range from 0 to 7 and allow you to partition traffic in up to six useable classes of service. (Settings 6 and 7 are reserved for internal network use.)

DiffServ

Differentiated services (DiffServ) are a new model that supersedes—and is backward-compatible with—IP precedence. DiffServ redefines the ToS byte as the DiffServ field and uses six prioritization bits that permit classification of up to 64 values (0 to 63), of which 32 are commonly used. A DiffServ value is called a DSCP.

With DiffServ, packet classification is used to categorize network traffic into multiple priority levels or classes of service. Packet classification uses the DSCP traffic descriptor to categorize a packet within a specific group to define that packet. After the packet has been defined (classified), the packet is then accessible for QoS processing on the network.

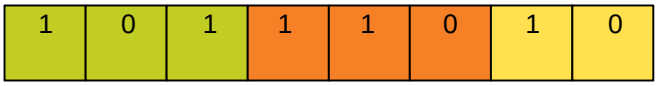
.

https://t.me/learningnets

Per-Hop Behaviour

Per hop behavior	Value	Service
Default	000 XXX	Best effort
Expedited Forwarding (EF)	101 110	Low delay
Assured Forwarding	XXX XX0	Guaranteed bandwidth

Example of ToS byte (EF):



- IP Precedence field = 101 (bits 5-7)
- DSCP field = 101110 (bits 2-7)
- ECN field = 10 (bits 0-1)

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

These PHBs are defined by IETF standards:

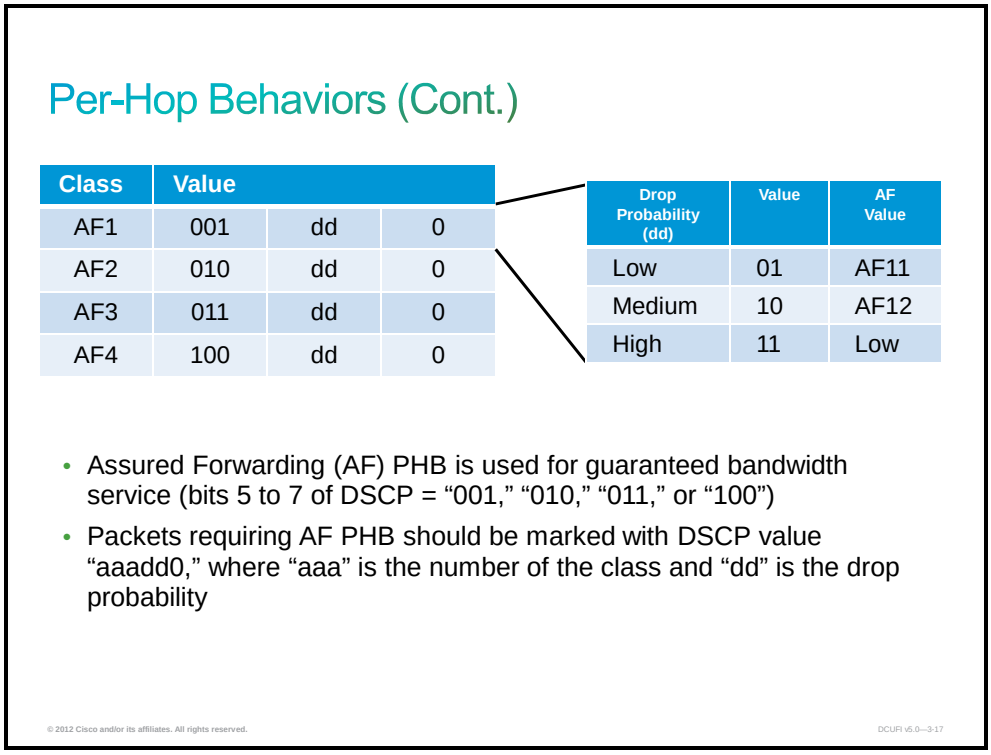
- **Default PHB:** Used for best-effort service (bits 5 to 7 of DSCP = “000”)
- **Expedited Forwarding (EF) PHB:** Used for low-delay service (bits 2 to 7 of DSCP = “101110”)
- **Assured Forwarding (AF) PHB:** Used for guaranteed bandwidth service (bits 5 to 7 of DSCP = “001,” “010,” “011,” or “100”)
- **Class-Selector PHB:** Used for backward compatibility with non-DiffServ-compliant devices (RFC 1812-compliant devices [bits 2 to 4 of DSCP = “000”])

For example, if the ToS byte equals 10111010, then the IP precedence field is 101, the DSCP field is 101110, and the ECN field is 10. DSCP 101110 is recommended for the EF PHB.

The EF PHB is identified based on the following information:

- The EF PHB ensures a minimum departure rate. The EF PHB provides the lowest possible delay to delay-sensitive applications.
- The EF PHB guarantees bandwidth. The EF PHB prevents starvation of the application if there are multiple applications using EF PHB.
- The EF PHB polices bandwidth when congestion occurs. The EF PHB prevents starvation of other applications or classes that are not using this PHB.
- Packets requiring EF should be marked with DSCP binary value “101110” (46 or 0x2E).

https://t.me/learningnets



The AF PHB is identified based on the following information:

- The AF PHB guarantees a certain amount of bandwidth to an AF class.
- The AF PHB allows access to extra bandwidth, if available.
- Packets requiring AF PHB should be marked with DSCP value “aaadd0,” where “aaa” is the number of the class and “dd” is the drop probability.

There are four standard defined AF classes. Each class should be treated independently and should have allocated bandwidth that is based on the QoS policy. Each AF class is assigned an IP precedence and has three drop probabilities: low, medium, and high. The AF number is created as follows:

AF*xy*: Assured Forwarding (RFC 2597), where *x* corresponds to the IP precedence value (only 1–4 are used for AF classes), and *y* corresponds to the Drop Preference value (either 1, 2, or 3)

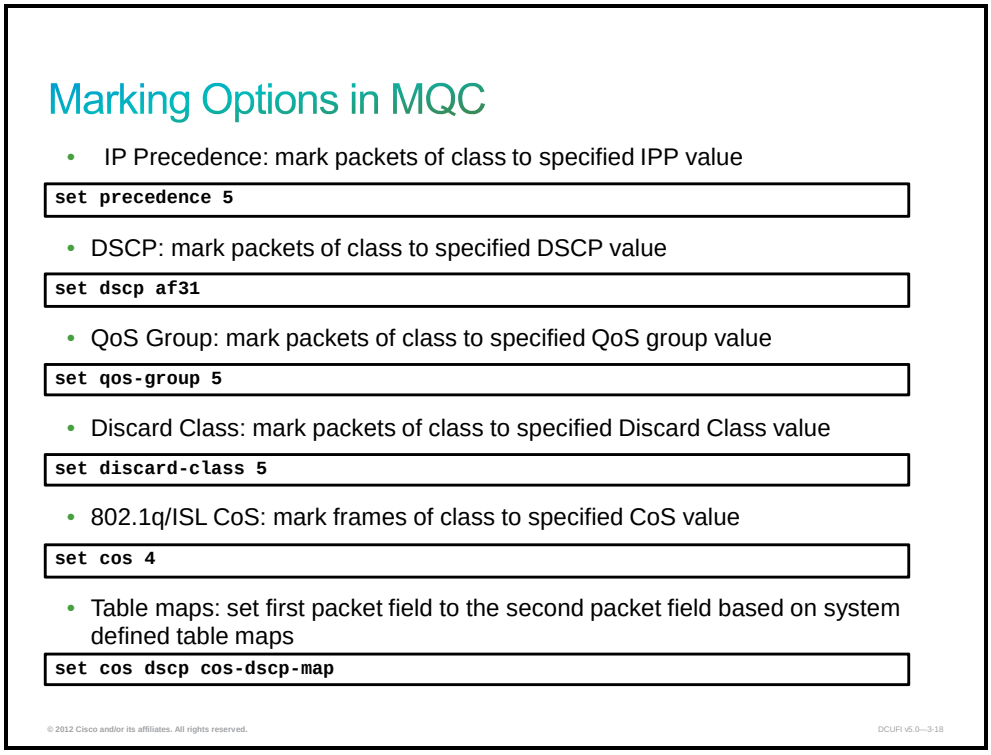
The following table map shows the binary and decimal representation of DSCP and shows the IP precedence values and PHB for all DSCP values.

DSCP (Binary)	DSCP (Decimal)	IP Precedence	Per-Hop Behavior
000000	0	0	Default (Class Selector 0)
001000	8	1	Class Selector 1 (CS1)
001010	10	1	AF11
001100	12	1	AF12
001110	14	1	AF13
010000	16	2	Class Selector 2 (CS2)
010010	18	2	AF21
010100	20	2	AF22

https://t.me/learningnets

DSCP (Binary)	DSCP (Decimal)	IP Precedence	Per-Hop Behavior
010110	22	2	AF23
011000	24	3	Class Selector 3 (CS3)
011010	26	3	AF31
011100	28	3	AF32
011110	30	3	AF33
100000	32	4	Class Selector 4 (CS4)
100010	34	4	AF41
100100	36	4	AF42
100110	38	4	AF43
101000	40	5	Class Selector 5 (CS5)
101110	46	5	Expedited Forwarding (EF)
110000	48	6	Class Selector 6 (CS6)
111000	56	7	Class Selector 7 (CS7)

© 2012 Cisco Systems, Inc.
https://t.me/learningnets



IP Precedence

To set the precedence value in the packet header, use the **set precedence** command in policy-map class configuration mode. Syntax of this command is as follows:

set precedence [**tunnel**] *precedence-value*

The optional word **tunnel s**ets the IP precedence value in the tunnel encapsulation in this and in the next commands.

DSCP

The **set dscp** command cannot be used with the **set precedence** command to mark the *same* packet. The two values, DSCP and precedence, are mutually exclusive. A packet can have one value or the other, but not both.

To mark a packet by setting the DSCP value in the ToS byte, use the **set dscp** command in QoS policy-map class configuration mode.

set dscp [tunnel] *dscp-value*

QoS Group

To set a QoS group identifier (ID) that can be used later to classify packets, use the **set qos-group** command in policy-map class configuration mode. Syntax of this command is as follows:

setqos-group *value*

Discard class

A discard-class value has no mathematical significance. For example, the discard-class value 2 is not greater than 1. The value simply indicates that a packet marked with discard-class 2 should be treated differently than a packet marked with discard-class 1.

Syntax of this command is as follows:

setdiscard-class *value*

https://t.me/learningnets

CoS

To set the Layer 2 CoS value of an outgoing packet, use the **set cos** command in policy-map class configuration mode.

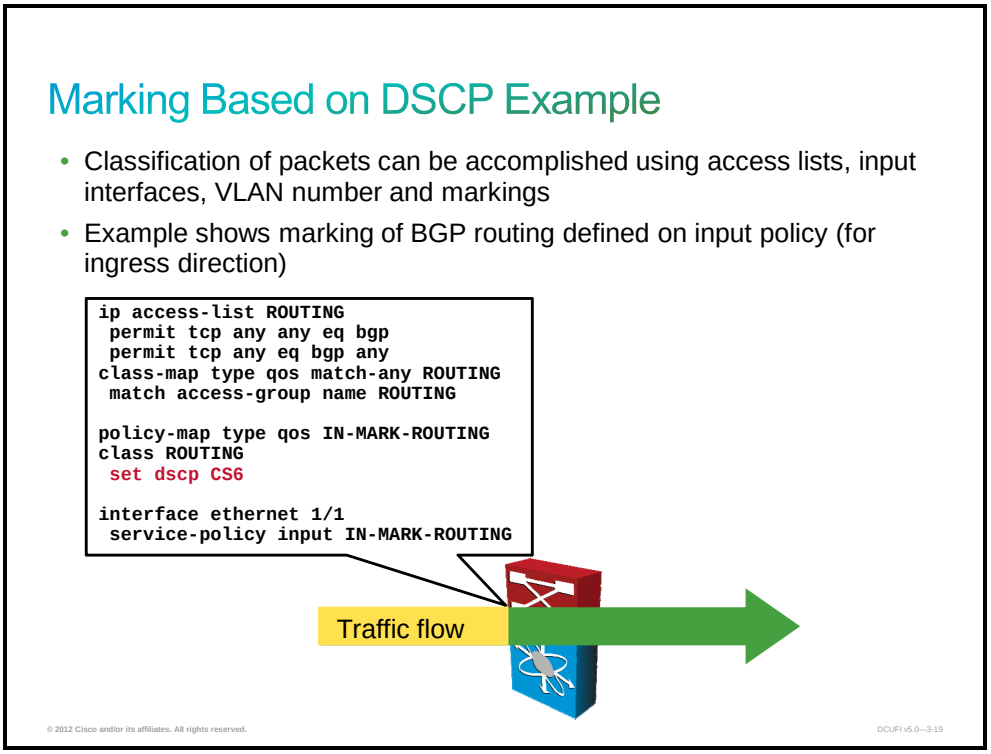
setcos *cos-value*

Table Map

Table map is used to set the first packet field to the second packet field based on system-defined table maps. The syntax of this command is as follows:

set *header-parameter* {*same-header-parameter* | *output-header-parameter*} **table** {*table-map-name* | *mutation-map*}.

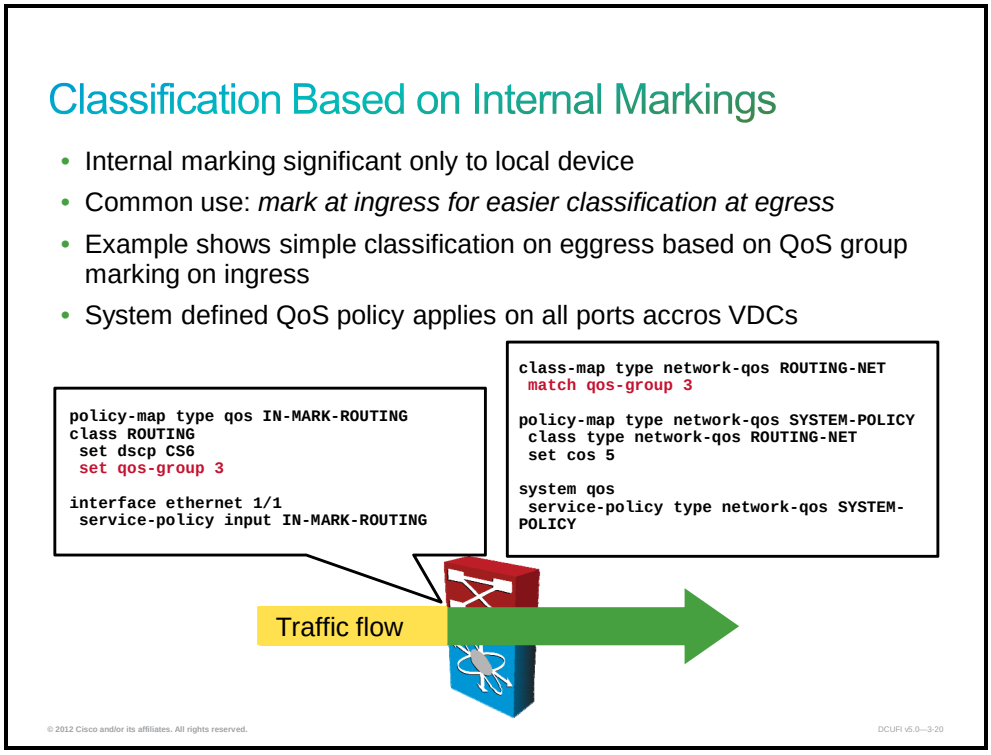
The *header-parameter* argument can be **cos, dscp, precedence** or **discard-class**. You can use a user-defined table map name or a system-defined table map name (*table-map-name* or *mutation-map* respectively).



This example shows how to mark the ROUTING class of traffic with a dscp value of CS6. This policy is applied in the input direction.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets



The **match qos-group** command is used by the class map to identify a specific QoS group value marking on a packet.

The *qos-group-value* argument is used as a marking only. The QoS group values have no mathematical significance. For instance, the *qos-group-value* of 2 is not greater than 1. The value simply indicates that a packet marked with the *qos-group-value* of 2 is different than a packet marked with the *qos-group-value* of 1. The treatment of these packets is defined by the user through the setting of QoS policies in QoS policy-map class configuration mode.

The QoS group value is local to the router, which means that the QoS group value that is marked on a packet does not leave the router when the packet leaves the router. If you need a marking that resides in the packet, use the IP precedence setting, IP DSCP setting, or another method of packet marking.

To identify a specific QoS group value as a match criterion, use the **match qos-group** command in class-map configuration mode. To remove a specific QoS group value from a class map, use the **no** form of this command.

match qos-group *qos-group-value*

Discard Class

A discard-class value has no mathematical significance. For example, the discard-class value 2 is not greater than 1. The value simply indicates that a packet marked with discard-class 2 should be treated differently than a packet marked with discard-class 1.

Packets that match the specified discard-class value are treated differently from packets marked with other discard-class values. The discard-class is a matching criterion only, used in defining PHB for dropping traffic.

https://t.me/learningnets

Verifying Packet Classification

- Verify packet marking by using command:
  - show table-map
  - show policy-map
- Verifying a policy-map applied on an interface:

```
switch(config)# show policy-map interface ethernet 1/1
Global statistics status: enabled
Ethernet1/1
Service-policy (qos) input: IN-MARK-ROUTING
policy statistics status: enabled
Class-map (qos): map (match-all)
0 packets, 0 bytes 5
 minute offered rate 0 bps, drop rate 0 bps
Match: cos 0
! <output omited…>
```

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

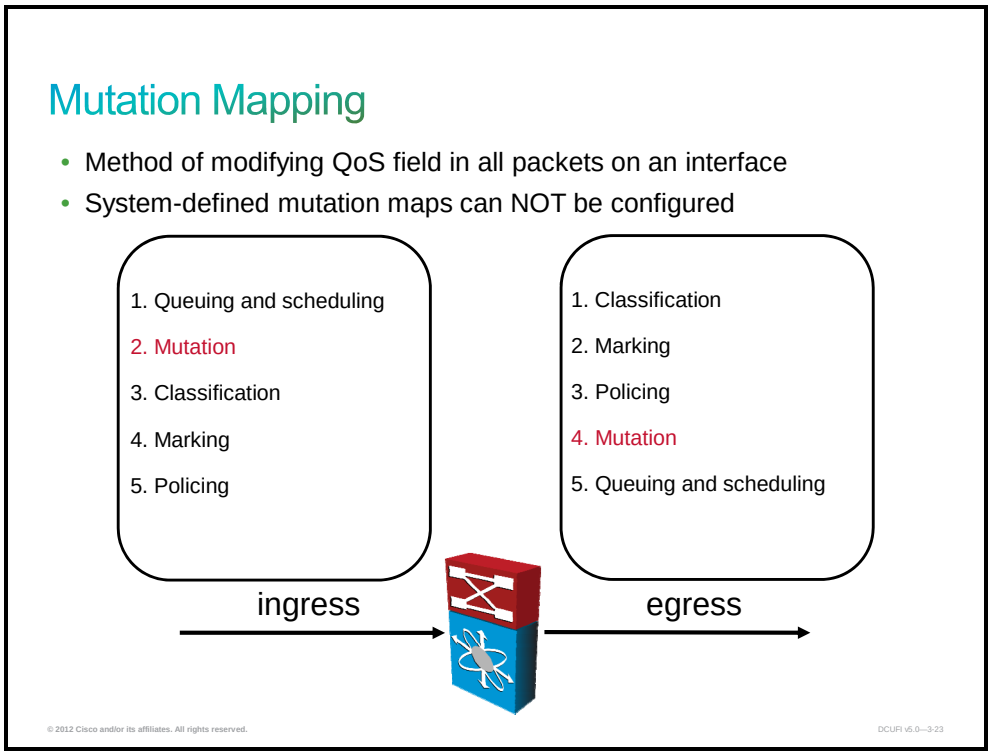
To verify packet classification, use one of the **show table-map** or **show policy-map interface** commands.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Mutation Mapping

This topic describes how to configure mutation mapping for QoS on the Cisco Nexus Family of switches.



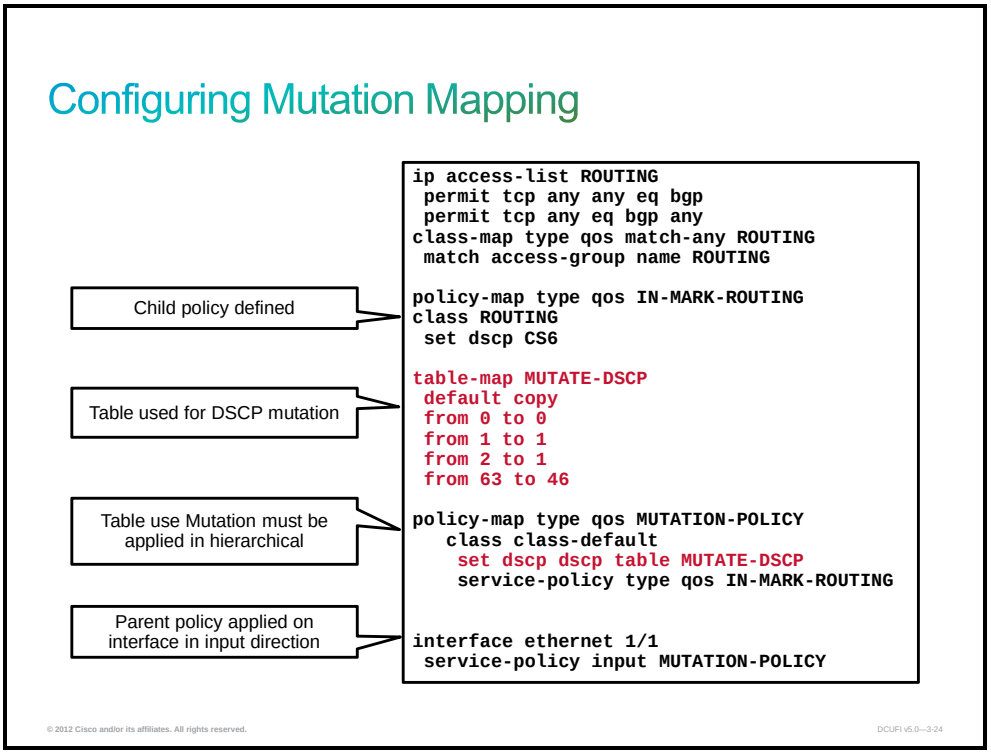
Mutation is the changing of packet header QoS fields on an interface. You can map IP precedence, DSCP, or CoS values to all incoming or outgoing packets. You can use mutation in policies that contain policing commands, but you cannot use mutation in queuing and scheduling commands. You use configurable, user-defined table maps for mutation.

On ingress, mutation mapping occurs before traffic classification and all other actions. On egress, mutation mapping occurs after traffic classification and before the other actions. You can apply mutation mapping to the CoS, DSCP, or IP precedence packet fields, or to the internal field discard class.

Mutation mapping has the following configuration guidelines and limitations:

- You use a hierarchical policy for mutation mapping. Hierarchical policies are not supported for any other use.
- The device supports only one level of hierarchy.
- You can configure up to 14 table maps for use in ingress interfaces and up to 15 table maps for use in egress interfaces.
- Before you delete a referenced policy map, you must first remove all references to that policy map.
- You can use only like parameters (for example, cos-cos) when you create a mutation map. Mutation maps with dissimilar types (for example, cos-dscp) are not supported.

https://t.me/learningnets



You use a hierarchical policy map to configure mutation mapping. In the mutation mapping policy map, you specify the field to mutate and the policy map to apply with the mutation. Next, only like parameters (for example, cos-cos) are permitted when you create a mutation map. Mutation maps with different types (for example, cos-dscp) are not supported.

In the example, child policy IN-MARK-ROUTING is defined to mark routing packets to a dscp value of CS6. Table map MUTATE-DSCP is used to perform a “replace DSCP fields of all packets on the interface.” If the value in the mutation map is not specified, the default behavior of the mapping is to copy from original to the modified (keywords *default copy*). The mutation is then applied on the parent policy, and child is applied on the all traffic that is passing through interface Ethernet 1/1 (keyword *class-default*).

.

© 2012 Cisco Systems, Inc.
https://t.me/learningnets

Verifying Mutation Mapping

- Verify configuration of mutation mapping by displaying policy map information:

```
N7K2-POD4# sh policy-map interface ethernet 4/5 input


Global statistics status :   enabled

Ethernet4/5

  Service-policy (qos) input:   MUTATION-POLICY

    Class-map (qos):   class-default (match-any)
      set dscp dscp table MUTATE-DSCP
      service-policy type qos IN-MARK-ROUTING
          Service-policy (qos) input:   IN-MARK-ROUTING

            Class-map (qos):   ROUTING (match-any)
      Match: access-group ROUTING
      set dscp cs6
```

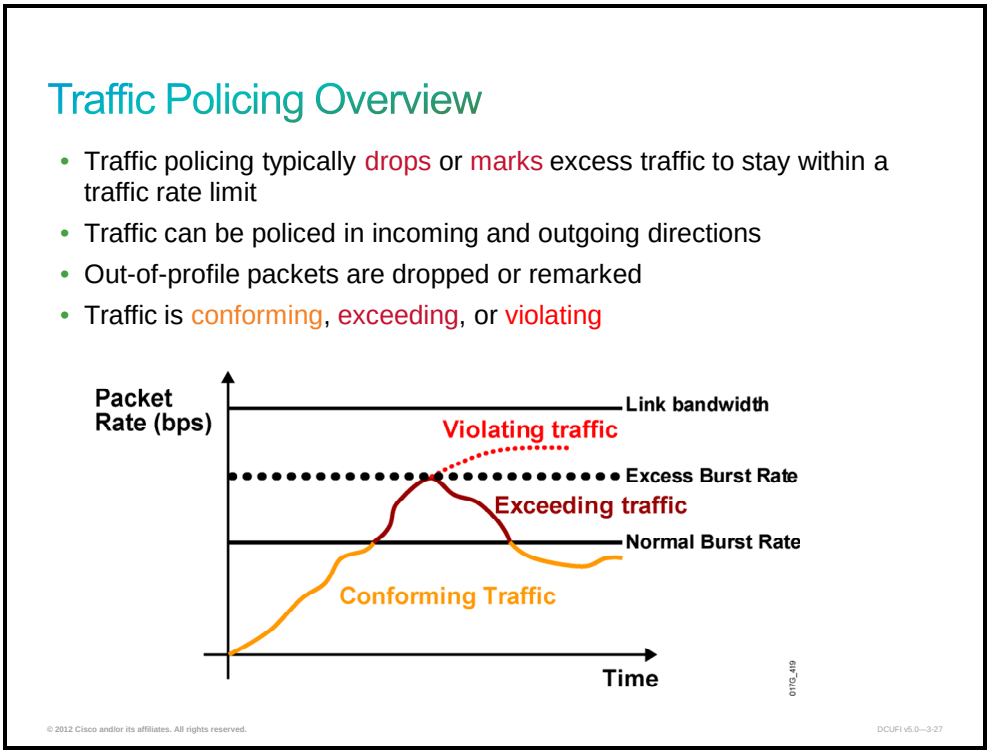
© 2012 Cisco and/or its affiliates. All rights reserved.

You can verify configuration of mutation mapping by displaying policy map configuration information on the particular interface.

https://t.me/learningnets

Policing

This topic describes how to configure policing on the Cisco Nexus Family of switches.



Traffic policing is often configured on interfaces at the edge of a network to limit traffic into or out of the network. Traffic policing drops or remarks excess traffic to control traffic flow within specified rate limits. Traffic policing does not introduce any delay to traffic that conforms to traffic policies. Traffic policing can cause more TCP retransmissions, because traffic in excess of the specified limit may be dropped. Traffic-policing mechanisms such as class-based policing have marking capabilities in addition to rate-limiting capabilities. Instead of dropping the excess traffic, traffic policing can alternatively mark and then send the excess traffic. This method allows the excess traffic to be re-marked with a lower priority before the excess traffic is sent out.

In the context of traffic policing, you can use three general levels of the traffic policing packet rate:

- Conforming traffic: This is a general term for traffic (or a class of traffic) that is in the normal defined traffic profile. Typically (but not always) this type of traffic is transmitted further without any other intervention.
- Exceeding traffic: This is a general term that is used for traffic (or a class of traffic) that is above the normal defined traffic profile but under the top allowed packet rate. Typically (but not limited) this type of traffic is remarked with lower priority QoS markings due to the exceeding nature of such traffic flows.
- Violating traffic: This is a general term that is used for traffic (or a class of traffic) that is above the top allowed traffic rate. Typically (but not always) this type of traffic is dropped by the policy.

© 2012 Cisco Systems, Inc.
https://t.me/learningnets

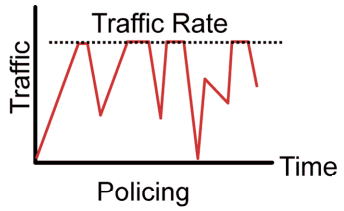
Policing vs. Shaping

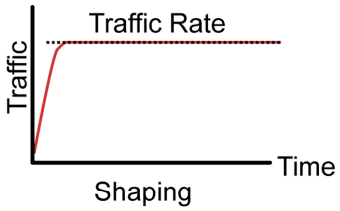
Policing:

- Incoming and outgoing directions
- Out-of-profile packets are dropped
- Dropping causes TCP retransmits
- Supports packet marking or re-marking
- Less buffer usage (shaping requires an additional shaping queuing system)

Shaping:

- Outgoing direction only
- Out-of-profile packets are queued until a buffer gets full
- Buffering minimizes TCP retransmits
- Marking or re-marking not supported
- Shaping supports interaction with Frame Relay congestion indication





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

Traffic policing and traffic shaping differ in the way they respond to traffic violations. Policing typically drops traffic, while shaping typically queues excess traffic.

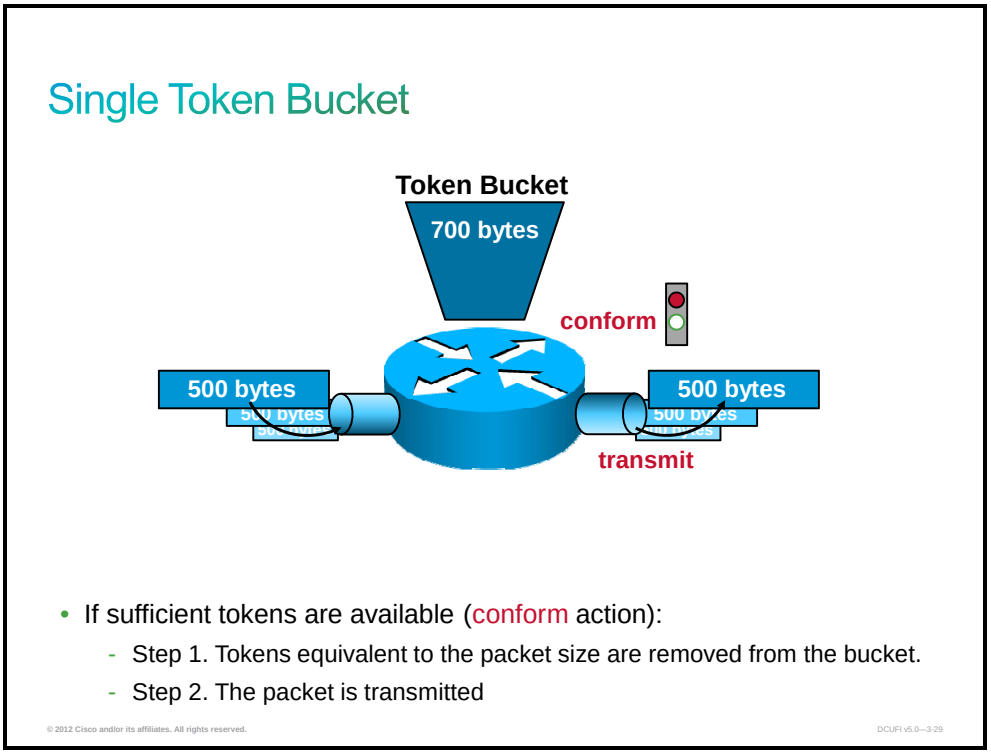
Shaping queues excess traffic by holding packets inside a shaping queue. Use traffic shaping to shape the outbound traffic flow when the outbound traffic rate is higher than a configured shape rate. Traffic shaping smooths traffic by storing traffic above the configured rate in a shaping queue. Therefore, shaping increases buffer utilization on a router and causes unpredictable packet delays.

You can apply policing to either the inbound or outbound direction, while you can apply shaping only in the outbound direction. Policing drops nonconforming traffic instead of queuing the traffic like shaping. Policing also supports marking of traffic. Traffic policing is more efficient in terms of memory utilization than traffic shaping because no additional queuing of packets is needed.

Both traffic policing and traffic shaping ensure that traffic does not exceed a bandwidth limit, but each mechanism has a different impact on the traffic:

- Policing drops packets more often, generally causing more retransmissions of connection-oriented protocols such as TCP.
- Shaping adds variable delay to traffic, possibly causing jitter.

https://t.me/learningnets



The token bucket is a mathematical model that is used by routers and switches to regulate traffic flow. The model has two basic components:

- **Tokens:** Each token represents permission to send a fixed number of bits into the network. Tokens are put into a token bucket at a certain rate.
- **Token bucket:** A token bucket has the capacity to hold a specified number of tokens. Each incoming packet, if forwarded, takes tokens from the bucket, representing the packet size. If the bucket fills to capacity, newly arriving tokens are discarded. Discarded tokens are not available to future packets. If there are not enough tokens in the token bucket to send the packet, the traffic conditioning mechanisms may take these actions:
  - Wait for enough tokens to accumulate in the bucket (traffic shaping)
  - Discard the packet **(**traffic policing)

Using a single token bucket model, the measured traffic rate can conform or it can exceed the specified traffic rate. The measured traffic rate is conforming if there are enough tokens in the single token bucket to transmit the traffic. The measured traffic rate is exceeding if there are not enough tokens in the single token bucket to transmit the traffic.

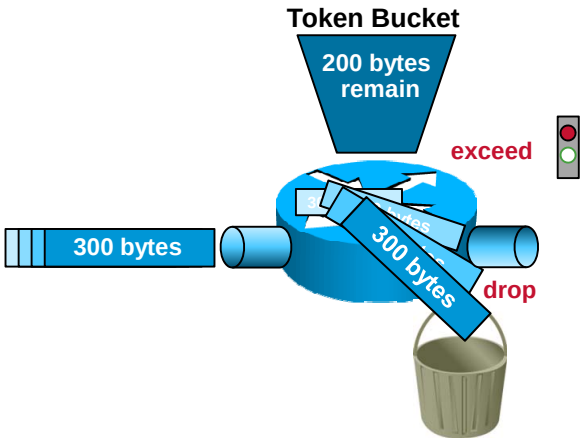
The figure shows a single token bucket traffic-policing implementation. Starting with a current capacity of 700 bytes worth of tokens that are accumulated in the token bucket, when a 500-byte packet arrives at the interface, its size is compared to the token bucket capacity (in bytes). The 500-byte packet conforms to the rate limit (500 bytes < 700 bytes), and the packet is forwarded; 500 bytes worth of tokens are taken out of the token bucket, leaving 200 bytes worth of tokens for the next packet.

.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Single Token Bucket (Cont.)



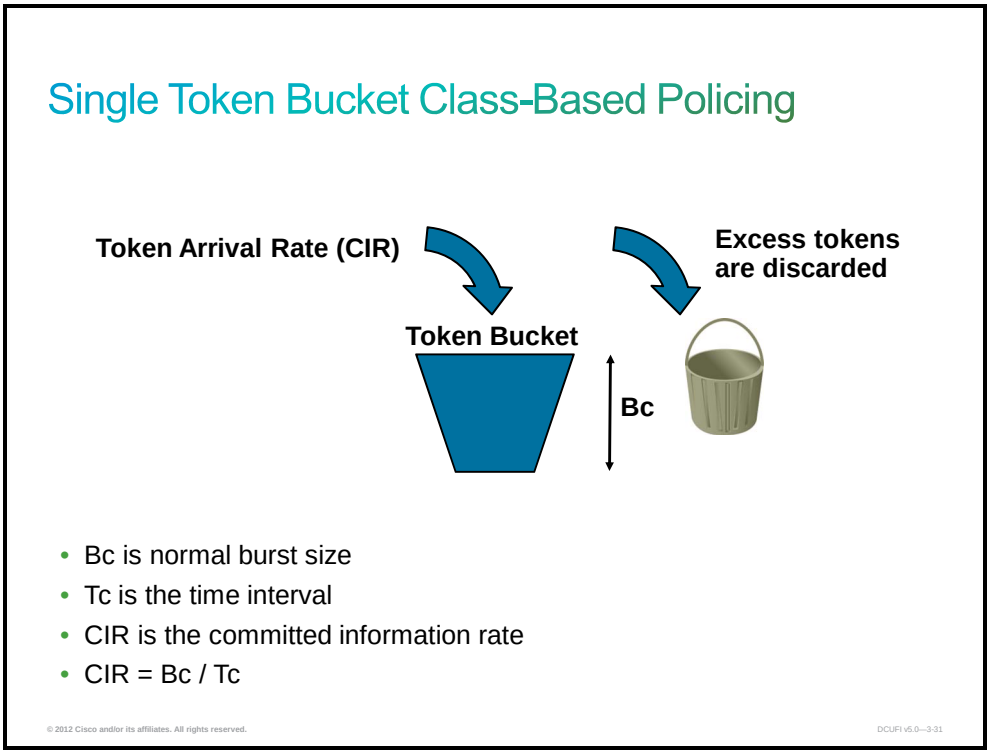
- If sufficient tokens are NOT available (exceed action):
  - Drop (or mark) the packet

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

When the next 300-byte packet arrives immediately after the first packet, and no new tokens have been added to the bucket (which is done periodically), the packet exceeds the rate limit. The current packet size (300 bytes) is greater than the current capacity of the token bucket (200 bytes), and the exceed action is performed. The exceed action can be to drop the packet or to remark the packet and then transmit it out.

.

https://t.me/learningnets



Token bucket operations rely on parameters such as the committed information rate (CIR), the normal burst size, or committed burst size (Bc), and the committed time interval, or committed rate measurement interval (Tc). The mathematical relationship between CIR, Bc, and Tc is as follows:

```
CIR (bps) = Bc (bits) / Tc (sec)
```

With traffic policing, new tokens are added into the token bucket based on the interpacket arrival rate and the CIR. Every time a packet is policed, new tokens are added back into the token bucket. The number of tokens that are added back into the token bucket is calculated as follows:

```
(Current Packet Arrival Time – Previous Packet Arrival Time) * CIR
```

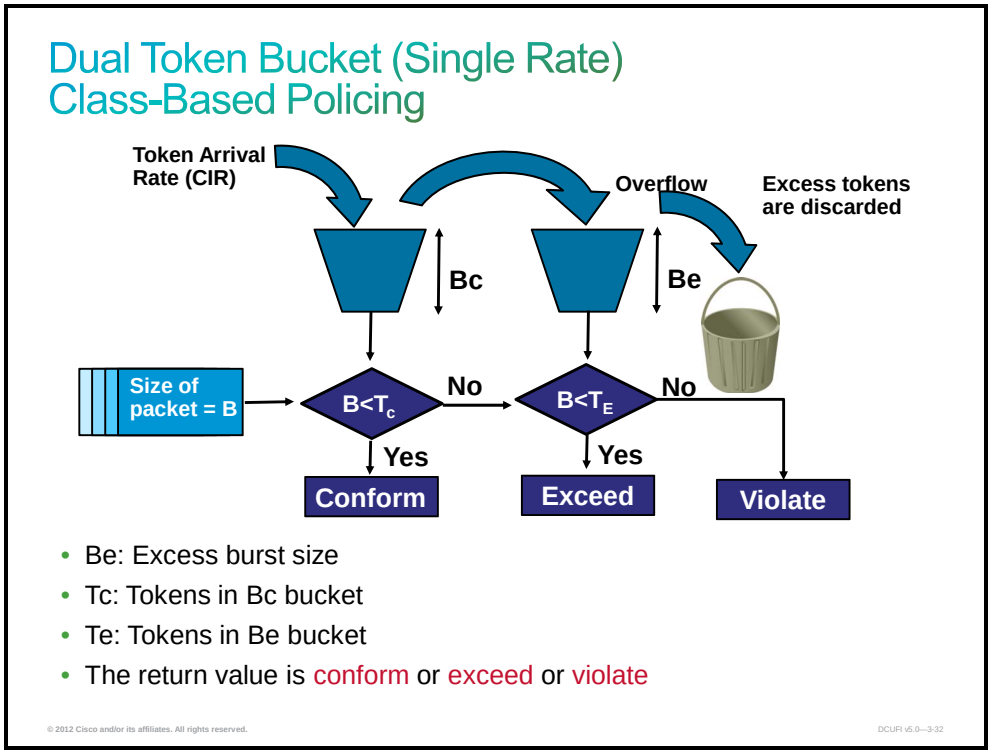
An amount (Bc) of tokens is forwarded without constraint in every time interval (Tc). For example, if 8000000 bits (Bc) worth of tokens are placed in the bucket every 250 ms (Tc), the router can steadily transmit 8000000 bits every 250 ms if traffic constantly arrives at the router.

```
CIR (normal burst rate) = 8000000 bits (Bc) / 0.25 seconds (Tc) = 32
Mbps
```

Without any excess bursting capability, if the token bucket fills to capacity (Bc of tokens), the token bucket will overflow and newly arriving tokens will be discarded. Using the example, in which the CIR is 32 Mb/s (Bc = 8000000 bits and Tc = 0.25 seconds), the maximum traffic rate can never exceed a hard rate limit of 32 Mb/s.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets



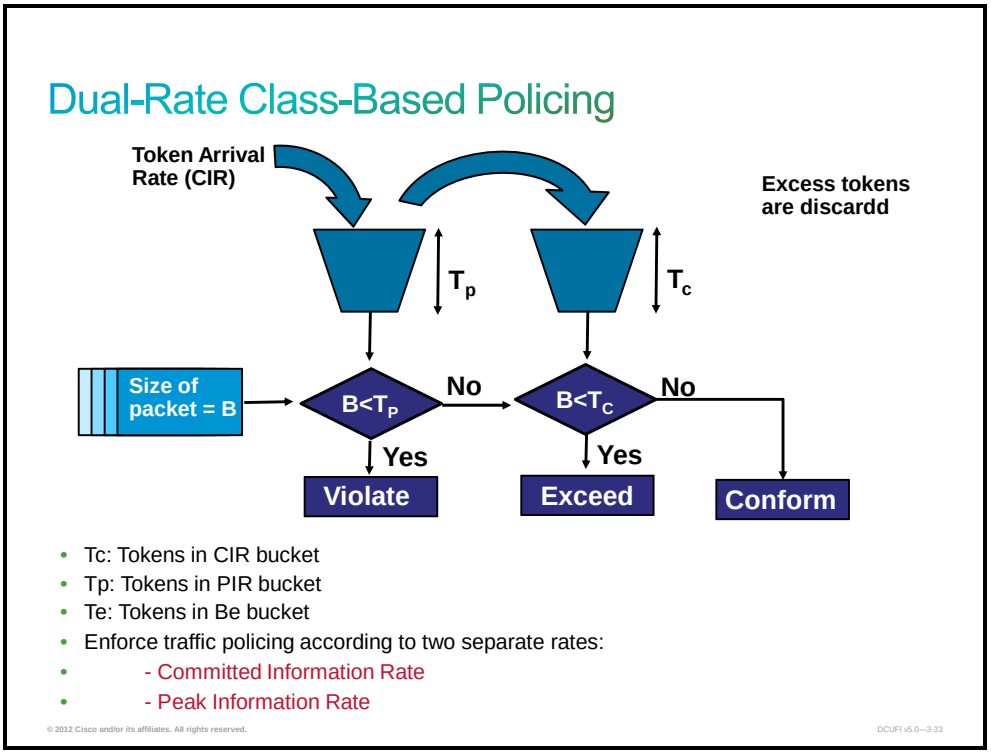
You can configure class-based traffic policing to support excess bursting capability. With excess bursting, after the first token bucket is filled to Bc, extra (excess) tokens can be accumulated in a second token bucket. Excess burst (Be) is the maximum amount of excess traffic over and above Bc that can be sent during the time interval *after* a period of inactivity. With a single rate-metering mechanism, the second token bucket with a maximum size of Be fills at the same rate (CIR) as the first token bucket. If the second token bucket fills up to capacity, no more tokens can be accumulated and the excess tokens are discarded.

When using a dual token bucket model, the measured traffic rate can be as follows:

- **Conforming:** There are enough tokens in the first token bucket with a maximum size of Bc.
- **Exceeding:** There are not enough tokens in the first token bucket, but there are enough tokens in the second token bucket with a maximum size of Be.
- **Violating:** There are not enough tokens in the first or second token bucket.

With dual token bucket traffic policing, the typical actions that are performed are sending all conforming traffic, re-marking (to a lower priority), sending all exceeding traffic, and dropping all violating traffic. The main benefit of using a dual token bucket method is the ability to distinguish between traffic that exceeds the Bc but not the Be. This method enables a different policy to be applied to packets in the Be category.

https://t.me/learningnets



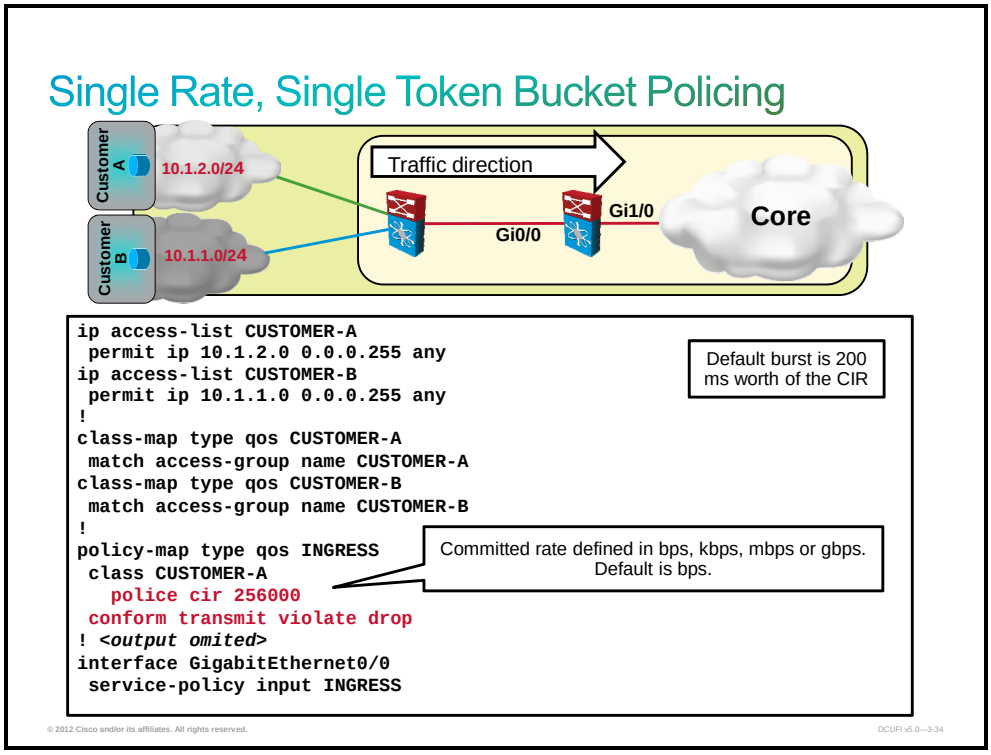
With dual-rate metering, traffic rate can be enforced according to two separate rates: CIR and peak information rate (PIR). Before this feature was available, you could meter traffic using a single rate that is based on the CIR with single or dual buckets. Dual-rate metering supports a higher level of bandwidth management and supports a sustained excess rate that is based on the PIR.

With dual-rate metering, the PIR token bucket fills at a rate that is based on the packet arrival rate, and the configured PIR and the CIR token bucket fills at a rate that is based on the packet arrival rate and the configured CIR.

When a packet arrives, the PIR token bucket is first checked to see if there are enough tokens in the PIR token bucket to send the packet. The violating condition occurs if there are not enough tokens in the PIR token bucket to transmit the packet. If there are enough tokens in the PIR token bucket to send the packet, then the CIR token bucket is checked. The exceeding condition occurs if there are enough tokens in the PIR token bucket to transmit the packet but not enough tokens in the CIR token bucket to transmit the packet. The conforming condition occurs if there are enough tokens in the CIR bucket to transmit the packet.

Dual-rate metering is often configured on interfaces at the edge of a network to police the rate of traffic that is entering or leaving the network. In the most common configurations, traffic that conforms is sent and traffic that exceeds is sent with a decreased priority, and traffic that violates is dropped. Users can change these configuration options to suit their network needs.

© 2012 Cisco Systems, Inc.
https://t.me/learningnets



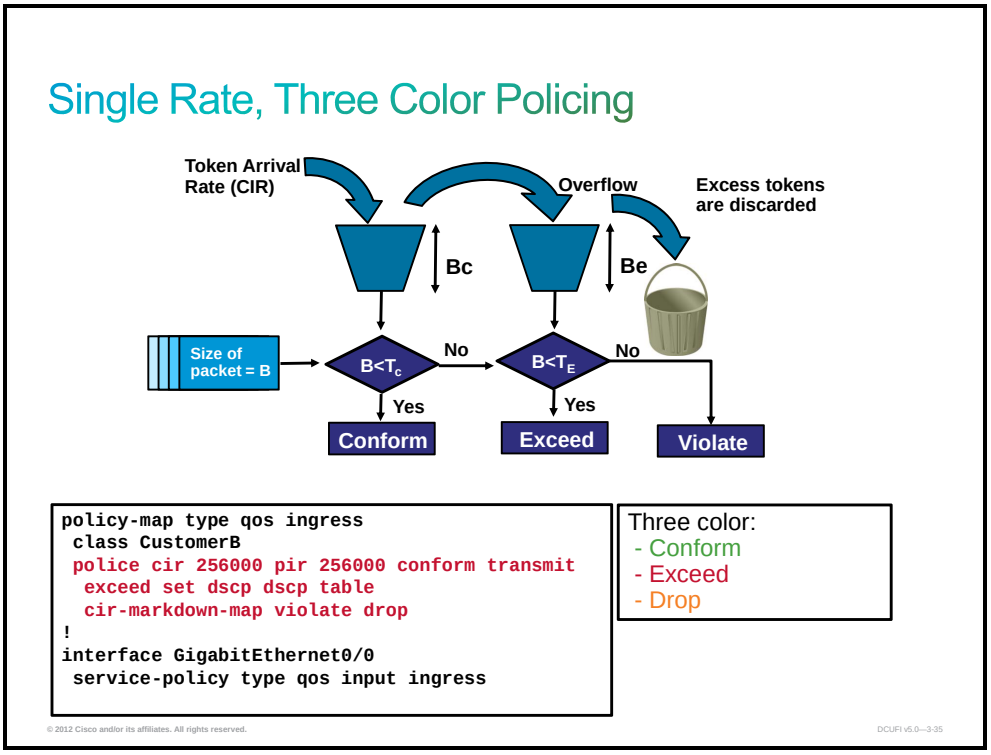
The class-based policing configuration example shows two configured traffic classes that are based on the traffic source IP address. Traffic that is originated by customer A is policed to a fixed bandwidth with no excess burst capability using a single token bucket. Conforming traffic is sent as is and exceeding traffic is dropped. In this case, the traffic from customer A is policed to a rate of 256 kb/s. The committed rate can be defined as a value in bits per second, kilobits per second, megabits per second, gigabits per second, and the default is bits per second.

A similar policy, not shown here, can be applied to traffic from Customer B, by adding another class in the ingress policy-map.

Because the violate action is not specified, this will use a single token bucket scheme and no excess bursting is allowed.

The default conform action is transmit, the default exceed action is drop. If no action is configured, the default action is taken.

https://t.me/learningnets



The class-based policing configuration example assumes that the traffic class Customer B is defined using any of the discussed methods, such as by matching the traffic using IP access groups, upstream MAC address, DSCP values, or any other parameters.

Traffic from customer B is policed to a fixed bandwidth with excess burst capability using a dual token bucket, by configuring a violate action. Conforming traffic will be sent as is, exceeding traffic will be marked according to dscp table *cir-markdown-map*, and transmitted. All violating traffic will be dropped.

© 2012 Cisco Systems, Inc.
https://t.me/learningnets

Verifying Policing

- You can verify policing configuration by showing policy map configuration:

```
N7K2-POD1# show policy-map interface ethernet 4/5 input


Global statistics status :   enabled

Ethernet4/5

  Service-policy (qos) input:   INGRESS

    Class-map (qos):   CUSTOMER-A (match-all)
      Match: access-group CUSTOMER-A
      police cir 256000 bps bc 200 ms
```

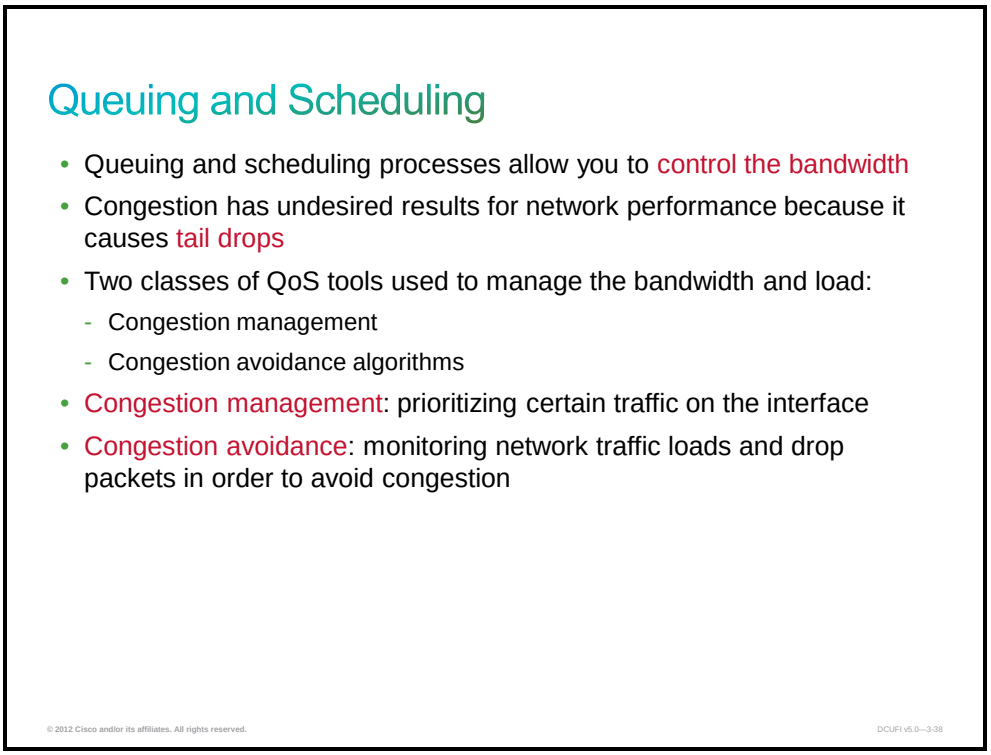
© 2012 Cisco and/or its affiliates. All rights reserved.

You can verify the policing configuration by showing the policy map that is applied to the particular interface.

https://t.me/learningnets

Queuing and Scheduling

This topic describes how to configure scheduling and queuing features on the Cisco Nexus Family of switches.



Speed or bandwidth mismatch is a common cause of congestion in a network. Speed mismatches can occur when traffic moves from a high-speed LAN environment (1000 Mb/s or higher) to lower-speed WAN links or in a LAN-to-LAN environment when, for example, a 10-Gb/s link feeds into a 1-Gb/s link. Other typical places of congestion are aggregation points. In a LAN environment, congestion resulting from aggregation often occurs at the distribution layer of networks, where the different access layer devices feed traffic to the distribution-level devices. Congestion has undesired results, because it causes packet drops.

Two classes of tools allow you to control the bandwidth and manage congestion and load:

- Congestion management
- Congestion avoidance

Congestion management uses queuing algorithms as one of the primary ways to manage congestion in a network. Network devices manage an overflow of arriving traffic by using a queuing algorithm to sort traffic and determine a method of prioritizing the traffic onto an output link. Each queuing algorithm was designed to solve a specific network traffic problem and has a particular effect on network performance. The QoS queuing features are enabled by default, so the device always enables a system default queuing policy. When users configure their own queuing policy and apply the new queuing policy to specified interfaces, the new queuing policy replaces the default queuing policy.

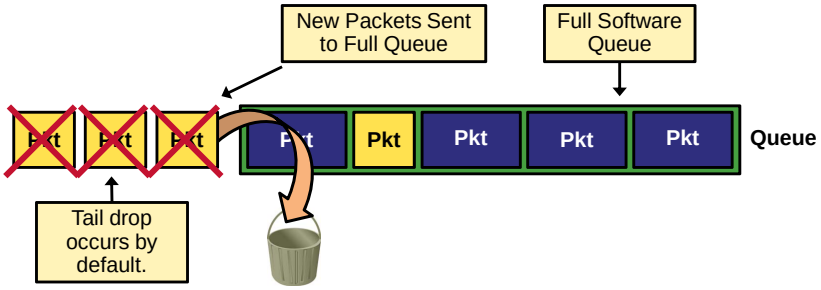
Congestion avoidance techniques monitor network traffic loads to anticipate and avoid congestion at common network bottleneck points. Congestion avoidance is achieved through packet dropping using a more complex dropping technique than simple tail drop. This lesson describes tail drop and the weighted random early detection (WRED) congestion avoidance, which is the Cisco implementation of random early detection (RED).

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Tail Drop

- Congestion occurs when the queue is full:
  - Additional incoming packets are tail-dropped.
  - Dropped packets may degrade application performance.
- Tail drop drawbacks:
  - TCP synchronization
  - TCP starvation
  - No differentiated drop



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

When an interface on a switch or router cannot transmit a packet immediately, the packet is queued. Packets are then taken out of the queue and eventually transmitted on the interface.

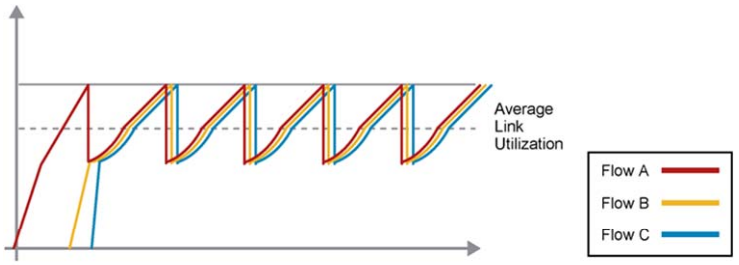
If the arrival rate of packets to the output interface exceeds the router capability to buffer and forward traffic, the queues increase to their maximum length and the interface becomes congested. Tail drop is the default queuing response to congestion. Tail drop treats all traffic equally and does not differentiate between classes of service. Applications may suffer performance degradation due to packet loss that is caused by tail drop. When the output queue is full and tail drop is in effect, all packets that are trying to enter (at the tail of) the queue are dropped until the congestion is eliminated and the queue is no longer full.

The simple tail-drop scheme does not work well in environments with many TCP flows or in environments in which selective dropping is required.

https://t.me/learningnets

TCP Synchronization

- Multiple TCP sessions start at different times.
- TCP window sizes are increased.
- Tail drops cause many packets of many sessions to be dropped at the same time.
- TCP sessions restart at the same time (synchronized).



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

Understanding the network interaction between TCP stack intelligence and dropping is required to implement a more efficient and fair dropping scheme, especially in the networks with a high number of applications that require differentiated treatment.

Tail drop has the following shortcomings:

- When congestion occurs, dropping affects most of the TCP sessions, which simultaneously back off and then restart again. This action causes inefficient link utilization at the congestion point (TCP global synchronization).
- TCP starvation, in which all buffers are temporarily seized by aggressive flows, and normal TCP flows experience buffer starvation.

There is no differentiated drop mechanism, and therefore premium traffic is dropped in the same way as best-effort traffic.

A Cisco Nexus switch can manage a large number of concurrent TCP sessions. It is likely that when traffic exceeds the queue limit, it exceeds this limit because of the bursty nature of packet networks. However, there is also a high probability that excessive traffic depth that is caused by packet bursts is temporary and that traffic does not stay excessively deep except at points where traffic flows merge, or at edge routers.

If the receiving switch drops all traffic that exceeds the queue limit, as is done with tail drop by default, many TCP sessions simultaneously go into slowstart. So, traffic temporarily slows down to the extreme and then all flows slowstart again. This activity creates a condition called global synchronization.

Global synchronization occurs as waves of congestion crest only to be followed by troughs during which the transmission link is not fully used. Global synchronization of TCP hosts can occur because packets are dropped all at once. It occurs when multiple TCP hosts reduce their transmission rates in response to packet dropping. When congestion is reduced, their transmission rates are increased. The most important point is that the waves of transmission known as global synchronization result in significant link underutilization.

https://t.me/learningnets

Random Early Detection (RED)

- Tail drop can be avoided if congestion is prevented.
- RED:
  - Mechanism that randomly drops packets before a queue is full.
  - Increases drop rate as the average queue size increases.
- RED results:
  - TCP sessions slow down to the approximate rate of output-link bandwidth.
  - Average queue size is small (much less than the maximum queue size).
  - TCP sessions are desynchronized by random drops.

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

RED is a dropping mechanism that randomly drops packets before a queue is full. The dropping strategy is based primarily on the average queue length—that is, when the average size of the queue increases, RED will be more likely to drop an incoming packet than when the average queue length is shorter.

Because RED drops packets randomly, it has no per-flow intelligence. The rationale is that an aggressive flow will represent most of the arriving traffic and therefore it is likely that RED will drop a packet of an aggressive session. RED therefore punishes more aggressive sessions with higher statistical probability and is, therefore, able to somewhat selectively slow down the most significant cause of congestion. Directing one TCP session at a time to slow down allows for full utilization of the bandwidth, rather than utilization that manifests itself as crests and troughs of traffic.

As a result of implementing RED, the problem of TCP global synchronization is much less likely to occur, and TCP can utilize link bandwidth more efficiently. In RED implementations, the average queue size also decreases significantly, as the possibility of the queue filling up is reduced. This is because of very aggressive dropping in the event of traffic bursts, when the queue is already quite full.

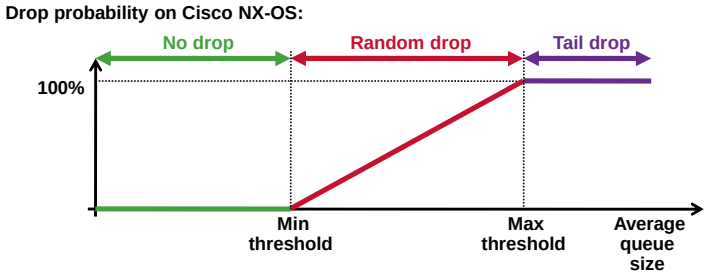
RED distributes losses over time and normally maintains a low queue depth while absorbing traffic spikes. RED can also use markers, such as DSCP, to establish different drop profiles for different classes of traffic, in which case it is referred to as WRED.

https://t.me/learningnets

RED Profiles

RED modes:

- No drop: When the average queue size is between 0 and the minimum threshold.
- Random drop: When the average queue size is between the minimum and the maximum threshold.
- Tail drop: When the average queue size is at maximum threshold or above.



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

A RED traffic profile is used to determine the packet-dropping strategy and is based on the average queue length. The probability of a packet being dropped is based on two thresholds that are contained within the RED profile:

- **Minimum threshold:** When the average queue length is equal or above the minimum threshold, RED starts dropping packets. The rate of packet drop increases linearly as the average queue size increases, until the average queue size reaches the maximum threshold.
- **Maximum threshold:** When the average queue size is above the maximum threshold, all packets are dropped.

The minimum threshold value should be set high enough to maximize the link utilization. If the minimum threshold is too low, packets may be dropped unnecessarily. The difference between the maximum threshold and the minimum threshold should be large enough to avoid global synchronization. If the difference is too small, many packets may be dropped at once, resulting in global synchronization.

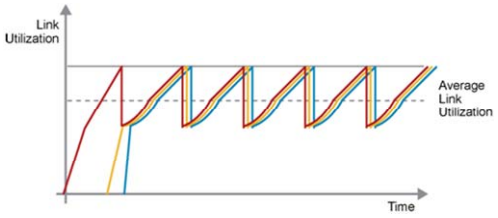
Based on the average queue size, RED has three dropping modes:

- When the average queue size is between 0 and the configured minimum threshold, no drops occur and all packets are queued.
- When the average queue size is between the configured minimum threshold and the configured maximum threshold, random drops occur, which is linearly proportional to the mark probability denominator and the average queue length.
- When the average queue size is at or higher than the maximum threshold, RED performs full (tail) drop in the queue. This situation is unlikely because RED should slow down TCP traffic ahead of congestion. If much non-TCP traffic is present, RED cannot effectively drop traffic to reduce congestion, and tail drops are likely to occur.

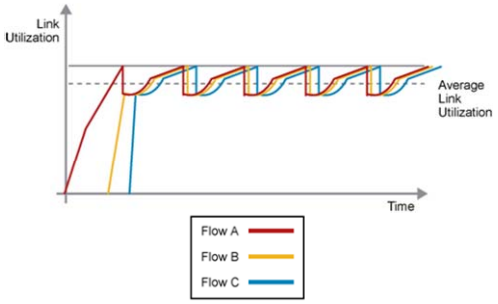
https://t.me/learningnets

TCP Traffic With RED

- Without RED:
  - TCP synchronization prevents average link utilization close to the link bandwidth.
  - Tail drops cause TCP sessions to go into slow-start.



- With RED:
  - Average link utilization is much closer to link bandwidth.
  - Random drops cause TCP sessions to reduce window sizes.



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

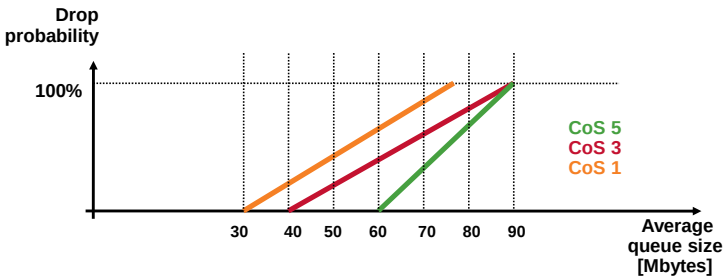
The first figure shows TCP throughput behavior that is compared to link bandwidth in a congested network scenario where the tail-drop mechanism is in use on a router. The global synchronization phenomenon causes all sessions to slow down when congestion occurs. All sessions are penalized when tail drop is used because it drops packets with no discrimination between individual flows. When all sessions slow down, congestion on the router interface is removed and all TCP sessions restart their transmission at roughly the same time. Again, the router interface quickly becomes congested, causing tail drop. As a result, all TCP sessions back off again. This behavior cycles constantly, resulting in a link that is generally underutilized.

The second figure shows TCP throughput behavior that is compared to link bandwidth in a congested network scenario in which RED has been configured on a router. RED randomly drops packets, influencing a few sessions at a time, before the interface reaches congestion. Overall throughput of sessions is increased, as well as average link utilization. Global synchronization is very unlikely to occur, because of selective, but random, dropping of adaptive traffic.

https://t.me/learningnets

Weighted Random Early Detection

- WRED can use multiple different RED profiles.
  - Drops less important packets more aggressively than important packets.
- Each profile is identified by:
  - Minimum threshold
  - Maximum threshold



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

WRED performs differentiated packet dropping based on packet markers, such as CoS. As with RED, WRED monitors the average queue length in the switch and determines when to begin discarding packets that are based on the length of the interface queue. When the average queue length is greater than the user-specified minimum threshold, WRED begins to randomly drop packets (both TCP and User Datagram Protocol [UDP] packets) with a certain probability. If the average length of the queue becomes larger than the maximum threshold, WRED reverts to a tail-drop packet discard strategy.

WRED can selectively discard lower-priority traffic when the interface becomes congested, and can provide differentiated performance characteristics for different classes of service.

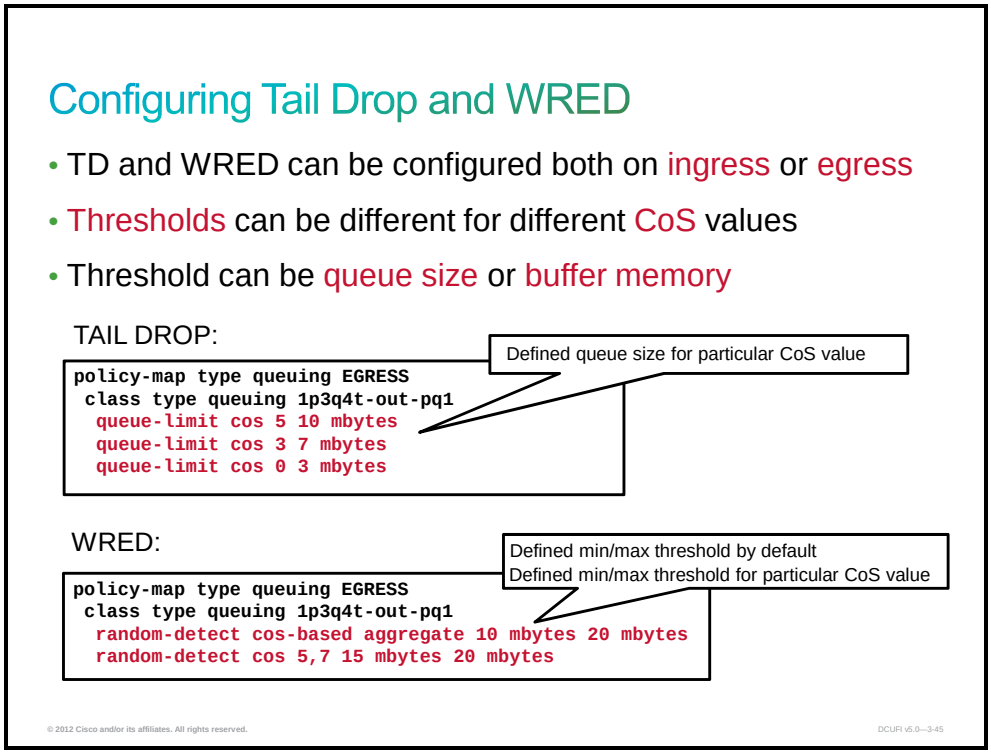
WRED is only useful when the bulk of the traffic is TCP traffic. With TCP, dropped packets indicate congestion, so the packet source reduces its transmission rate. With other protocols, packet sources might not respond or might resend dropped packets at the same rate, and so dropping packets might not decrease congestion.

WRED is more often used in the core than in the edge and the access network. Access and edge switches mark packets. WRED uses these assigned values to determine how to treat different types of traffic.

WRED is not recommended for voice and video. WRED will not throttle back voice traffic because voice traffic is UDP-based. The network itself should be designed not to lose voice packets because lost voice packets result in reduced voice quality.

The figure illustrates differentiated RED profiles. Each class has the minimum and maximum thresholds set to specific CoS values.

https://t.me/learningnets



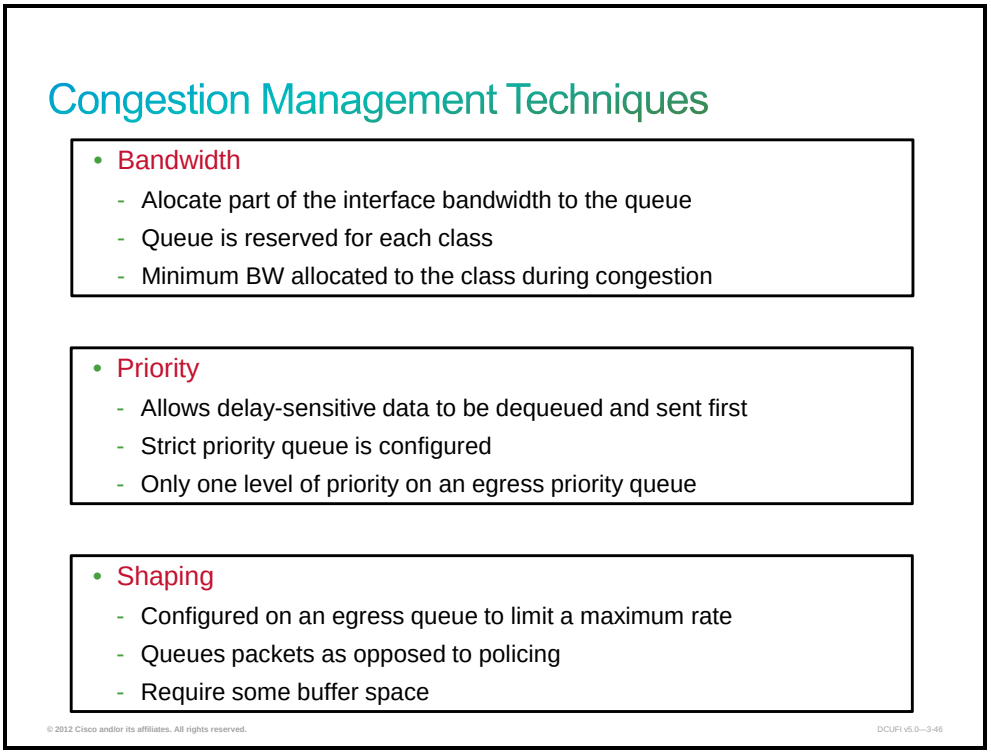
Congestion avoidance features (tail drop and WRED) can be configured both on ingress or egress ports.

For tail drop, thresholds can be configured for different CoS values. This capability means that higher CoS value can have maximum queue size or allocated buffer percentage size that is larger than the lower CoS value. In the example that is provided, queue size for CoS 5 value is 10 Mb/s, for CoS, it is 3 – 7 Mb/s and for CoS 0, it is 3 Mb/s. The device will drop the packets of CoS 1 first, when the queue reaches 3 Mb/s.

As was mentioned in the WRED theory explanation, the frequency of dropped packets increases as the queue size exceeds the minimum threshold. When the maximum threshold is exceeded, all packets for the CoS value are dropped.

The next example illustrates a WRED configuration scenario. Using the first command, you configure WRED for all CoS values that are not configured by a CoS-specific random-detect second command. Using a second command, you can specify minimum and maximum thresholds that are used to drop packets from the queue for specific CoS values. You can configure thresholds by the number of packets, number of bytes, the duration of time at the underlying interface minimum guaranteed link rate, or as the percentage of queue size and the default threshold is in packets. You must configure the minimum and maximum thresholds to be of the same type. For example, you must use *mbytes* for both minimum and maximum thresholds.

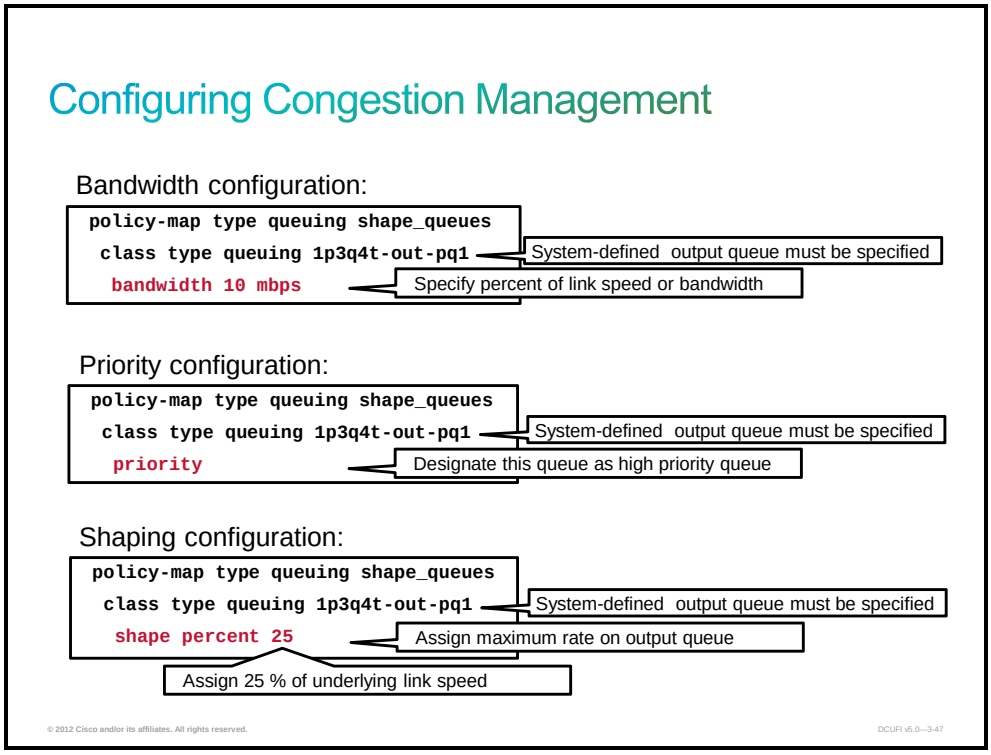
https://t.me/learningnets



Next, congestion management methods can be configured in a policy map:

- **Bandwidth allocation:** Allocates minimum bandwidth that is reserved for the class in the moments of congestion on the interface. Bandwidth is allocated to the queue on per-class basis. This bandwidth can be configured also as a percent of the remaining bandwidth that is calculated as available bandwidth subtracted by all reserved bandwidth on the interface.
- **Priority queue:** One or more of the queues can be strict priority queues. This setting allows delay-sensitive data to be dequeued first and sent out to the interface. Typical classes on which priority reservation can be applied are voice and interactive video.
- Shaping: Can be configured to limit maximum rate on an egress queue. As opposed to policing, shaping queues packets rather than drops them, and as a result, some additional buffer space is required to store those queued packets.

© 2012 Cisco Systems, Inc.
https://t.me/learningnets



When configuring bandwidth, policing, and shaping, you must use one of the system-defined output queues. In the example that is provided, *1p3q4t-out-pq1* is selected for the output queue.

Using the **bandwidth** command, you are assigning the bandwidth to the output queue *1p3q4t-out-pq1*. The example shows a data rate that is configured for the bandwidth that is provided, but it can be configured as a percentage of the link.

In the next example, using the **priority** keyword, you designate the system-defined *1p3q4t-out-pq1* as a priority queue that has strict priority over the other classes.

As opposed to the **bandwidth** command, using the **shape** command, you configure the maximum rate that is allowed for the *1p3q4t-out-pq1* class. In the example that is provided, 25 percent of the underlying link speed is used as above limit for the traffic class on which the shaping is applied.

https://t.me/learningnets

Verifying Queuing and Scheduling Configuration

- You can verify queuing and scheduling configuration by showing type queuing policy map configuration:

```
N7K2-POD1# show policy-map interface ethernet 4/5 input


Global statistics status :   enabled

Ethernet4/5

! <Output omitted…>

  Service-policy (queuing) input:   default-4q-8e-in-policy

    Class-map (queuing):   2q4t-8e-in-q1 (match-any)
      queue-limit percent 10
      bandwidth percent 50
      queue dropped pkts : 0

    Class-map (queuing):   2q4t-8e-in-q-default (match-any)
      queue-limit percent 90
      bandwidth percent 50
      queue dropped pkts : 0
```

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

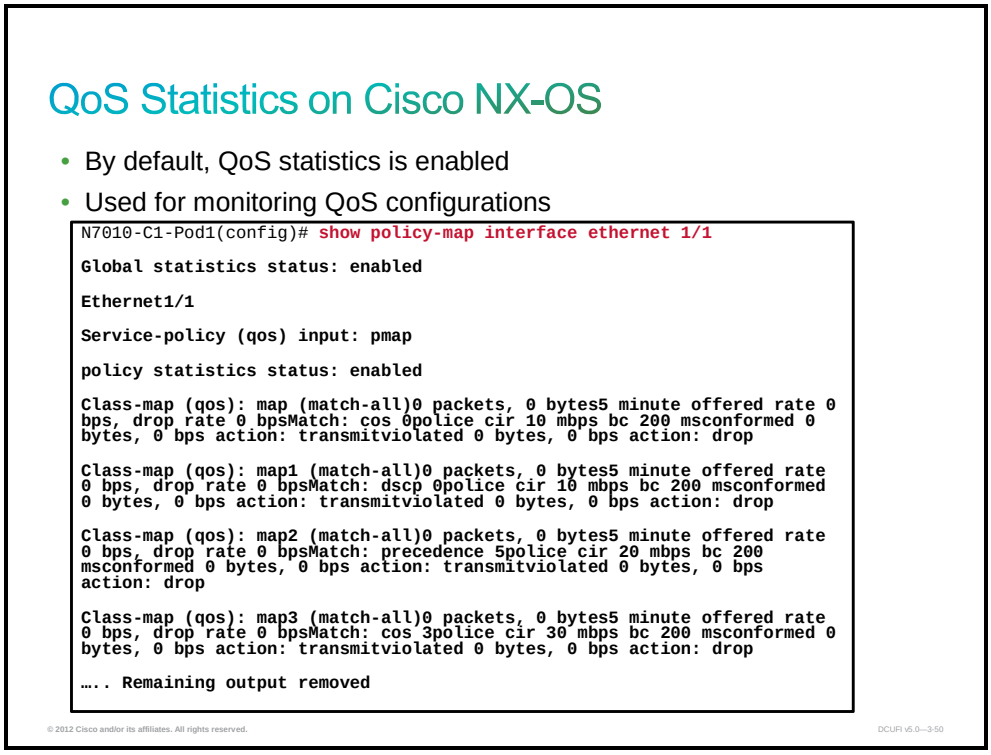
The **show policy-map interface** command displays the configuration of all classes that are configured for all service policies on the specified interface. This configuration includes all WRED parameters that are implementing the drop policy on the specified interface.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Monitoring QoS

This topic describes how to monitor QoS statistics on the Cisco Nexus Family of switches.



QoS statistics can be monitored, displayed, and cleared on the Cisco NX-OS devices. To monitor QoS statistics on the Cisco Nexus 7000 Series Switches you should be logged into the correct VDC.

By default, QoS statistics are enabled, but you can disable QoS statistics on an interface or all interfaces if required. QoS statistics can be displayed for all interfaces or a specific interface, data direction, or a QoS type.

The figure provides an example of how to display the QoS statistics and the type of information that might be displayed.

https://t.me/learningnets

Summary

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

Summary

- Goal of QoS is desirable flow of traffic through a network; the Cisco Nexus 7000 Series switch provides a rich set of data center QoS features.
- Using MQC, class maps, policy maps and service policies are defining QoS policies that apply on particular interface or network wide.
- Classification identify and categorize traffic into different classes; Marking is used for "coloring" packet using traffic descriptors
- Mutation mapping is the process of modifying a QoS field in all packets on an interface either ingress or egress.
- Policing typically drops excess traffic to stay within the limit; alternatively, it can re-mark, then send excess traffic) are mechanisms that use classification to limit traffic rate.
- Queuing and scheduling mechanisms allow you to control the bandwidth on an interface
- QoS statistics are monitored using show policy map interface command

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

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Lesson 5

Configuring Security Features

Overview

In this lesson, you will learn how to evaluate the security features that are available on the Cisco Nexus Family of switches and identify which features should be implemented into a solution. You will also learn how to identify different security threats and which techniques and technologies can be used to mitigate them.

Objectives

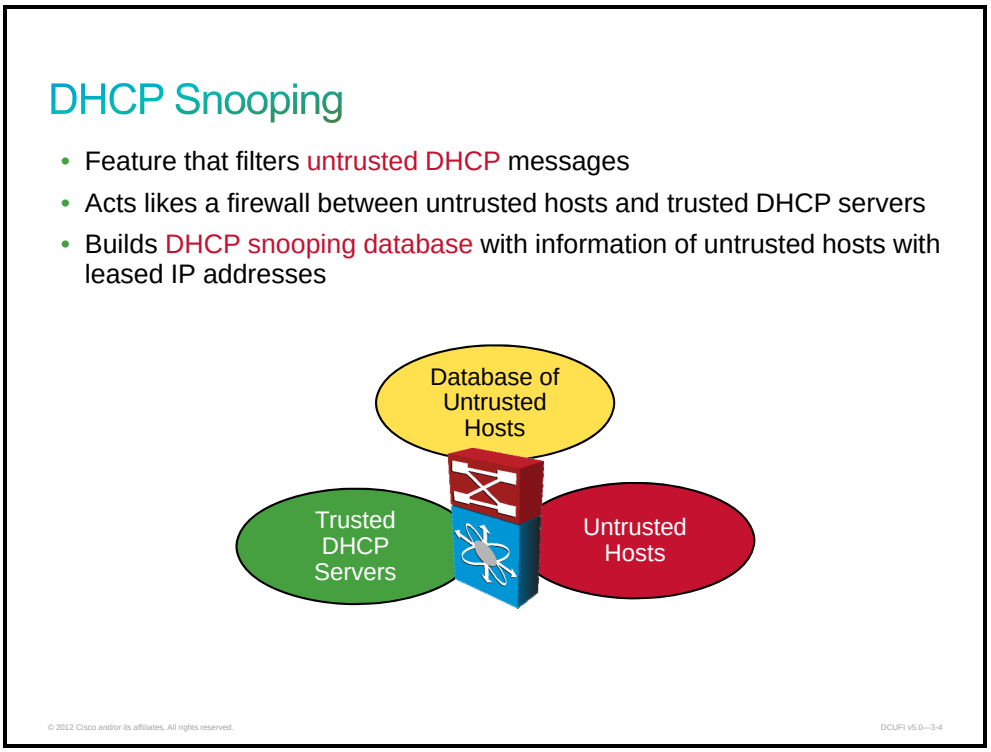
Upon completing this lesson, you will be able to evaluate the security features that are available on the Cisco Nexus Family of switches to identify which features should be implemented into a solution. This ability includes being able to meet these objectives:

- Configure DHCP snooping on the Cisco Nexus Family of switches and verify the configuration
- Configure DAI on the Cisco Nexus Family of switches and verify the configuration
- Configure IP Source Guard on the Cisco Nexus Family of switches and verify the configuration
- Configure uRPF on the Cisco Nexus Family of switches and verify the configuration
- Configure traffic storm control on the Cisco Nexus Family of switches and verify the configuration
- Configure port security on the Cisco Nexus Family of switches and verify the configuration
- Configure CoPP on the Cisco Nexus Family of switches and verify the configuration
- Configure and use the Cisco TrustSec solution on the Cisco Nexus Family of switches

https://t.me/learningnets

DHCP Snooping

This topic describes how to configure DHCP snooping on the Cisco Nexus Family of switches and verify the configuration.



The DHCP snooping security feature provides a security mechanism that filters untrusted DHCP messages and passes trusted DHCP messages. It acts like a firewall between untrusted hosts and trusted DHCP servers.

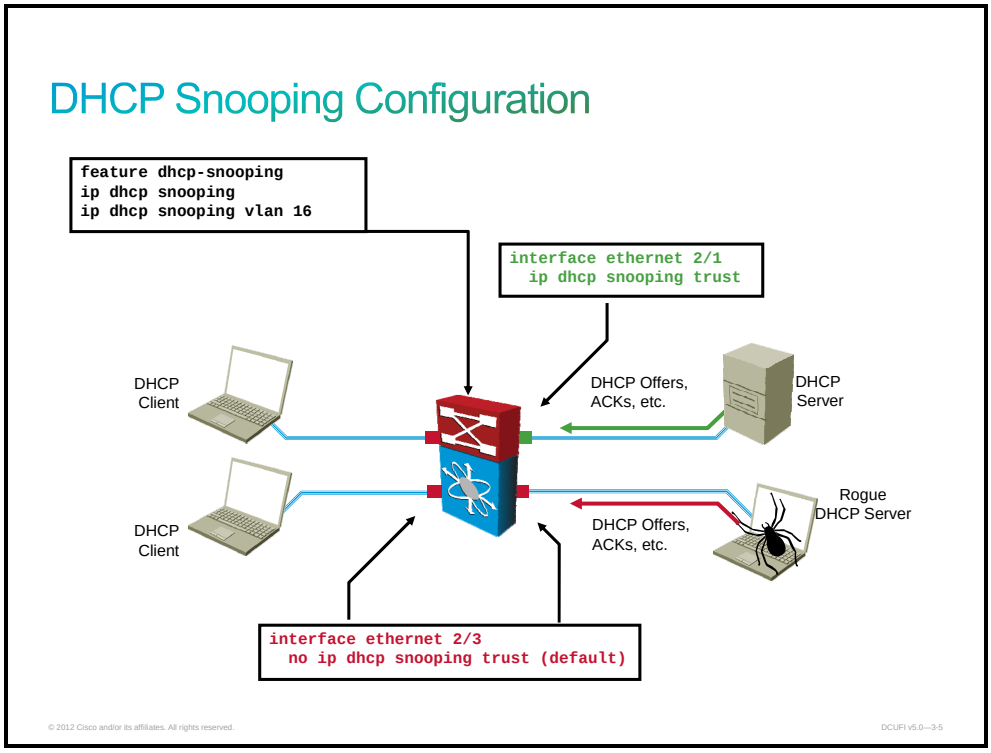
DHCP snooping is supported on the Cisco Nexus 7000 Series Switches and the Cisco Nexus 5000 and 5500 Platform switches, and performs these functions:

- Assigns a level of trust to a switch port
- Validates messages that are received from untrusted sources and discards invalid messages
- Builds and maintains a DHCP snooping binding database of untrusted hosts with leased IP addresses, and uses this database to validate messages that are received from untrusted hosts

In virtual device contexts (VDCs), the DHCP snooping binding database is unique for each VDC. There is no limit to the database size.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets



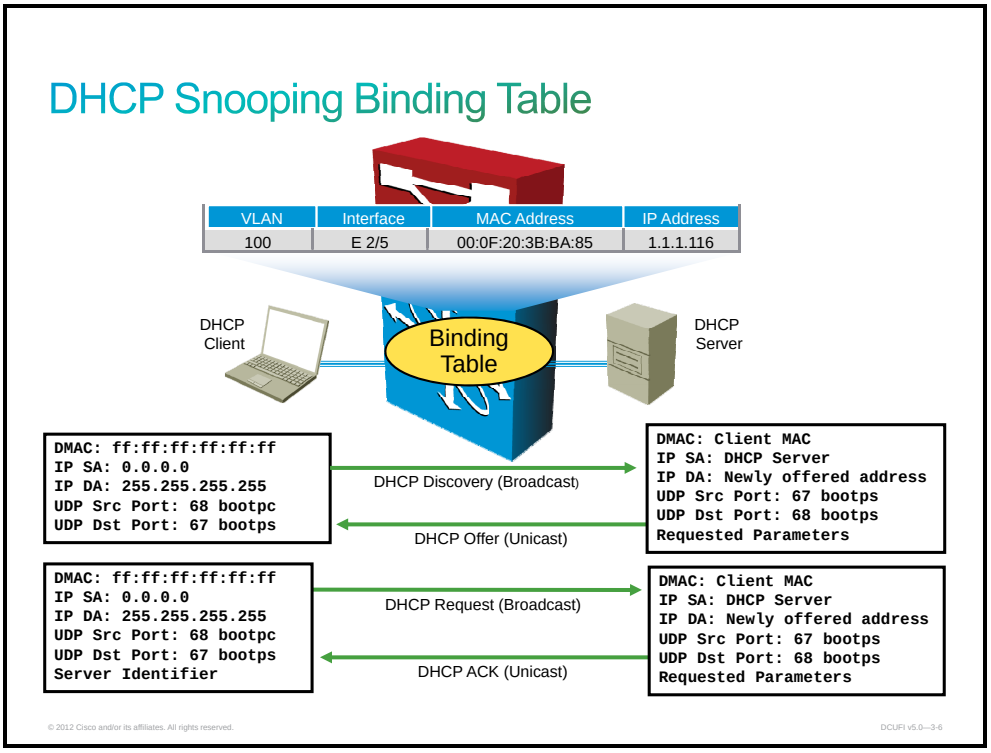
On interfaces with DHCP snooping enabled, the device validates the DHCP packet that is received on the untrusted interfaces and either forwards or drops the packet. Packets are dropped if any of the following circumstances occur:

- Receipt of a DHCP response packet on an untrusted interface
- Receipt of a DHCP packet on an untrusted interface where the source MAC address and the DHCP client hardware address do not match, if DHCP snooping MAC address verification is enabled
- Receipt of a DHCPRELEASE or DHCPDECLINE message on an untrusted interface with an entry in the DHCP snooping binding table that does not match the interface on which the message was received
- Receipt of a DHCP packet that includes a relay agent IP address other than 0.0.0.0

Option 82 supports metropolitan Ethernet-access environments where DHCP centrally manages the IP addresses for many subscribers. Option 82 permits identification of a subscriber device even though it is connected to the same port on the access device, thus providing unique identification for each subscriber. When option 82 is enabled, the Cisco Nexus Operating System (NX-OS) device adds the option 82 information into the DHCP packet. This process permits the DHCP server to assign an IP address that is based on any parameters that are configured. The DHCP server must be option 82-capable to support this functionality.

For complete list of configuration commands refer to *Cisco Nexus 7000 Series NX-OS Security Configuration Guide, Release 6.x* at http://www.cisco.com/en/US/docs/switches/datacenter/sw/6_x/nx-os/security/configuration/guide/b_Cisco_Nexus_7000_NX-OS_Security_Configuration_Guide__Release_6.x.html.

https://t.me/learningnets



This figure shows the building of the DHCP snooping database.

This database does not contain information about hosts that are interconnected with a trusted interface. A trusted interface is one that is configured to receive only messages from within the network and is connected to the DHCP server or another switch.

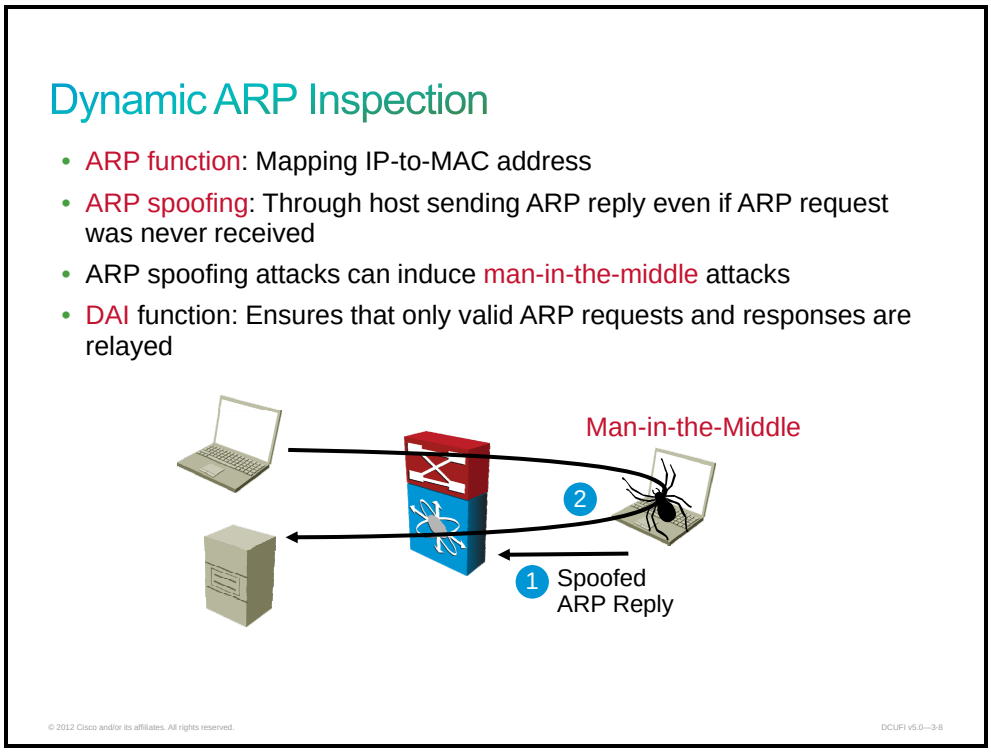
All entries in the database include the MAC address of the host, leased IP address, lease time, binding type, and the VLAN number and interface information that is associated with that host.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Dynamic ARP Inspection

This topic describes how to configure Dynamic ARP Inspection (DAI) on the Cisco Nexus Family of switches and verify the configuration.

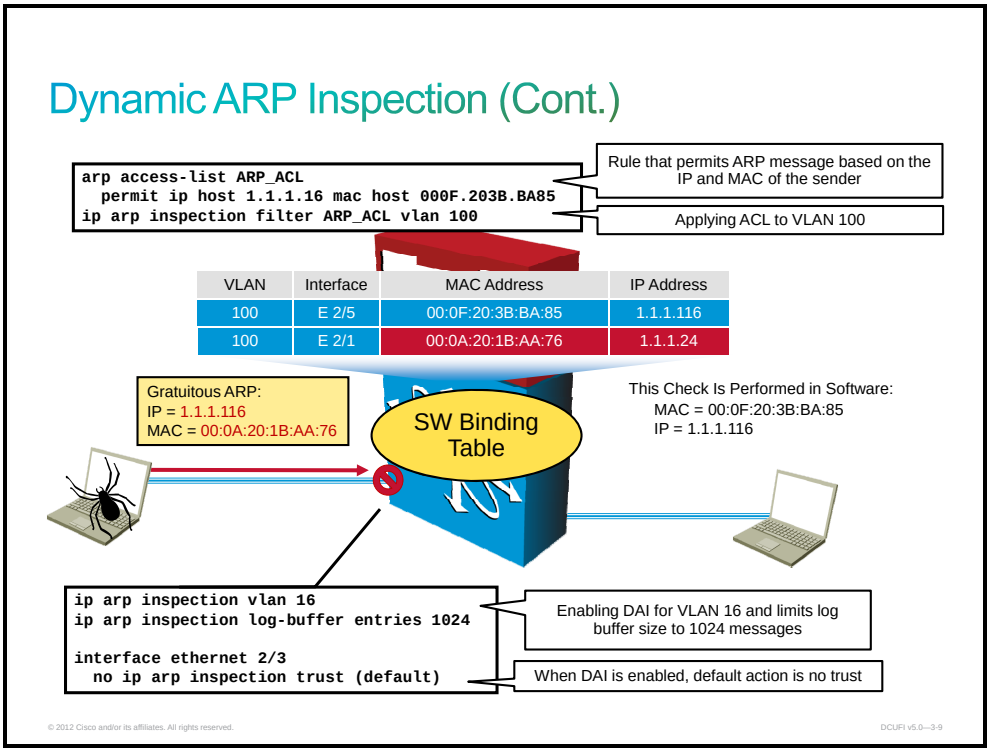


The Address Resolution Protocol (ARP) mechanism enables IP communication within a Layer 2 broadcast domain by IP-to-MAC address mapping.

ARP spoofing attacks and ARP cache poisoning can occur when ARP permits a reply from a host even if an ARP request was never received. This occurrence could cause all traffic from a device under attack to flow through the attacker computer (man-in-the-middle attack) before being forwarded to a router, switch, or host.

The purpose of DAI is to ensure that only valid ARP requests and responses are relayed. If DAI is enabled and configured correctly on the Cisco NX-OS device, the device intercepts all ARP requests and responses on untrusted ports. It verifies that these intercepted packets have a valid IP-to-MAC address binding before updating the local ARP cache. Invalid ARP packets are dropped.

https://t.me/learningnets



The validity check of a valid IP-to-MAC address binding is performed by checking the DHCP snooping binding database.

DAI is used to protect against ARP spoofing attacks and ARP cache poisoning:

- Intercepts and validates all ARP requests and responses on untrusted ports
- Rate-limits ARP traffic, 15-p/s default, configurable

In a typical network configuration, switch ports that are connected to host ports are untrusted, and switch ports that are connected to switches are trusted.

In the example shown, the ARP Inspection feature is configured for VLAN 100. ARP messages are permitted from hosts that are defined in ARP_ACL. Nonstatic entries are checked via DHCP snooping database. DAI can determine the validity of an ARP packet based on valid IP-to-MAC address bindings that are stored in a DHCP snooping binding database.

Using the command **ip arp inspection log-buffer**, you can configure the maximum buffer that is reserved for DAI logging messages. When DAI is enabled, default action on the switch port is *not* to trust ARP messages on that port.

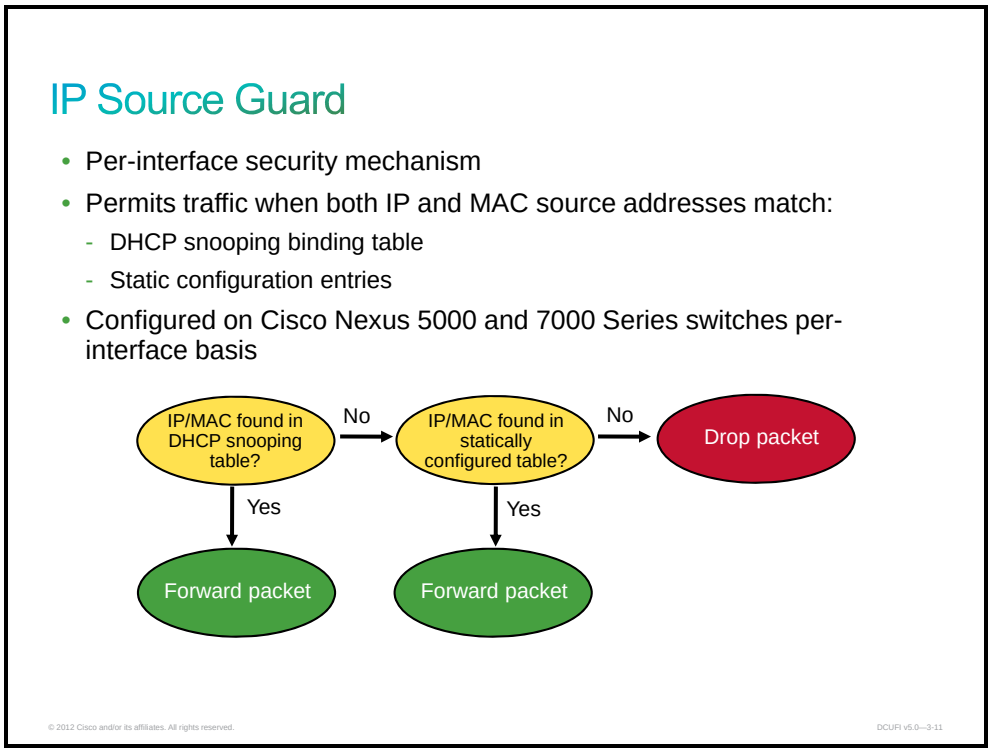
For complete list of configuration commands refer to *Cisco Nexus 7000 Series NX-OS Security Configuration Guide, Release 6.x* at http://www.cisco.com/en/US/docs/switches/datacenter/sw/6_x/nx-os/security/configuration/guide/b_Cisco_Nexus_7000_NX-OS_Security_Configuration_Guide__Release_6.x.html.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

IP Source Guard

This topic describes how to configure IP Source Guard on the Cisco Nexus Family of switches and verify the configuration.



IP Source Guard is a security mechanism that filters per-interface traffic that permits IP traffic only when the IP address and MAC address of each packet matches one of two sources of IP and MAC address bindings, entries in the DHCP snooping binding table, or static IP source entries that are configured.

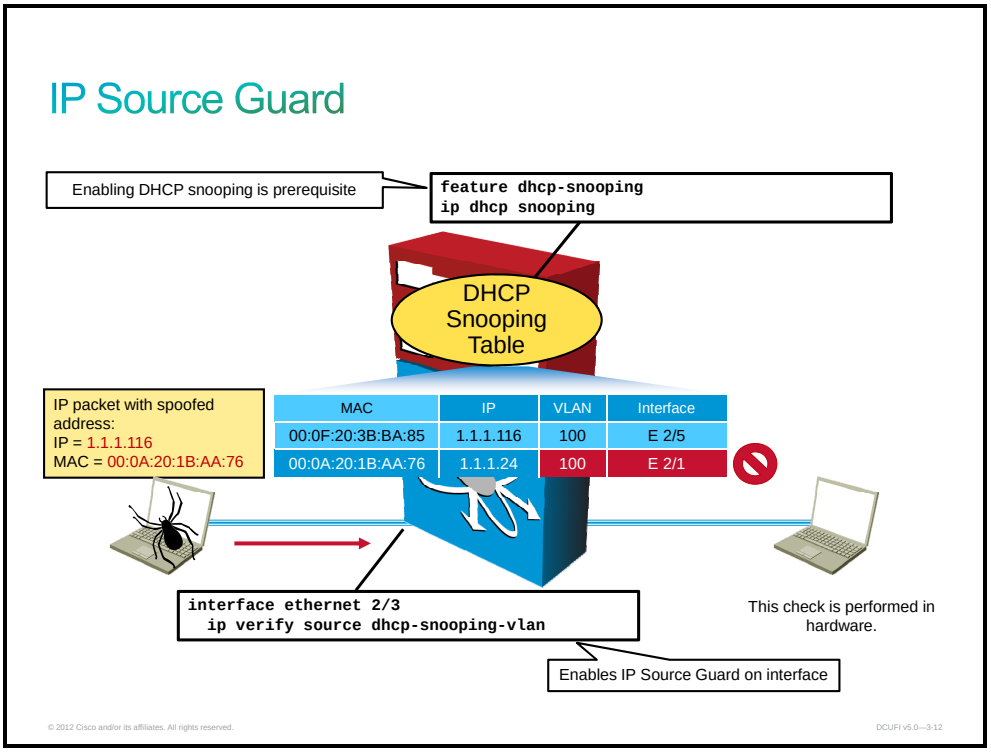
This feature helps to prevent spoofing attacks because an attacker would have to spoof both the IP address and the MAC address of a valid host to gain unauthorized access to the network.

The DHCP snooping binding database is downloaded into the forwarding engine:

- The DHCP snooping binding database is fully supported in the hardware without using access control list (ACL) ternary content addressable memory (TCAM).
- The database supports dynamic and static DHCP binding.
- The hardware-based MAC-IP binding check has a higher priority over the ACL result.
- Binding of the same IP address to multiple MAC addresses is not supported.
- IP Source Guard is not supported on trunk ports.

IP-MAC address bindings are unique for each VDC on the Cisco Nexus 7000 Series Switches, and the bindings in one VDC cannot affect the bindings in another. Cisco NX-OS Software does not limit the binding database size per VDC.

https://t.me/learningnets



Configuration of IP Source Guard is straightforward. The DHCP snooping feature can be enabled globally and it is a prerequisite for the IP Source Guard feature. You can verify IP DHCP snooping binding table entries by issuing the command **show ip dhcp snooping binding**.

IP Source Guard is supported on the Cisco Nexus 7000 Series and the Cisco Nexus 5000 and 5500 Platform switches and is enabled on a per-interface basis using **ip verify source dhcp-shooping-vlan** command.

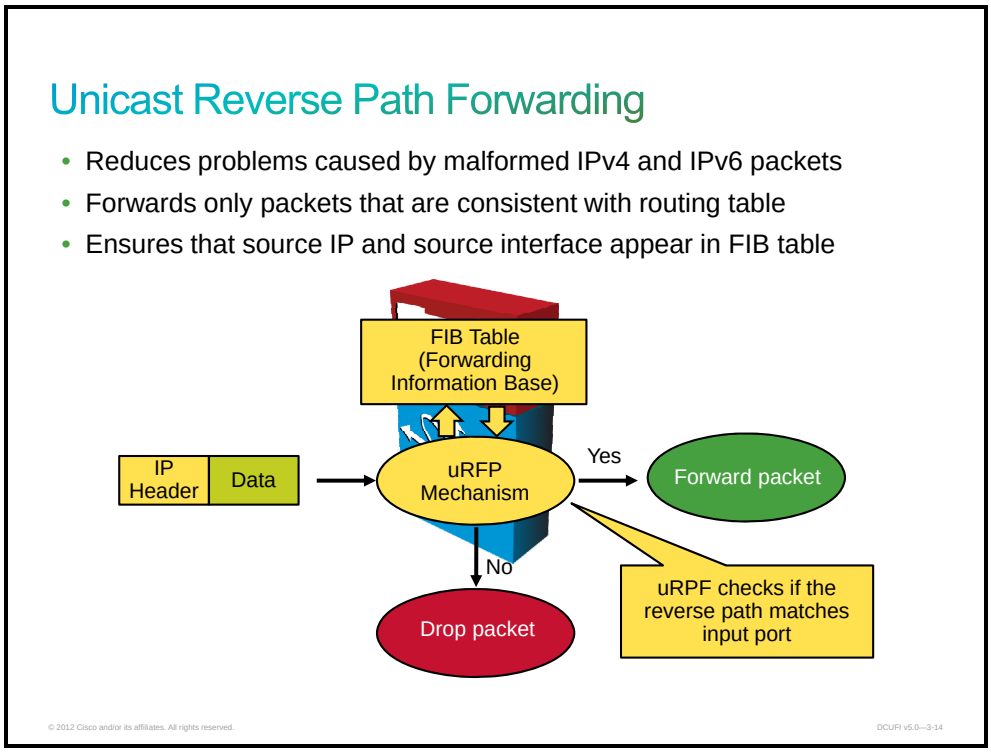
For complete list of configuration commands refer to *Cisco Nexus 7000 Series NX-OS Security Configuration Guide, Release 6.x* at http://www.cisco.com/en/US/docs/switches/datacenter/sw/6_x/nx-os/security/configuration/guide/b_Cisco_Nexus_7000_NX-OS_Security_Configuration_Guide__Release_6.x.html.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Unicast Reverse Path Forwarding

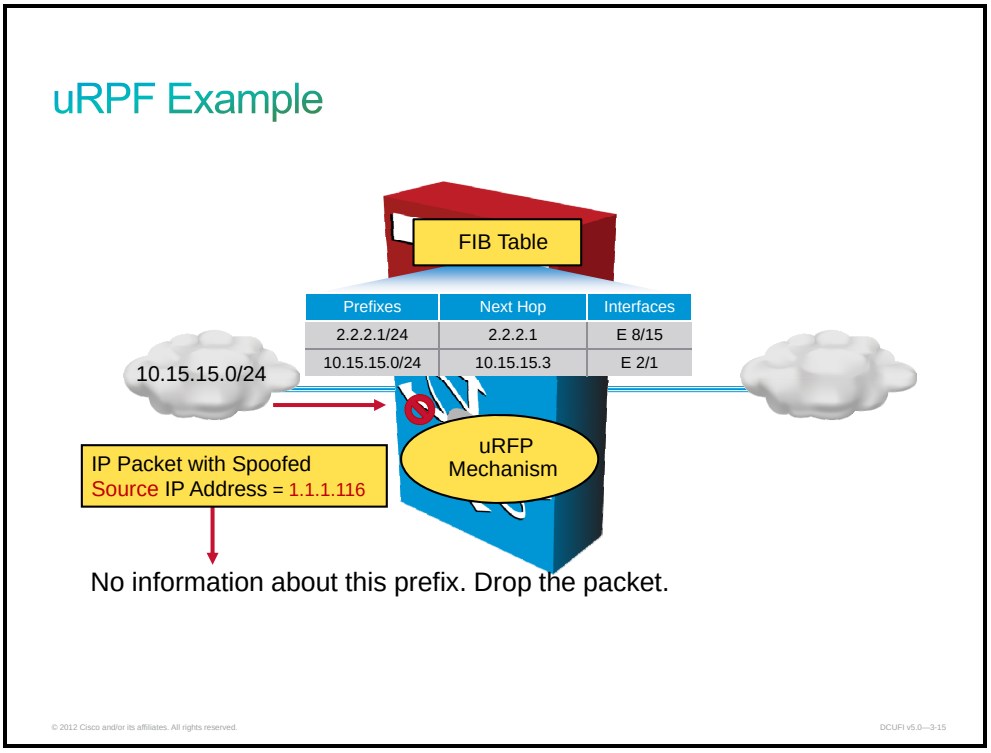
This topic describes how to configure Unicast Reverse Path Forwarding (uRPF) on the Cisco Nexus Family of switches and verify the configuration.



The uRPF feature reduces problems that are caused by the introduction of malformed or forged (spoofed) IP version 4 (IPv4) or IP version 6 (IPv6) source addresses into a network. It does this by discarding IPv4 or IPv6 packets that lack a verifiable IP source address.

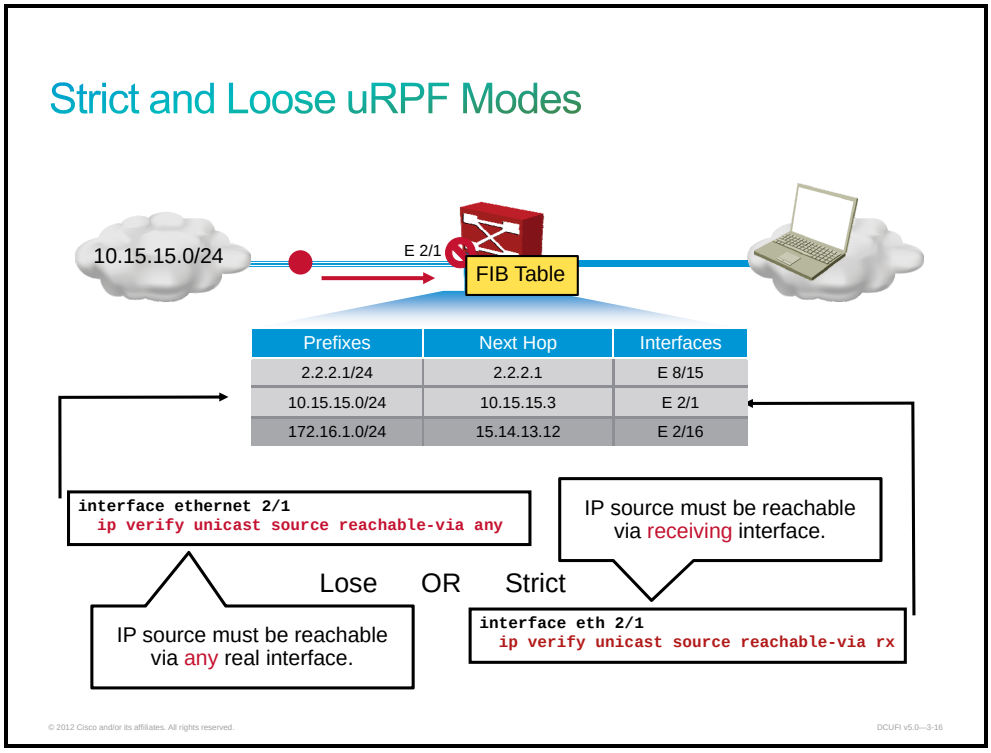
The uRPF checks to see if a packet that is received on a switch interface arrives on the best return path (return route) to the source of the packet. The uRPF accomplishes this check by doing a reverse lookup. If the packet was received from one of the best reverse path routes, the packet is forwarded as normal. If there is no reverse path route on the same interface from which the packet was received, the source address may have been modified. If the uRPF does not find a reverse path for the packet, the packet is dropped or forwarded, depending on whether an ACL is specified in the interface configuration.

https://t.me/learningnets



In the example shown in the diagram, the packet with the source IP address 1.1.1.16 is received on E2/1 interface. The uRPF mechanism checks in the forwarding information base (FIB) table to see if there is a return route to 1.1.1.16 through the E2/1 interface. If there is a matching path, the packet is forwarded. Otherwise, it is discarded.

© 2012 Cisco Systems, Inc.
https://t.me/learningnets



Unicast RPF can be configured in two modes on an ingress interface:

- **Strict uRPF mode**: This mode checks to verify that the packet source address and the ingress interface through which the packet is received matches one of the uRPF interfaces in the FIB match. If this check fails, the packet is dropped. You can use the uRPF mode to check where you expect packet flows to be symmetrical.
- **Loose uRFP mode**: this mode checks the packet source address in the FIB, and returns a match. If the FIB lookup result indicates that the source is reachable through at least one real interface, the packet is forwarded. The ingress interface through which the packet is received is not required to match any of the interfaces in the FIB result.

For complete list of configuration commands refer to *Cisco Nexus 7000 Series NX-OS Security Configuration Guide, Release 6.x* at http://www.cisco.com/en/US/docs/switches/datacenter/sw/6_x/nx-os/security/configuration/guide/b_Cisco_Nexus_7000_NX-OS_Security_Configuration_Guide__Release_6.x.html.

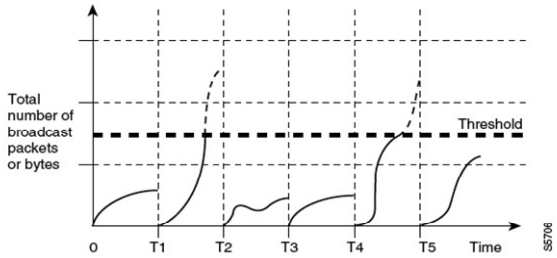
https://t.me/learningnets

Traffic Storm Control

This topic describes how to configure traffic storm control on the Cisco Nexus Family of switches and verify the configuration.

Traffic Storm Control

- A traffic storm occurs when packets flood the LAN.
- Negative effects: excessive traffic, poor network performance.
- Traffic storm control mechanism monitors levels of:
  - Unicast
  - Broadcast
  - Multicast
- When the ingress traffic reaches configured threshold on the port, traffic storm control drops the traffic until the interval ends.



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

A traffic storm is a state of the network that occurs when packets flood the LAN, creating excessive network traffic and degrading network performance. Traffic storm control is used to prevent disruptions on Layer 2 ports by a broadcast, multicast, or unicast traffic storm on physical interfaces.

Thresholds are used for the traffic storm control mechanism to limit the number of packets that are generated over the time interval. Time interval and threshold allow this mechanism to work with different levels of granularity. A higher level of threshold allows more packets to pass through.

Traffic storm control on a switch in the Cisco Nexus Family of switches is implemented in the hardware. The traffic storm control mechanism monitors packets that pass from a Layer 2 interface to the switching fabric. The mechanism determines if the packet is unicast or broadcast, tracks the current count of packets within the 1-second interval, and filters out subsequent packets when a threshold is reached.

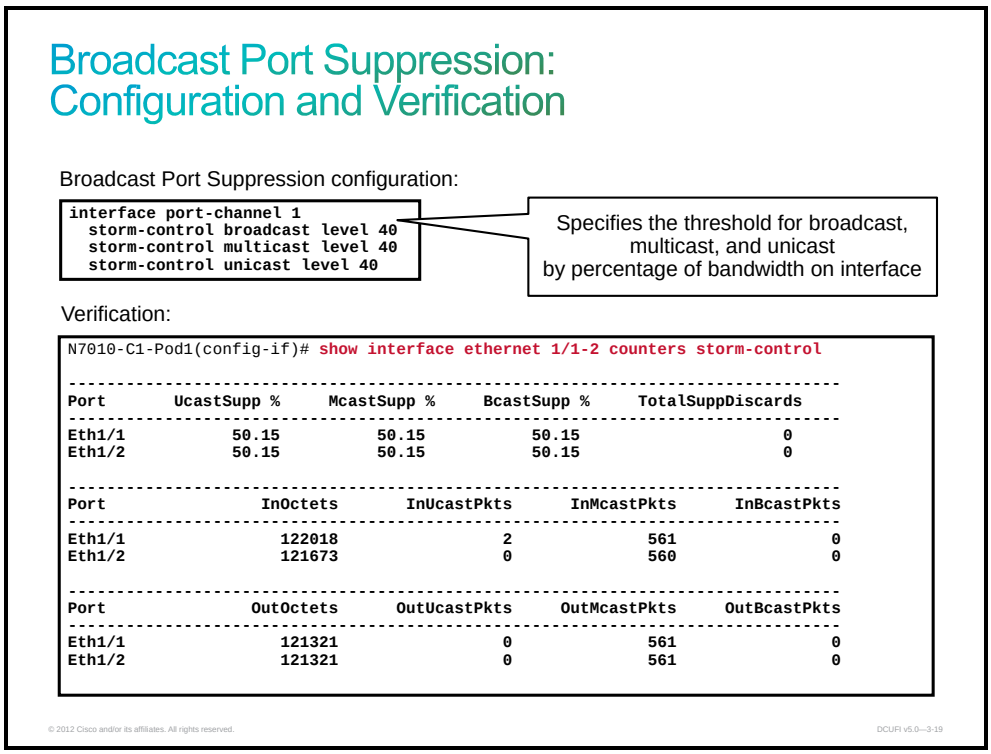
Traffic storm control uses a method that measures bandwidth to measure traffic load. When you configure the threshold, you set the percentage of available total bandwidth that the controlled traffic can use.

These are some examples of traffic storm control:

- If the broadcast traffic storm control is enabled, and broadcast traffic exceeds the level within the 1-second interval, traffic storm control drops all broadcast traffic until the end of that interval.
- If both broadcast and multicast traffic storm control mechanisms are enabled, and the sum of broadcast and multicast traffic exceeds the level within the 1-second interval, traffic storm control drops all broadcast and multicast traffic until the end of that interval.

https://t.me/learningnets

- If both broadcast and multicast traffic storm control is enabled, and broadcast traffic exceeds the level within the 1-second interval, traffic storm control drops all broadcast and multicast traffic until the end of that interval.
- If both broadcast and multicast traffic storm control is enabled, and multicast traffic exceeds the level within the 1-second interval, traffic storm control drops all broadcast and multicast traffic until the end of that interval.



You need to be aware of a few guidelines and limitations when configuring traffic storm control:

- Traffic storm control can be configured on a port-channel interface.
- Traffic storm control must *not* be configured on interfaces that are members of a port-channel interface, because the interfaces will fall into a suspended state.
- Level is configured as a percentage of the total interface bandwidth:
  - The level can be in the range of 0 to 100 percent.
  - 100 percent means no traffic storm control.
  - 0 percent drops all ingress traffic.

Because of hardware limitations and the method by which packets of different sizes are counted, the level percentage is an approximate value. Depending on the sizes of the frames arriving at the port, the actual configured level might differ from the configured level by several percentage points.

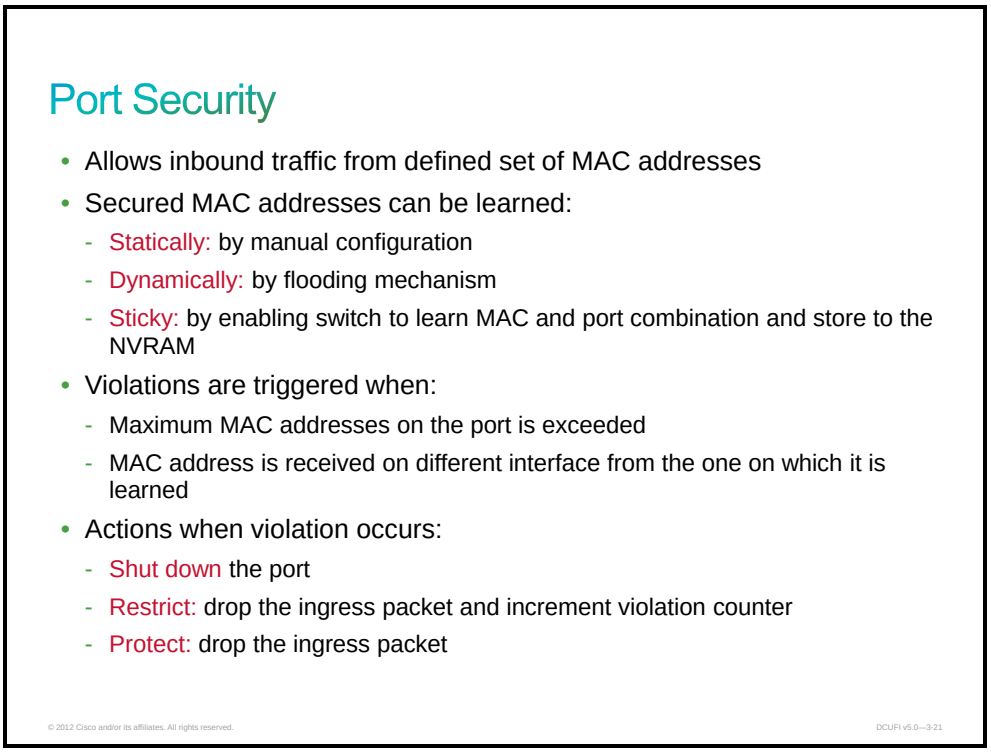
In the example shown, the same threshold level of 40 percent is configured for all of three types: unicast, broadcast, and multicast.

For complete list of configuration commands refer to *Cisco Nexus 7000 Series NX-OS Security Configuration Guide, Release 6.x* at http://www.cisco.com/en/US/docs/switches/datacenter/sw/6_x/nx-os/security/configuration/guide/b_Cisco_Nexus_7000_NX-OS_Security_Configuration_Guide__Release_6.x.html.

https://t.me/learningnets

Port Security

This topic describes how to configure port security on the Cisco Nexus Family of switches and verify the configuration.



Port security allows you to configure physical or port channel interfaces that allow inbound traffic from only a restricted set of MAC addresses. The MAC addresses in the trusted MAC address set are called secure MAC addresses.

Restricting which interface a MAC address is associated with also means that that MAC address is not permitted on another interface within the same VLAN.

Secure MAC addresses are configured per-interface and can be learned in these ways:

- **Statically:** Manually configures the association of a MAC address with an interface.
- **Dynamically:** Enables port security and permits the device to learn the MAC address-to-port combination. This method does not maintain the dynamic address in the table if the device is restarted.
- **Sticky:** Enables port security and permits the device to learn the MAC address-to-port combination, but store the entry in NVRAM so that it is maintained during a device restart.

If a violation occurs, the device can be configured to perform an action against that packet. Violation conditions include the following:

- MAC addresses at the port exceeds the maximum threshold
- MAC address arrives on an interface in the same VLAN, but on a different interface from the one on which it was learned

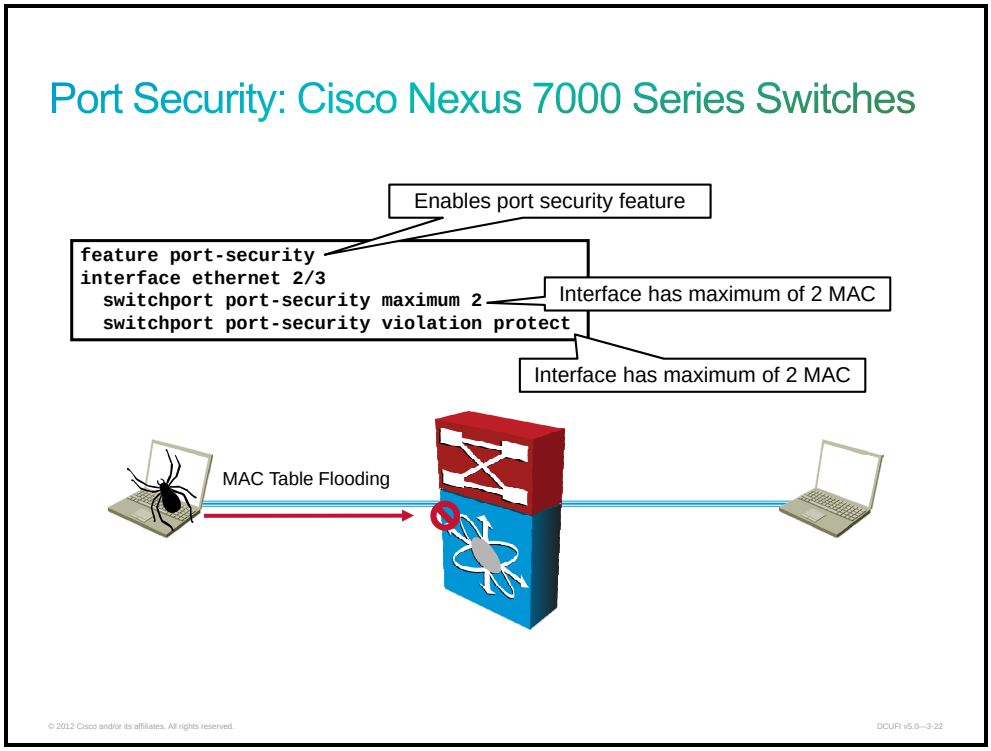
© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Violation actions include the following:

- **Shutdown:** Put the interface into an error disabled state. Shutdown is the default action.
- **Restrict:** Drop the ingress traffic and maintains a count of dropped packets.
- **Protect:** Continue learning until threshold is reached, and then disable learning and drop all ingress traffic from nonsecure MAC addresses.

Port security can be configured on the same interface as IEEE 802.1X, where port security secures the MAC addresses that 802.1X authenticates.



When configuring port security, first of all, you need to enable the port security feature on Cisco NX-OS Software.

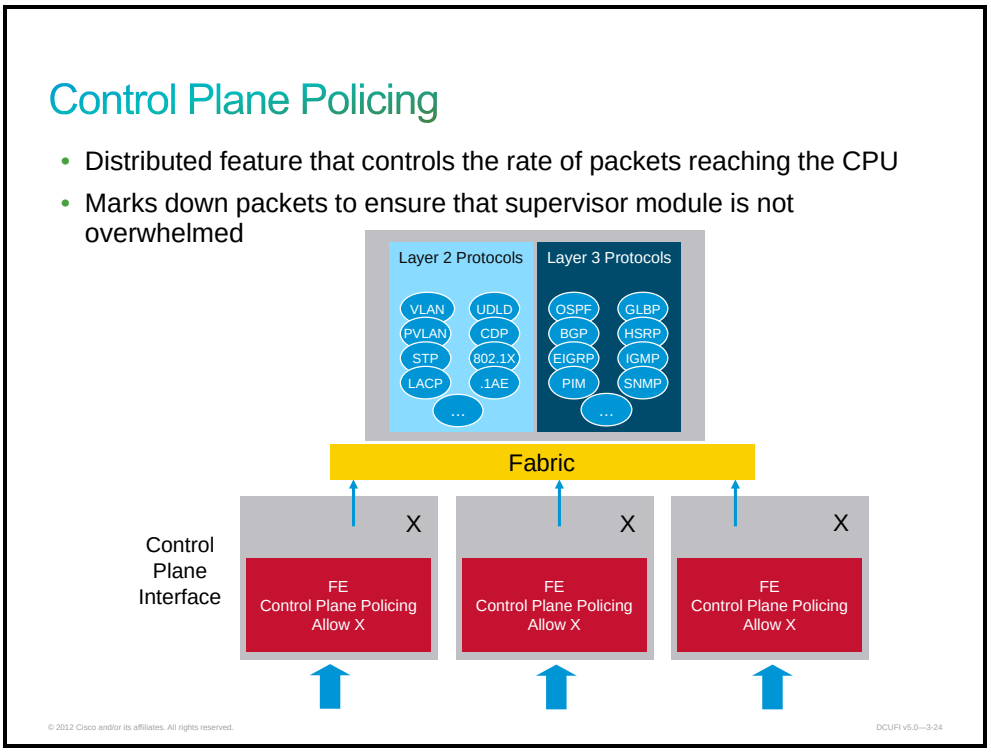
In the example shown, port security is configured on Ethernet 2/3 port. Also, two MAC addresses are allowed to be sources of traffic on the port, and this setting is configured by the **switchport port-security maximum 2** command. The violate action is in this case *protect*, and can also be *restrict*. The only difference between these two violate actions are that violation counters are incremented in **restrict** command, while they are not incremented in **protect** command.

For complete list of configuration commands refer to *Cisco Nexus 7000 Series NX-OS Security Configuration Guide, Release 6.x* at http://www.cisco.com/en/US/docs/switches/datacenter/sw/6_x/nx-os/security/configuration/guide/b_Cisco_Nexus_7000_NX-OS_Security_Configuration_Guide__Release_6.x.html.

https://t.me/learningnets

Control Plane Policing

This topic describes how to configure Control Plane Policing (CoPP) on the Cisco Nexus Family of switches and verify the configuration.



CoPP protects the control plane and separates it from the data plane, ensuring network stability, reachability, and packet delivery. The Cisco NX-OS device provides CoPP to prevent denial of service (DoS) attacks from impacting performance.

The supervisor module divides the traffic that it manages into three functional components, or planes, as follows:

- **Data plane:** Manages all data traffic
- **Control plane:** Manages all routing protocol control traffic
- **Management plane:** Manages all components that are associated with device management

The management and control planes are on the supervisor module, and both are critical to the operational capabilities of the network and device. If there are any disruptions or attacks against the supervisor module, the network can be seriously affected.

CoPP is a distributed, hardware-based feature that controls the rate at which packets are allowed to reach the CPU. By segregating the traffic into different classes, Cisco NX-OS Software is able to apply device policies or mark down packets to ensure that the supervisor module is not overwhelmed.

- Exception logic separates data plane and control plane packets.
- Classification identifies DoS attack packets.
- Service policies mark, drop, or police.
- Independent control-plane interface.
- Configuration is MQC-based

On the Cisco Nexus 7000 Series Switches, CoPP can be configured only in the default VDC, with the CoPP configuration being applied to all VDCs on the Cisco NX-OS device.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

CoPP Enhancements on Cisco Nexus 7000

Enhancements to CoPP:

- Multicast traffic
- ARP packets
- Layer 2 broadcast packets
- IP unicast with DMAC
- Certain packets redirected to the CPU
- Matching of packets generating exceptions and redirections
- Configurable policy map for packets per second

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

The Cisco Nexus 7000 Series Switches introduce new enhancements for CoPP:

- Multicast traffic, preserving the bridging on the ingress VLAN
- ARP packets, including ARP broadcast storms
- Layer 2 broadcast packets
- IP unicast with dynamic MAC address having I/G bit set
- Packets that are redirected to CPU:
  - Packets with IP options set
  - DHCP snooping, DAI, port-security packets

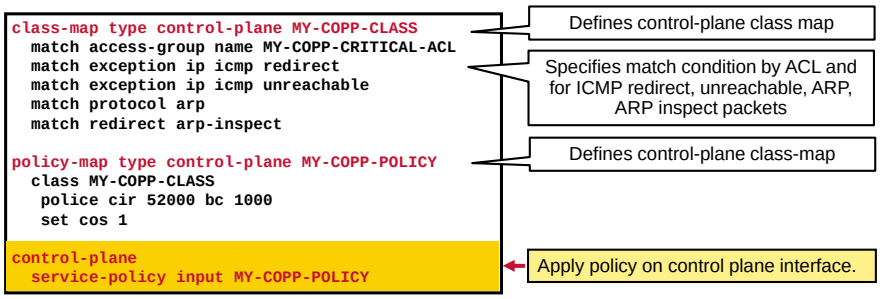
CoPP supports IPv4, IPv6, ARP, and MAC ACLs, and can match packets generating exceptions and redirections. The rate in the policy map can be configured as packets per second (p/s), but statistics will still be shown in bytes per second (B/s).

The implementation of CoPP in the Cisco Nexus Family devices is strictly hardware-based.

https://t.me/learningnets

CoPP Configuration

- Control plane class-map type must be used.
- Policing must be configured in control plane policy-map.
- CoPP can be modified only in default VDC.
- Beginning with Cisco NX-OS Release 5.2, the CoPP best practice (default) policy is read-only. If you want to change policy, you must copy policy and then apply changes to copy.



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

The first step in CoPP configuration is creation of the control plane type class-map. You can use system-defined class map names, but you cannot use **class-default**, **match-all**, or **match-any** as class map names.

In the example shown in the figure, MY-COPP-CLASS is defined. The match criteria can be ACL or type of packet, as shown in the example.

The policy map for CoPP includes policing parameters. If the policer for the class is not defined, the default action is drop.

In the example shown, policing is performed on MY-COPP-CLASS with policing parameters of 52000 b/s and burst size of 1000 bytes.

Finally, the policy map MY-COPP-POLICY is applied on the control plane.

You can configure CoPP only in the default virtual device context (VDC), but the CoPP configuration applies to all VDCs on the Cisco NX-OS device.

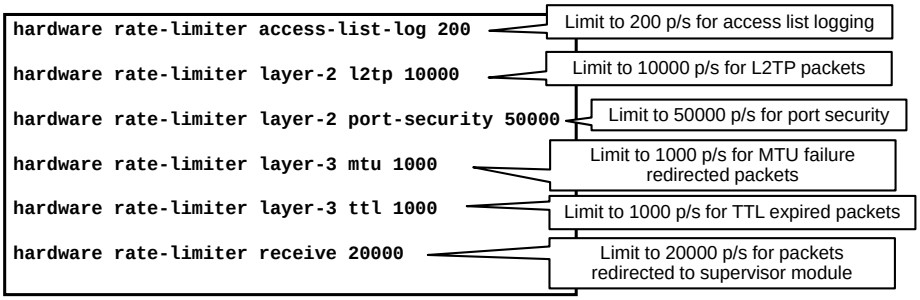
Beginning with Cisco NX-OS Software Release 5.2, the CoPP best practice policy is read-only. If you want to modify its configuration, you must copy it. Copied policies are treated as user configurations.

For complete list of configuration commands refer to *Cisco Nexus 7000 Series NX-OS Security Configuration Guide, Release 6.x* at http://www.cisco.com/en/US/docs/switches/datacenter/sw/6_x/nx-os/security/configuration/guide/b_Cisco_Nexus_7000_NX-OS_Security_Configuration_Guide__Release_6.x.html.

https://t.me/learningnets

CoPP: Cisco Nexus 7000 Rate Limiters

- Rate limiters prevent redirected packets from overwhelming the supervisor module.
- You can define limits per type of packet.
- Configuration limits the number of packets per second (p/s).



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

Rate limiters prevent redirected packets from overwhelming the supervisor engine. When the number of the packet limit that is defined in the Rate Limiter configuration for the specific type of packet has been reached within a 1-second time interval, the subsequent packets for that type will be dropped.

In the example shown, the number of packets per 1-second interval are set as the limit for the following type of packets:

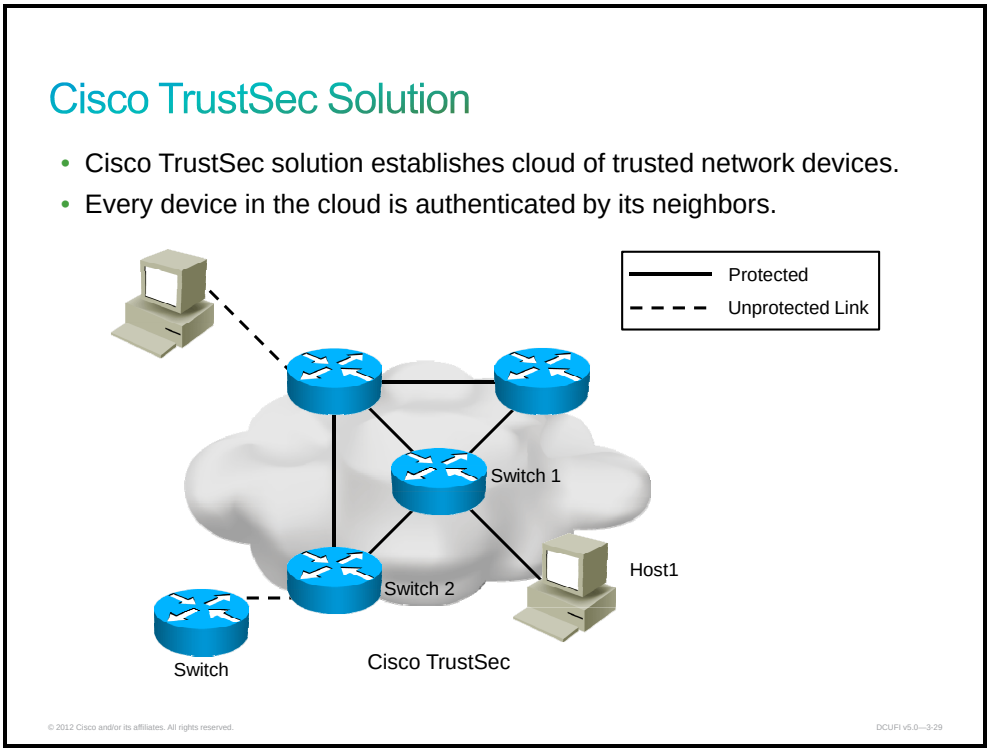
- ACL logging
- Layer 2 Tunneling Protocol (L2TP) packets
- Port security packets
- Maximum transmission unit (MTU) failure packets
- Time to Live (TTL) Expired packets
- Redirected packets to supervisor module

For complete list of configuration commands refer to *Cisco Nexus 7000 Series NX-OS Security Configuration Guide, Release 6.x* at http://www.cisco.com/en/US/docs/switches/datacenter/sw/6_x/nx-os/security/configuration/guide/b_Cisco_Nexus_7000_NX-OS_Security_Configuration_Guide__Release_6.x.html.

https://t.me/learningnets

Cisco TrustSec

This topic describes Cisco TrustSec architecture on the Cisco Nexus Family of switches.



The Cisco TrustSec security architecture builds secure networks by establishing clouds of trusted network devices. Every device in the cloud is authenticated by its neighbors. Communication on the links between devices in the cloud is secured with a combination of encryption, message integrity checks, and data-path replay protection mechanisms.

Major components of the Cisco TrustSec architecture include the following:

- **Authentication:** Verification of a device before permitting it to join the Cisco TrustSec network
- **Authorization:** Level of access the device has to the Cisco TrustSec network resources
- **Access control:** Access policies that are applied on a per-packet basis using the source tags on each packet
- **Secure communication:** Encryption, integrity, and data-path reply protection for packets flowing over links within the Cisco TrustSec network

The three entities within the Cisco TrustSec network include the following:

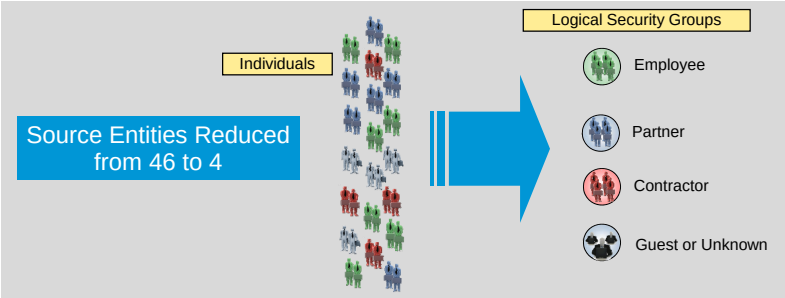
- **Supplicants:** Devices attempting to join the Cisco TrustSec network
- **Authenticators:** Devices already part of the Cisco TrustSec network
- **Authorization server:** Device that provides authentication information, authorization information, or both

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Cisco TrustSec: Security Group Tags

- Cisco TrustSec provides scalable access via Security Group Tags (SGTs).
- The SGTs represent logical groups of users and servers with similar sets of privileges.
- The SGTs are 16-bit (2-byte) tags supporting up to 64 K (65,536) logical groups.
- Per-user SGTs are not recommended because their use would defeat the purpose of logical groups.



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

As packets enter the network, Cisco TrustSec uses device and user identification information for classifying the packets. The classification is maintained by the use of tags on the ingress into the Cisco TrustSec network. These tags allow packets to be properly identified for applying security policies along the path.

The tag is also called the Security Group Tag (SGT). The SGT allows the network to enforce access control policies by enabling the endpoint device to act on the SGT for filtering purposes.

Each security group has a unique 16-bit SGT assigned to it. The number of SGTs in the Cisco NX-OS device is limited to the number of authenticated network entities. The SGT scope is global within a Cisco TrustSec network.

After the authentication process has occurred, packets that originate from that device are tagged with the SGT. That tag is carried throughout the network within the Cisco TrustSec header. This tag is referred to as the source SGT and is applied on the ingress port. The tag informs the egress port which access control policy should be applied.

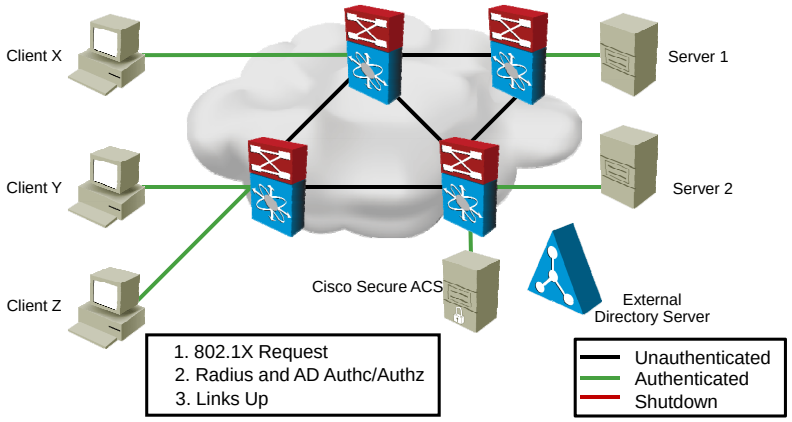
A collection of information or policies that are known as environment data allows a device to function as a Cisco TrustSec cloud, although some of the data might be manually configured on the device. The device uses RADIUS to download the following environment data from the authentication server:

- **Server lists:** List of servers that the client can use for future RADIUS requests
- **Device SGT:** Security group to which this device belongs
- **Expiry timeout:** How often the Cisco TrustSec device should refresh the environment data

https://t.me/learningnets

Cisco TrustSec: Admission Control

- Security Group ACLs (SGACLs) are used to bind roles and permissions based on assigned security group.
- After 802.1X authentication, SGT, and SGACL are applied to the link.
- SGT is applied on ingress switch port.



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

SGACLs (Security Group Access Control Lists) are used to control the operations that users can perform based on assigned security groups. Grouping permissions into a role simplifies the management of the security policy. As users are added to the Cisco NX-OS device, they immediately receive the correct permissions by being placed into one or more security groups. Security groups can be modified as required.

The following sequence of events takes place when a link between a supplicant and an authenticator first comes up:

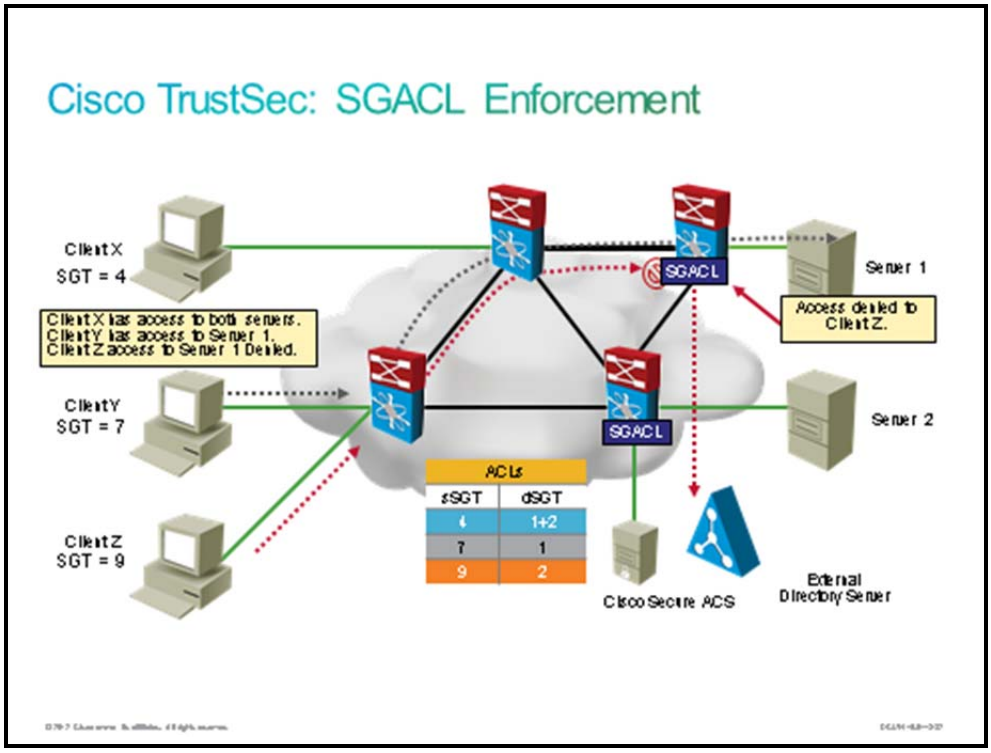
1. Authentication using 802.1X
2. Authorization in which each side obtains policies such as SGT and ACLs that are to be applied to the link
3. Security Association Protocol (SAP) negotiation where the Extensible Authentication Protocol over LAN (EAPOL)-key exchange is used between the supplicant and the authenticator to negotiate a cipher suite, exchange security parameter indexes (SPIs), and manage the keys

When these three events are successful, a security association (SA) is established.

The SGT is applied on the ingress switch port. Security group and role-based ACL policies are applied on SGTs (permit, deny, log, police, re-mark, span, redirect, and so on).

© 2012 Cisco Systems, Inc.

https://t.me/learningnets



The SGT is applied on the ingress switch port and is put in the Cisco TrustSec security header. The tag is maintained throughout the Cisco TrustSec network so that the relevant security policies can be applied to that packet.

In the example shown here, there are five servers in the network with a statically or dynamically assigned SGT. Client X has SGT = 4, Client Y has SGT = 7, and Client Z has SGT = 9. On the other side, Server 1 and Server 2 have SGT = 1 and SGT = 2 respectively.

From the Client X, Y, and Z point of view, servers have different privileges according to the SGACL that is defined on ingress ports on the switches that connect Server 1 and Server 2.

Client X has the privilege to access both Server 1 and Server 2, Client Y has access to only Server 1, and Client Z has access to Server 2.

https://t.me/learningnets

Summary

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

Summary

- DHCP snooping acts like a firewall between untrusted and trusted DHCP servers.
- DAI ensures that only valid ARP requests and responses are relayed.
- IP Source Guard permits IP traffic when the IP address and MAC address of each packet matches entries in a binding database.
- uRPF discards IPv4 and IPv6 packets that lack a verifiable IP source address.
- Traffic storm control is used to prevent disruptions on Layer 2 ports by broadcast, multicast, or unicast packets.
- Port security allows inbound traffic from only a restricted set of MAC addresses.
- CoPP protects the control plane to prevent DoS attacks from impacting performance.
- The Cisco TrustSec solution is a security architecture for building secure networks.

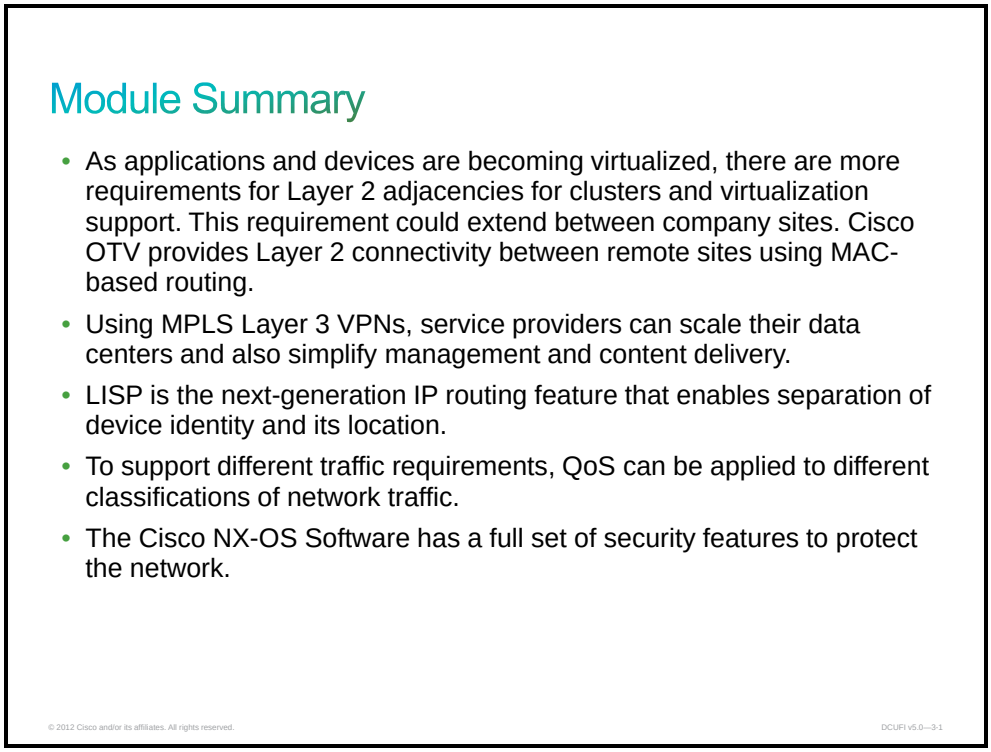
© 2012 Cisco and/or its affiliates. All rights reserved.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Module Summary

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



The Cisco Nexus Operating System (NX-OS) Software supports a comprehensive set of security features that are set to protect the network against degradation or failure and against data loss or compromise. There are five main aspects of the security portfolio—traffic integrity, control plane protection, access control, admission control, and data confidentiality. For businesses requiring secure network infrastructures in their data center, Cisco TrustSec can provide security and confidentiality by using IEEE 802.1X authentication of devices and encryption of traffic.

Cisco Overlay Transport Virtualization (OTV) provides Layer 2 connectivity between remote network sites by using MAC address-based routing and IP-encapsulated forwarding across a transport network. Cisco OTV provides support for applications requiring Layer 2 adjacency, such as clusters and virtualization. Cisco OTV can be deployed on edge devices in each site, and it requires no changes to the sites or transport network.

Quality of service (QoS) classifies network traffic for policing and prioritizing of traffic flows so that traffic congestion can be avoided in the network. It does not provide more bandwidth, but rather manages traffic that is using the bandwidth. Before you apply QoS features, you should understand the application requirements. Classification of traffic is performed through class maps. Policy maps apply the required policies to the classified traffic and service policies apply the policy map to a specific interface or range of interfaces as required.

Multiprotocol Label Switching (MPLS) has been widely adopted by network operators to provide a scalable solution for network segmentation and traffic separation. Using MPLS Layer 3 VPNs, service providers can scale their data centers and also simplify management and content delivery.

Locator/ID Separation Protocol (LISP) is the next-generation IP routing feature that enables separation of device identity and its location. Creation of separate namespaces for endpoint

https://t.me/learningnets

identification and location of an endpoint creates space for greater scalability of the Internet routing system and improved efficiency of multihomed locations.

References

For additional information, refer to these resources:

- Cisco Systems, Inc. *Cisco Nexus 5000 Series NX-OS Layer 2 Switching Configuration Guide, Release 5.x*. San Jose, California, July 2012. http://www.cisco.com/en/US/docs/switches/datacenter/nexus5000/sw/layer2/521_n1_1/b_5k_Layer2_Config_521N11.html
- Cisco Systems, Inc. *Cisco Nexus 5000 Series NX-OS Security Configuration Guide, Release 5.x*. San Jose, California, July 2012. http://www.cisco.com/en/US/docs/switches/datacenter/nexus5000/sw/security/521_n1_1/b_5k_Security_Config_521N11.html
- Cisco Systems, Inc., *Cisco Nexus 7000 Series Command Reference Master Index*. San Jose, California, September 2012. http:/www.cisco.com/en/US/docs/switches/datacenter/sw/6_x/nx-os/master/index/6_xMasterIndex.html
- Cisco Systems, Inc., *Cisco Nexus 7000 Series NX-OS LISP Command Reference*. San Jose, California, September 2012. http://www.cisco.com/en/US/docs/switches/datacenter/sw/6_x/nx-os/lisp/command/reference/preface.html
- Cisco Systems, Inc. *Cisco Nexus 7000 Series NX-OS OTV Configuration Guide*. San Jose, California, June 2012. http://www.cisco.com/en/US/docs/switches/datacenter/sw/nx-os/OTV/config_guide/b_Cisco_Nexus_7000_Series_NX-OS_OTV_Configuration_Guide.html
- Cisco Systems, Inc. *Cisco Nexus 7000 Series NX-OS Security Configuration Guide, Release 6.x*. San Jose, California, August 2012. http://www.cisco.com/en/US/docs/switches/datacenter/sw/6_x/nx-os/security/configuration/guide/b_Cisco_Nexus_7000_NX-OS_Security_Configuration_Guide__Release_6.x.html

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

Module Self-Check

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

Q1) Which feature can be configured using a configuration session? (Source: Configuring Security Features)

A) access control lists
B) all security features
C) DHCP snooping
D) port security

Q2) Which security feature can restrict inbound traffic from a specific set of MAC addresses known as secure MAC addresses? (Source: Configuring Security Features)

A) DHCP snooping
B) Dynamic ARP Inspection
C) port security
D) IP Source Guard

Q3) You are implementing DHCP snooping and the support of metropolitan Ethernet-access environments where DHCP centrally manages the IP addresses for many subscribers.

Which option should be supported on the DHCP server so that it permits the DHCP server to assign an IP address that is based on any parameters that are configured? (Source: Configuring Security Features)

A) 80
B) 81
C) 82
D) 83

Q4) Which two disadvantages are disadvantages of traditional data center interconnect technologies? (Choose two.) (Source: Understanding Overlay Transport Virtualization)

A) overhead associated with additional headers
B) complex operations
C) low bandwidth
D) transport-dependent

Q5) Match the technology to the characteristics. (Source: Understanding Overlay Transport Virtualization)

A) Cisco OTV
B) Traditional DCI technologies

_____ 1. MAC address learning based on flooding

_____ 2. control plane-based MAC learning

_____ 3. dynamic encapsulation

_____ 4. pseudowires and tunnels

_____ 5. complex dual-homing

_____ 6. native automated dual-homing

https://t.me/learningnets

Q6) Which statement best describes a Cisco OTV overlay interface? (Source: Understanding Overlay Transport Virtualization)

A) An overlay interface is one of the uplinks of the Cisco OTV edge device.
B) An overlay interface is one of the interfaces on the edge device that faces the site and carries at least one of the VLANs that are extended through Cisco OTV.
C) An overlay interface is a new virtual interface that contains the entire Cisco OTV configuration.
D) An overlay interface is a special type of Layer 3 interface that sends and receives packets on a virtual IP-based ring.

Q7) Cisco OTV edge devices can perform fragmentation to allow jumbo frames to be carried across an IP transport network with a smaller MTU. (Source: Understanding Overlay Transport Virtualization)

A) true
B) false

Q8) Which protocol is used as the Cisco OTV control protocol? (Source: Understanding Overlay Transport Virtualization)

A) OSPF
B) IS-IS
C) BGP
D) MSDP
E) CFS

Q9) Which five commands need to be configured on the Cisco OTV overlay interface for Cisco OTV to be fully operational? (Choose five.) (Source: Understanding Overlay Transport Virtualization)

A) **otv join-interface interface ethernet 1/1**
B) **otv site-vlan 200**
C) **otv extend-vlan 100-199**
D) **otv control-group 239.1.1.1**
E) **otv isis group 239.1.1.1**
F) **otv ssm-groups 232.1.1.0/24**
G) **otv data-group 232.1.1.0/24**
H) **ip igmp version 3**
I) **ip pim sparse-mode**
J) **no shutdown**

Q10) Which entity is part of the Cisco TrustSec network? (Source: Configuring Security Features)

A) local switch
B) supplicant
C) authenticator
D) authorization server

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Q11) When you are implementing QoS, which type of traffic should be classified so that no frames are dropped due to its exchange-based behavior? (Source: Configuring Quality of Service)

A) SCSI I/O
B) file transfer
C) SMTP traffic
D) voice

Q12) When policing traffic on a Cisco Nexus 7000 Series switch, how many colors or conditions can be applied? (Source: Configuring QoS)

A) 1
B) 2
C) 3
D) 4

Q13) What is the default type of class-map when no specific type is specified? (Source: Configuring QoS)

A) type policing
A) type queuing
B) type scheduling
C) type QoS

Q14) Which statement is correct regarding mutation mapping? (Source: Configuring QoS)

A) On ingress, mutation mapping occurs after traffic classification and all other actions.
B) On egress, mutation mapping occurs after traffic classification and all other actions.
C) On ingress, mutation mapping occurs before traffic classification and after all other actions.
D) On egress, mutation mapping occurs after traffic classification and before the other actions.

Q15) Which type of policy map is used to mark, shape and queue packets? (Source: Configuring QoS)

A) type policing
B) type queuing
C) type scheduling
D) type QoS

Q16) Which two forwarding bases are used in MPLS data plane? (Choose two.) (Source: Configuring MPLS)

A) routing table
B) FIB
C) LFIB
D) topology table

Q17) LDP is used for label allocation and distribution in MPLS network. (Source: Configuring MPLS)

A) True
B) False

https://t.me/learningnets

Q18) Which Cisco Nexus card has MPLS functionality support on it? (Source: Configuring MPLS)

A) All M I/O series of line card
B) F1 cards
C) F2 cards
D) FCoE-based cards

Q19) Which three commands must be used to install the MPLS feature set on Cisco Nexus switches? (Choose three.) (Source: M03L02 – Configuring MPLS)

A) **install ldp**
B) **install feature-set mpls**
C) **feature-set mpls**
D) **use mpls feature-set**
E) **feature mpls ldp**
F) **install mpls-l3vpn**

Q20) Which two benefits are benefits of using LISP? (Choose two.) (Source: Configuring LISP)

A) mobility
B) reduction of BGP table
C) optimization
D) scalability

Q21) What is the RLOC? (Source: Configuring LISP)

A) IPv4 or IPv6 addresses that are used in the source or destination address fields in the inner LISP header
B) IPv4 or IPv6 address of the customer router that de-encapsulates a LISP packet
C) a block of IP addresses that are assigned to a site by each service provider
D) router that accepts IP packets and performs encapsulation into the LISP header and performs an EID-to-RLOC mapping lookup

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Module Self-Check Answer Key

Q1) A

Q2) C

Q3) C

Q4) B, D

Q5) 1-B
2-A
3-A
4-B
5-B
6-A

Q6) C

Q7) B

Q8) B

Q9) A, C, D, G, J

Q10) C

Q11) A

Q12) C

Q13) D

Q14) D

Q15) B

Q16) B, C

Q17) A

Q18) A

Q19) B, C, E

Q20) A, B

Q21) B

https://t.me/learningnets

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Module 4

Cisco Nexus Storage Features

Overview

The Cisco Nexus switches provide a wide range of Fibre Channel and Fibre Channel over Ethernet (FCoE) switching features. The Fibre Channel networks use a set of frames, protocols, and configurable flow control mechanisms. You will be able to explain the various protocols that are used with Fibre Channel. The N-Port Virtualization (NPV) and N-Port ID Virtualization (NPIV) features help implement a scalable environment with virtual server applications and address the problem of blade switch explosion.

The FCoE protocol is extensively used in the data center environments. The FCoE requires that converged network adapters (CNAs) and the FCoE protocol software stack are used on the hosts that are attached to the network. Alongside FCoE, the FCoE Initialization Protocol (FIP) is used to initialize communications.

To carry lossless FCoE traffic over the Ethernet infrastructure, a set of Data Center Bridging (DCB) enhancements have been designed for Cisco Unified Fabric I/O consolidation. Some of the parameters are configurable on Cisco Nexus switches.

In this module, you will learn how to implement FCoE on Cisco Nexus switches in various scenarios, including Enhanced virtual port channel (vPC), Cisco Adapter Fabric Extender (FEX), and FCoE NPV solutions.

Module Objectives

Upon completing this module, you will describe the Fibre Channel Protocol (FCP), the FCoE protocol, and the DCB enhancements. In addition, you will be able to configure FCoE on the Cisco Nexus 5000 Series Switches, including the Cisco Nexus 5000 and 5500 Platform switches, the Cisco Nexus 7000 Series Switches and Cisco MDS Family of switches. This ability includes being able to meet these objectives:

- Describe the FCP, including Fibre Channel addressing, flow control, and zoning
- Explain the FCoE protocol, the adapters, software stack, and FIP and FCoE configuration
- Identify the relevant DCB enhancements that are required to support I/O consolidation in the data center
- Configure the FCoE protocol on a Cisco Nexus switch.
- Configure the relevant SAN switching features on the Cisco Nexus and Cisco MDS Families of switches to support the design requirements.

https://t.me/learningnets

- Configure the Cisco Nexus and Cisco MDS Families of switches to support NPIV, NPV, and FCoE NPV based on a given implementation plan
- Identify the SAN management tools that can be used to manage the Cisco Nexus 5000 and 7000 Series Switches and the Cisco MDS 9000 Series Multilayer Switches

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Lesson 1

Describing Fibre Channel Protocol

Overview

In this lesson, you will learn about Fibre Channel networks, frames, and flow control. You will be able to explain the various protocols that are used with Fibre Channel. You will learn how to use N-Port Virtualization (NPV) and N-Port ID Virtualization (NPIV).

Objectives

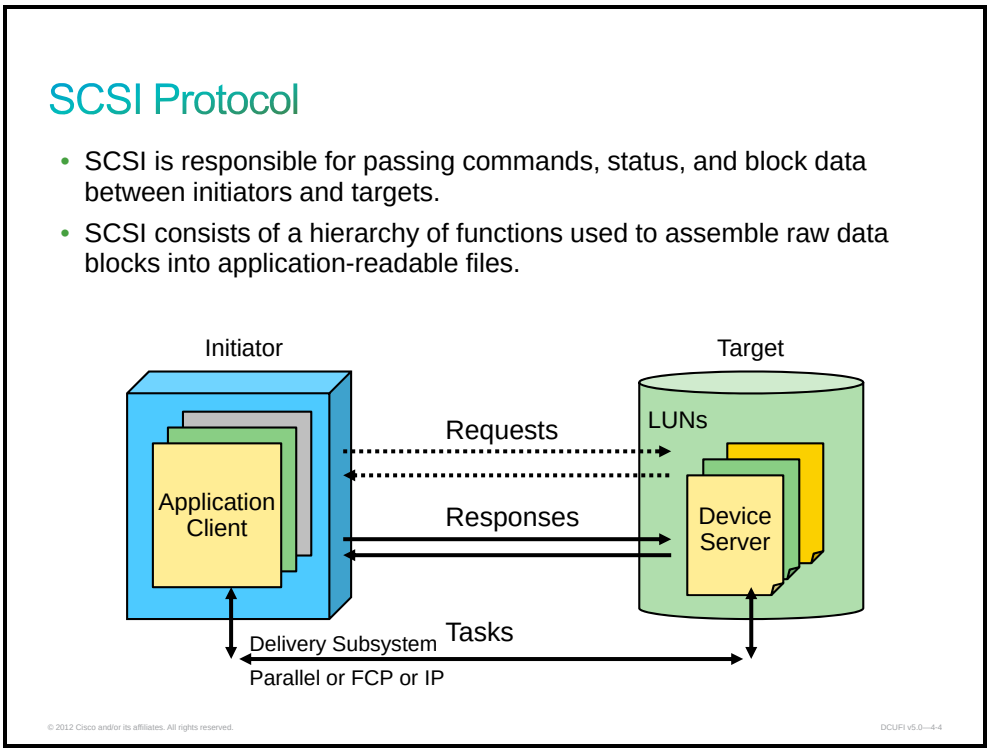
Upon completing this lesson, you will be able to describe the Fibre Channel Protocol (FCP), including Fibre Channel addressing, flow control, and zoning. This ability includes being able to meet these objectives:

- Explain the function of the SCSI protocol in a storage environment and its operation and limitations
- Describe the Fibre Channel technology
- Explain how a device performs a FLOGI on a Fibre Channel SAN
- Explain the addressing scheme used with the FCP
- Explain how VSANs and IVR are used within a Fibre Channel SAN
- Explain how the FSPF protocol is used on the SAN
- Explain how zoning is used within a Fibre Channel SAN
- Determine when and how to use NPIV
- Determine when and how to use the NPV mode

https://t.me/learningnets

SCSI Protocol

This topic explains the function of the Small Computer System Interface (SCSI) protocol in a storage environment and its operation and limitations.



SCSI is responsible for passing commands, status information, and block data between platforms and storage devices. One function of operating systems is to hide the complexity of the computing environment from the end user. Management of system resources, including memory, peripheral devices, display, and context switching between concurrent applications is generally concealed behind the user interface. For I/O operations for peripheral devices such as disk, tape, optical storage, printers, and scanners, these functions are provided by the SCSI protocol. This protocol is embedded in a device driver or a host adapter.

Because the SCSI protocol layer sits between the operating system and the peripheral resources, it has different functional components. Applications typically access data as files or records. Although these may be stored on disk or tape media in the form of data blocks, retrieval of the file requires a hierarchy of functions. These functions assemble raw data blocks into a coherent file that can be manipulated by an application.

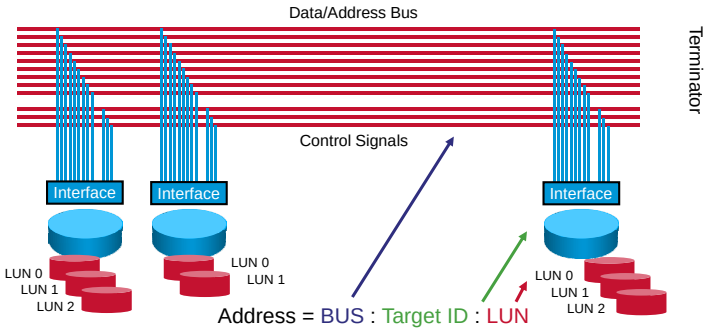
The SCSI architecture defines the relationship between initiators (hosts) and targets (for example, disks) as a client/server exchange. The application client resides in the host and represents the upper layer application, file system, and operating system I/O requests. The server sits in the target device, responding to requests.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

SCSI Multidrop Topology

- SCSI uses a parallel bus architecture:
  - Data sent simultaneously over multiple wires
  - Half-duplex—data travels in one direction at a time
- Device must assume exclusive control over the bus to communicate:
  - Term “simplex” channel—one device can transmit at a time



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

All of the devices on a SCSI bus are connected to a single cable. This setup is called a multidrop topology and has the following characteristics:

- Data bits are sent in parallel on separate wires.
- Control signals are sent on a separate set of wires.
- Only one device at a time can transmit. The device must have exclusive use of the bus.
- A special circuit that is called a terminator must be installed at the end of the cable. The cable must be terminated to prevent unwanted electrical effects from corrupting the signal.
- A multidrop topology has inherent limitations, including the following:
  - Parallel transmission of data bits allows more data to be sent in a given time period, but complicates transmitter-receiver synchronization.
  - Incorrect cable termination can cause intermittent, difficult-to-trace errors.

SCSI was designed to support a few devices at most. SCSI devices use hard addressing:

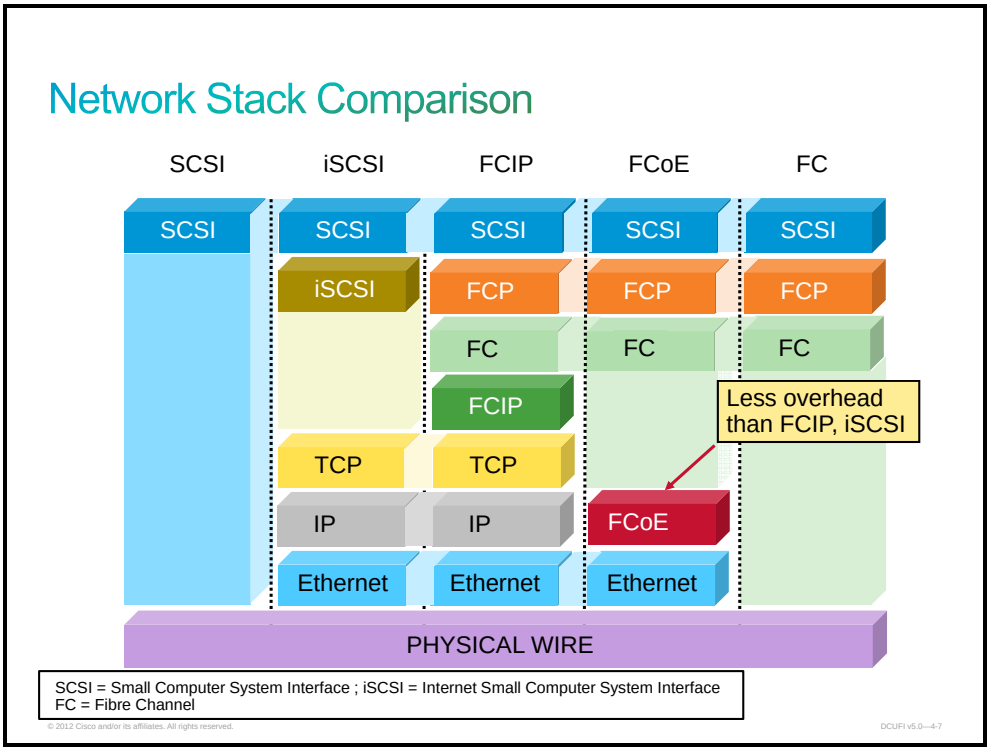
- Each device has a series of jumpers that determine the physical address of the device, or SCSI ID. The ID is software-configurable on some devices.
- Each device must have a unique ID. Before adding a device to the cable, the administrator must know the ID of every other device that is connected to the cable. The administrator must select a unique ID for this new device.
- The ID of each device determines its priority on the bus. For example, the SCSI target with ID 7 always has a higher priority than the SCSI initiator with ID 6. Because a transmitting device must have exclusive use of the bus, ID 6 must wait until ID 7 has finished.

SCSI defines the bus, target, and logical unit number (LUN) triad. The bus represents one of several potential SCSI interfaces that are installed in the host, each supporting a separate string of disks. The target represents a single disk controller on the string. The LUN designation allows for additional disks that are governed by a controller (for example, a Redundant Array of Independent Disks [RAID] device).

https://t.me/learningnets

Fibre Channel

This topic describes the Fibre Channel technology.



The figure shows different elements of competing network stacks. While all stacks depend on SCSI mechanism 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 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 SCSI is recovering from a loss. Because Fibre Channel carries SCSI, it inherits the requirement for an underlying lossless network.

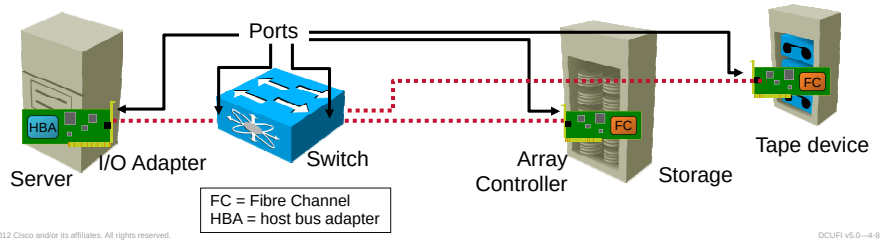
The ability of Fibre Channel over Ethernet (FCoE) to work seamlessly with existing infrastructure makes it an evolutionary technology, one that data centers can deploy at the pace and to the extent that best serves their needs. FCoE allows IP and Fibre Channel network traffic to be carried over existing FCoE-aware drivers, network interface cards (NICs), and switches, allowing 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 Gigabit 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

Fibre Channel

- Fibre Channel: industry standard used for storage networking
- Mapped to various protocols:
  - Fibre Channel Protocol (FCP)
  - SCSI
  - FICON
- Fibre Channel ports are intelligent interface points:
  - Embedded in an I/O adapter
  - Embedded in an array or tape controller
  - Embedded in a fabric switch



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

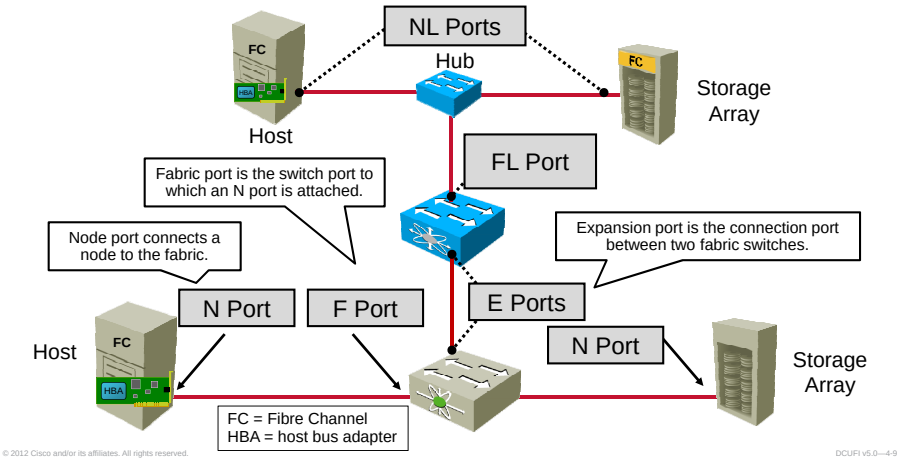
Fibre Channel is a standards-based technology that is primarily used for storage networking. It can be mapped to a number of protocols, such as FCP, SCSI, IBM Fiber Connection (FICON), and others.

The Fibre Channel infrastructure ensures communications between various interface points, called Fibre Channel ports. They may be embedded in an I/O adapter, an array or tape controller, or in a fabric switch.

https://t.me/learningnets

Fibre Channel Ports

- Common ports: node (N), fabric (F), expansion (E)
- Other ports: node loop (NL) and fabric loop (FL)
  - Node and fabric ports in one-way ring topology called arbitrated loop



The most common Fibre Channel ports are node port (N Port), fabric port (F Port) and expansion port (E Port). The node port is embedded in an I/O adapter, an array or tape controller, and connects the node to the fabric. The F Port is embedded in a fabric switch and connects to an N port. The E Port is also embedded in a switch but connects to another switch.

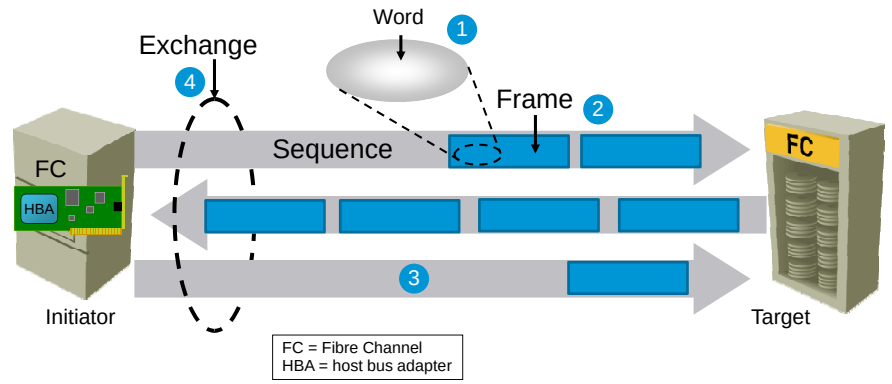
Other port types include the node loop port (NL Port) and fabric loop ports (FL Ports) that are found in the Fibre Channel Arbitrated Loop (FC-AL). FC-AL enables the devices to be connected in a one-way loop fashion in a ring topology.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Fibre Channel Framing

1. Word: Smallest data unit of data, 32 bits encoded into a 40-bit form
2. Frame: Package of multiple words, equivalent to an IP packet
3. Sequence: Unidirectional series of frames
4. Exchange: Series of sequences sent between two nodes



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

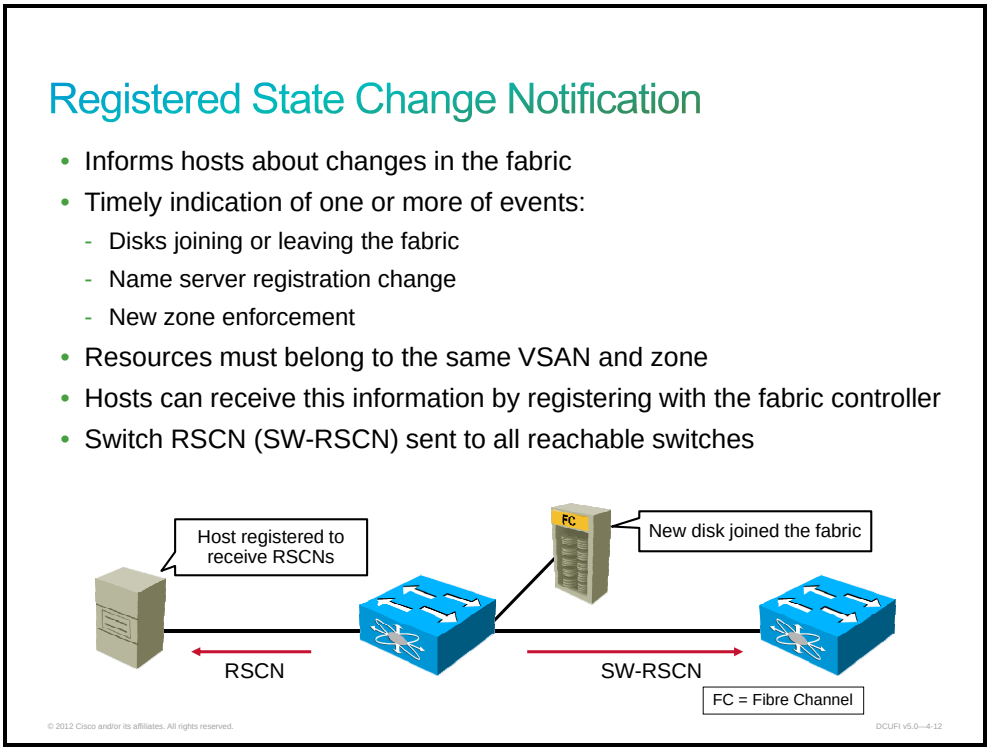
This figure shows a transaction between a host (initiator) and a storage device (target). It has the following characteristics:

- The smallest unit of data is a word. Words consist of 32 bits (4 bytes) of data that are encoded into a 40-bit form by the 8-bit or 10-bit encoding process.
- Words are packaged into frames. A Fibre Channel frame is equivalent to an IP packet.
- A sequence is a series of frames that are sent from one node to another node. Sequences are unidirectional. In other words, a sequence is a set of frames that are issued by one node.
- An exchange is a series of sequences that are sent between two nodes. The exchange is the mechanism that is used by two ports to identify and manage a discrete transaction. The exchange defines an entire transaction, such as a SCSI read or write request. An exchange is opened whenever a transaction is started between two ports and is closed when the transaction ends. A Fibre Channel exchange is equivalent to a TCP session.

https://t.me/learningnets

Fibre Channel Login and Communication

This topic shows how a device performs a fabric login (FLOGI) on a Fibre Channel SAN.



The Registered State Change Notification (RSCN) is a Fibre Channel service that informs hosts about changes in the fabric. The RSCNs are exchanged about fabric events that occur within the virtual storage area network (VSAN) and zone. Hosts can receive this information by registering with the fabric controller (through Service Control Response [SCR]). These notifications provide a timely indication of one or more of the following events:

- Disks joining or leaving the fabric
- A name server registration change
- A new zone enforcement
- IP address change
- Any other similar event that affects the operation of the host

In parallel, a switch RSCN (SW-RSCN) is sent to registered hosts and to all reachable switches in the fabric.

Note	The switch sends an RSCN to notify registered nodes that a change has occurred. It is up to the nodes to query the name server again to obtain the new information. The details of the changed information are not delivered by the switch in the RSCN that is sent to the nodes.

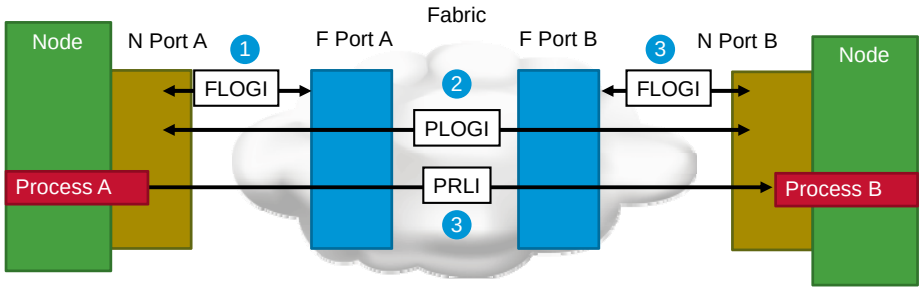
© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Fabric Login (FLOGI)

Three processes prerequisite to node port communication:

1. N Port must log into its attached F Port—fabric login (FLOGI)
2. N Port must log in to its target N Port—port login (PLOGI)
3. N Port must exchange ULP support information with its target N Port:
   - Ensure that the initiator and target process can communicate.
   - Known as the process login (PRLI).



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

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

1. The N Port must log into its attached F Port. This process is known as the fabric login (FLOGI).

1. The N Port must log into its target N Port. This process is known as the port login (PLOGI).

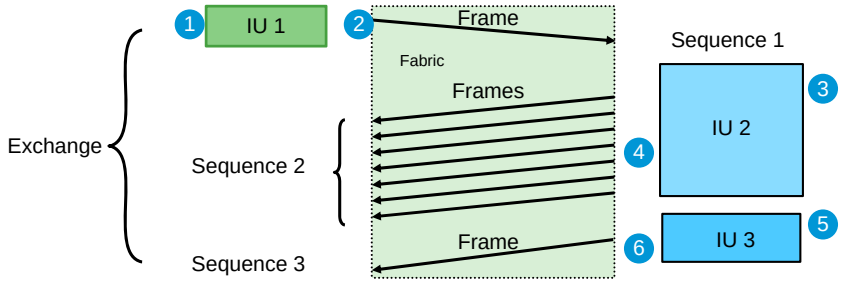
2. The N Port must exchange information about upper-layer protocol (ULP) support with its target N Port to ensure that the initiator and target process can communicate. This process is known as the process login (PRLI).

Note	Upper-layer processes such as SCSI TARGET discovery and SCSI REPORT LUNs also need to take place before data frames can be exchanged over an FCP session. They are not covered in this lesson.

https://t.me/learningnets

Fibre Channel Read Operation

1. Information Unit (IU) 1 carries SCSI read request (FCP_CMD).
2. Initiator FC-2 layer sends IU 1 as a single frame.
3. Target retrieves requested data (FCP_DATA) and packages as IU 2.
4. Target FC-2 layer sends IU 2 as frame sequence.
5. Target generates status command (FCP_RSP) and packages as IU 3.
6. Target FC-2 layer sends IU 3.



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

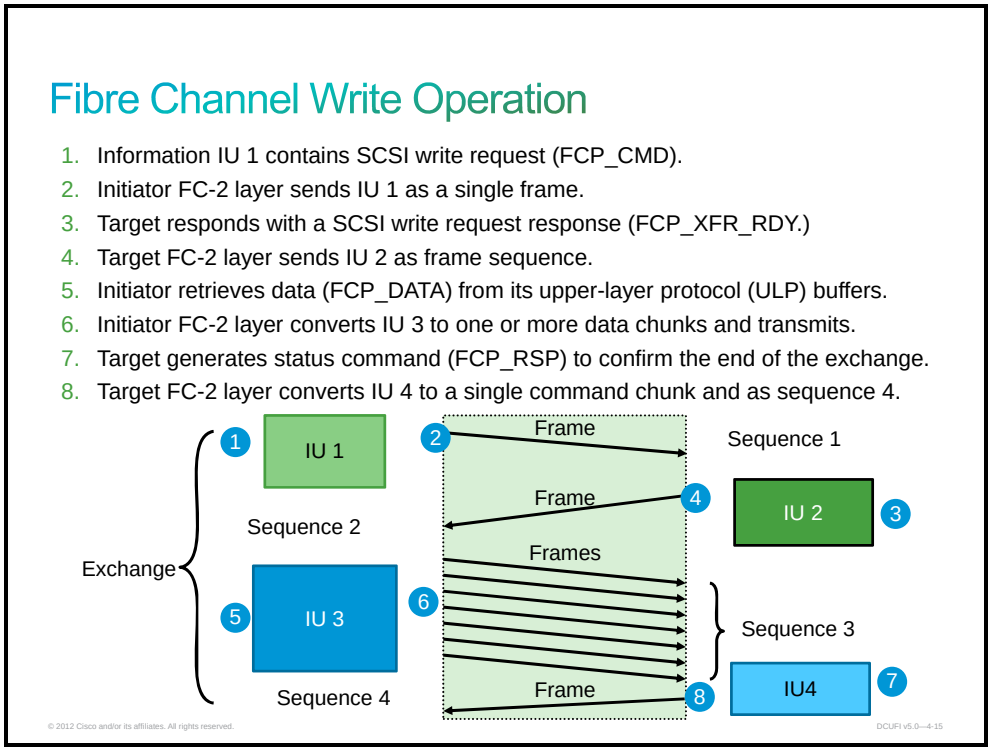
The figure shows a SCSI- FCP read operation, which consists of the following actions:

1. The initiator node generates a SCSI read request (FCP_CMD), which is packaged as information unit (IU) 1.
2. The initiator FC-2 layer converts IU 1 to a single command chunk and sends it across the fabric as a single frame. This process constitutes sequence 1.
3. The target node processes IU 1, retrieves the requested data (FCP_DATA) from storage, and packages the data as IU 2.
4. The target FC-2 layer converts IU 2 to one or more data chunks and sends them across the fabric. This process constitutes sequence 2.
5. The target node then generates a status command (FCP_RSP) that informs the initiator that the requested data transmission is complete. The status command is packaged as IU 3.
6. The target FC-2 layer converts IU 3 to a single command chunk and sends it across the fabric. This process constitutes sequence 3.

At this point, the I/O operation is complete. The collection of three sequences constitutes a single exchange.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets



The figure shows a SCSI-FCP write operation, which consists of the following actions:

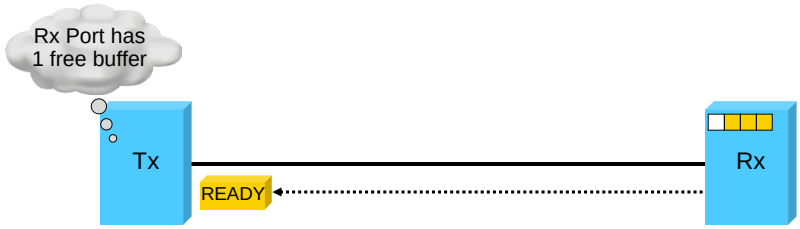
1. The initiator node generates a SCSI write request (FCP_CMD), which is packaged by the FC-4 layer as IU 1.
2. The initiator FC-2 layer converts IU 1 to a single command chunk and sends it across the fabric as a single frame. This process constitutes sequence 1.
3. The target node responds with a SCSI write request response (FCP_XFR_RDY), which is packaged as IU 2. The write request response is sent to confirm that the buffers have been released and are ready to accept data.
4. The target FC-2 layer converts IU 2 to a single command chunk and sends it across the fabric. This process constitutes sequence 2.
5. The initiator node retrieves the data (FCP_DATA) from its ULP buffers and packages it as IU 3.
6. The initiator FC-2 layer converts IU 3 to one or more data chunks and sends them across the fabric. This process constitutes sequence 3.
7. The target then generates a status command (FCP_RSP) to confirm the end of the exchange. The command is packaged as IU 4.
8. The target FC-2 layer converts IU 4 to a single command chunk and sends it across the fabric. This process constitutes sequence 4.

The collection of four sequences constitutes a single exchange.

https://t.me/learningnets

Fibre Channel Flow Control

- Fibre Channel uses a credit-based strategy.
  - 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.
- Benefits
  - 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. This means that the sender keeps track of the number of free buffers on the receiver side.

A credit-based strategy ensures that the 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 receive (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:

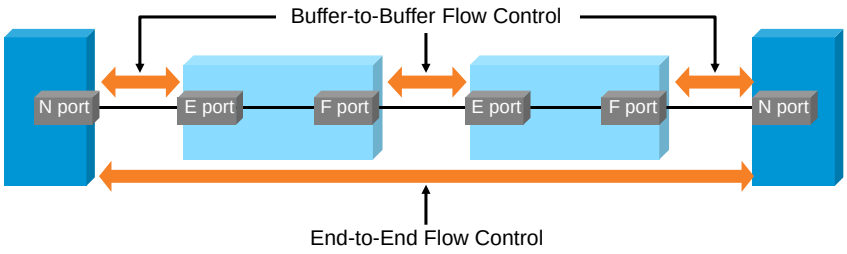
- The Tx port counts the number of free buffers at the Rx port.
- Before the Tx port can send a frame, the Rx port must notify the Tx port that the Rx port has a free buffer and is ready to accept a frame. When the Tx port receives the notification, it increments its count of the number of free buffers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

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Fibre Channel Flow Control Types

- 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:

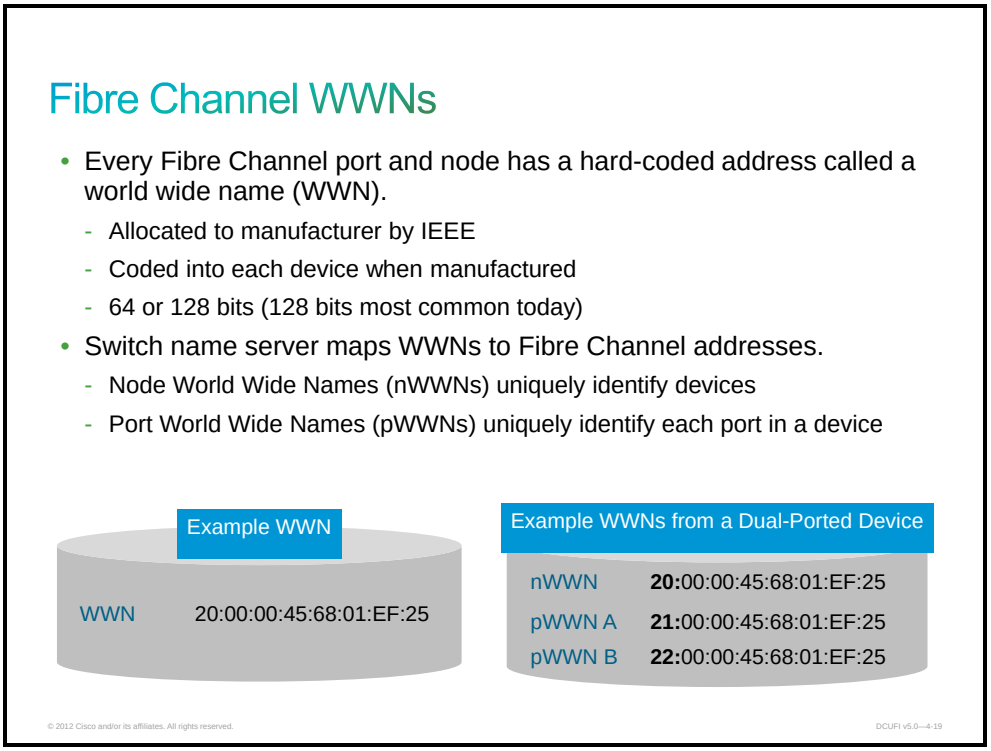
- Link level flow control takes place between two ports that are connected by a Fibre Channel link, such as an N Port and an F Port, two E Ports, or two loop ports (L Ports).
- End-to-end flow control takes place between the source node and the destination node.

Note that buffer-to-buffer 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.

https://t.me/learningnets

Fibre Channel Addressing Schemes

This topic explains the addressing scheme that is used with the Fibre Channel Protocol (FCP).



Fibre Channel uses world wide names (WWNs) and Fibre Channel IDs (FCIDs). WWNs are unique identifiers that are hardcoded into Fibre Channel devices. FCIDs are dynamically acquired addresses that, in a switch fabric, are routable.

Each 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 the following characteristics:

- Guaranteed to be globally unique
- Permanently associated with devices

These characteristics ensure that the fabric can reliably identify and locate devices. When a management service or application needs to quickly locate a device, the following occurs:

1. The service or application queries the switch name server service for the WWN of the target device.
2. The name server looks up and returns the current port address that is associated with the target WWN.
3. The service or application communicates with the target device using the port address.

There are two types of WWNs—node and port:

- Node world wide names (nWWNs) uniquely identify devices. Every host bus adapter (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.

© 2012 Cisco Systems, Inc.

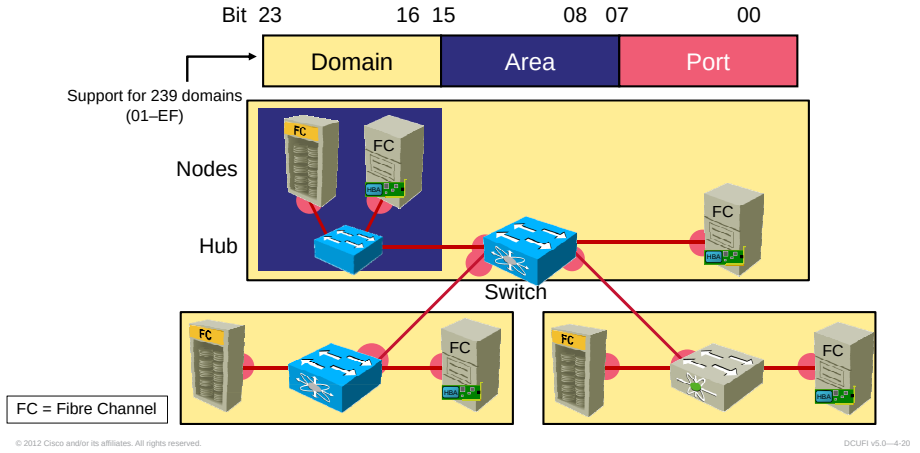
https://t.me/learningnets

nWWNs and pWWNs are both needed because devices can have multiple ports. On a single-port device, the nWWN and pWWN may be the same. On multiport devices, 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

Fibre Channel Address Format

- Domain ID defines a switch. Each switch receives a unique domain ID.
- Area ID identifies groups of ports within a domain.
- Port ID identifies individual device(s) on that port.



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

The 24-bit Fibre Channel address consists of three 8-bit elements:

- The domain ID is used to identify a switch in the Fibre Channel SAN. Each switch receives a unique domain ID.
- The area ID may be used to identify groups of ports within a domain. Areas can be used to group ports within a switch. They may also be used to uniquely identify fabric-attached arbitrated loops. Each fabric-attached loop receives a unique area ID.
- The port ID is used to identify the individual device(s) on that port.

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

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

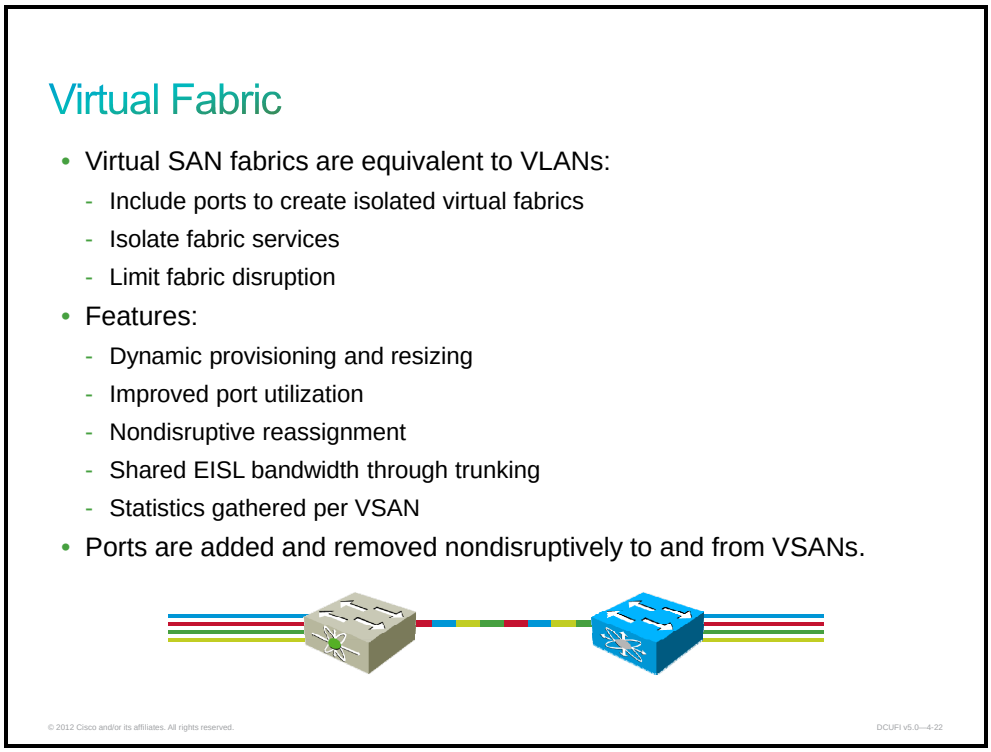
Each switch must have a unique domain ID so that there can be no more than 239 switches in a fabric. In practice, the number of switches in a fabric is far smaller because of the storage vendor qualifications.

Note	Cisco supports 80 domain IDs in a fabric. Storage manufacturers can further restrict this number.

https://t.me/learningnets

VSANs and IVR

This topic shows how VSANs and Inter-VSAN Routing (IVR) are used within a Fibre Channel SAN.



Virtual fabric functionality leverages the advantages of isolated SAN fabrics with capabilities that address the limitations of isolated SAN islands. VSANs provide a method for allocating ports within a physical fabric to create virtual fabrics. Independent physical SAN islands are virtualized onto a common SAN infrastructure. VSANs let you virtualize the physical switch into many virtual switches. Separate fabric services are available on each VSAN, as are statistics that are gathered on a per-VSAN basis.

You can use VSANs to deploy a collapsed fabric that provides a common redundant physical infrastructure for clean partitioning of different operating environments, fewer fabric switches to manage, and a network that is significantly more stable and manageable than a zoning-plus-best-practices approach.

Each VSAN provides its own Fibre Channel fabric services, which include these:

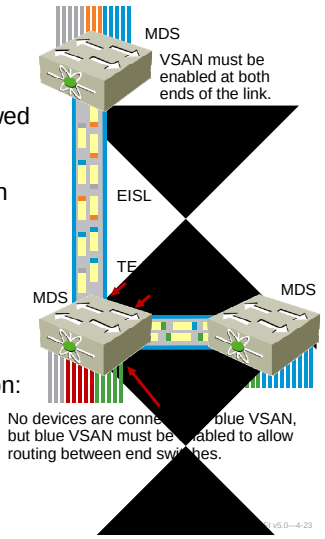
- FLOGI server
- Distributed name server
- Distributed zoning
- Fabric Shortest Path First (FSPF) routing protocol
- Management server

If a fabric service should crash or fail in one VSAN, it will not affect other VSANs. Therefore, applications within each VSAN are not affected by failures of other applications in other VSANs.

https://t.me/learningnets

TE Ports and EISL

- Equivalent to trunk ports
- TE Port
  - Carries tagged frames from multiple VSANs
  - Trunks all VSANs (1-4093) by default
  - VSAN-allowed list defines which frames are allowed
  - Can be optionally disabled for E Port operation
  - Has native VSAN assignment for E Port operation
  - Not to be confused with port channels
- EISL
  - Link created by connecting two TE Ports
  - Enhanced ISL functionality
  - Also carries per-VSAN control protocol information:
  - FSPF, distributed name server, zoning updates



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

Multiple VSANs can be transported over a single Fibre Channel link in the same way that multiple VLANs are carried over a trunk port. The required interfaces include trunking expansion ports (TE Ports) and Enhanced Inter-Switch Link (EISL) ports and are supported only on Cisco MDS 9000 Series Multilayer Switches.

TE Ports have the following characteristics:

- Can pass tagged frames belonging to multiple VSANs.
- By default, pass all VSAN traffic (1–4093). The passing of traffic for specific VSANs can be disabled.
- By default, are assigned as part of VSAN 1.
- Allow for the segregation of SAN traffic and should not be confused with port aggregation (trunking).

EISLs are interswitch links (ISLs) that connect two TE Ports:

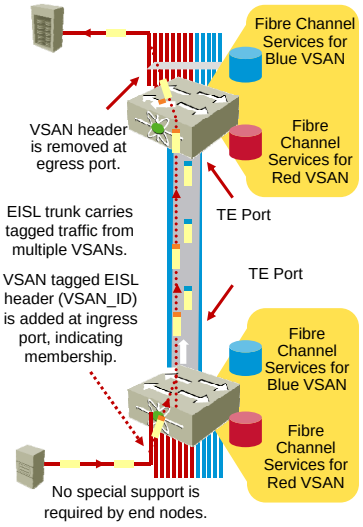
- An EISL is created when two TE Ports are connected.
- EISLs offer a superset of ISL functionality.
- EISLs carry per-VSAN control protocol information.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

VSAN Primary Functions

- Hardware-based isolation of traffic:
  - No special drivers or configuration are required for end nodes.
  - Frames are tagged at Fibre Channel ingress port and carried across EISL or PortChannel with trunking enabled.
- Independent fabric services for each VSAN:
  - Login server
  - Name server
  - Zone server
  - Management server
  - FSPF routing
  - Services run, managed, and configured independently



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

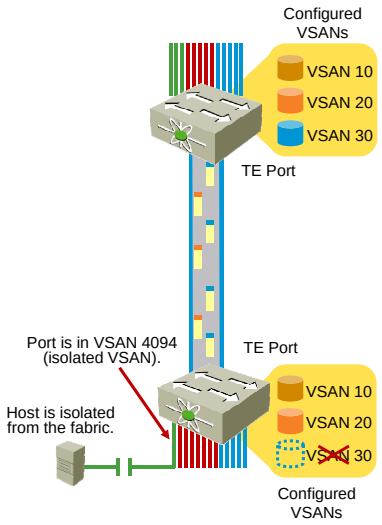
The VSAN feature provides two primary functions:

- Hardware-based isolation of tagged traffic belonging to different VSANs, which requires no special drivers or configuration at the end nodes, such as hosts, disks, and so on. Traffic is tagged at the Fibre Channel ingress port (Fx Port) and carried across EISL links between Cisco MDS 9000 Series switches. Because VSANs use explicit frame tagging, they can be extended over the metropolitan-area network (MAN) or WAN. The Cisco MDS 9000 IP Storage Services Module can add tags that are to be transported in Fibre Channel over IP (FCIP) for greater distances. Fibre Channel, and therefore VSANs, can easily be carried across dark fiber. However, VSANs add 8 bytes of header, which can be a concern for channel extenders. The channel extenders might consider it an invalid frame and drop it. Dense wavelength-division multiplexing (DWDM) switches might also count frames as invalid but might still pass the frames anyway. Qualification is still ongoing within Cisco to validate various extension methods.
- Creation of an independent instance of Fibre Channel fabric services for each newly created VSAN. These services include zone server, name server, management server, and principle switch selection. Each service runs independently on each VSAN and is independently managed and configured as well.

https://t.me/learningnets

VSAN Numbering

- VSAN 1:
  - Automatically configured by the switch as the default VSAN
  - All ports are originally in VSAN 1
  - Always present, cannot be deleted
- VSAN 2 through 4093:
  - User-configurable VSANs
  - 4000 VSANs per physical fabric
- VSAN 4094 (Isolated VSAN):
  - Used to isolate ports whose port-VSAN has been deleted
  - Not propagated across switches
  - Always present, cannot be deleted



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

There are certain rules that you must follow when creating VSANs. VSAN 1, for instance, is automatically configured by the switch as the default VSAN. All ports that are configured are originally put into VSAN 1 until they are configured into another VSAN number.

The VSAN numbers ranging from 2 through 4093 are the user-configurable VSANs.

VSAN 4094 is a reserved special VSAN that is called the isolated VSAN. It is used to temporarily isolate the ports whose VSAN has been deleted. VSAN 4094 is not propagated across switches; it is always present, and it cannot be deleted.

In the figure, VSAN 30 is not propagated across EISL, because it is not configured in the local switch, but it is configured on the remote switch. Instead of the host device on the local switch being able to connect to the remote switch, it has been placed in the isolated VSAN 4094 because the VSAN of the port VSAN 30, has been deleted from the local switch configuration.

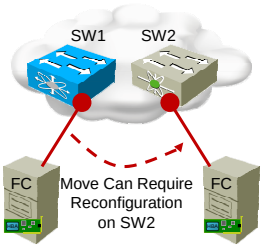
Note VSAN 0 and VSAN 4095 are reserved and cannot be used.

https://t.me/learningnets

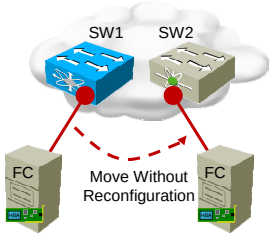
Port- and WWN-Based VSANs

Port-based VSANs

- VSAN membership based on physical switch port
- Switch-wide configuration
- Reconfiguration required when server or storage moves to another switch
- Switch port belongs to VSAN



- WWN-based VSANs
- VSAN membership based on pWWN of server or storage
- Fabric-wide distribution of configuration using Cisco Fabric Services
- No reconfiguration required when a host or storage moves
- Device port belongs to VSAN



FC = Fibre Channel

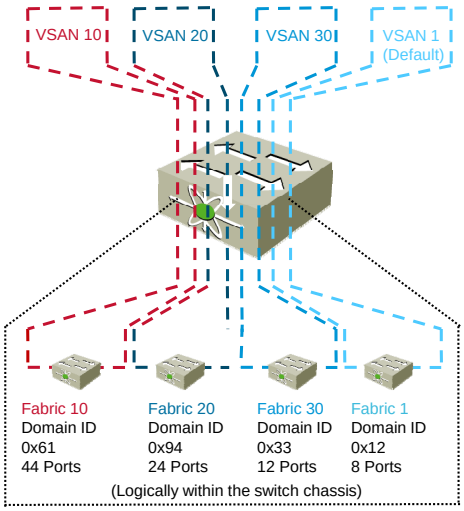
© 2012 Cisco and/or its affiliates. All rights reserved.

VSAN membership can now be defined based on the WWN of hosts and storage devices, or by switch port. With WWN-based VSAN membership, host and targets can be moved from one port to any other port anywhere in the Cisco MDS fabric without requiring manual reconfiguration of the port VSANs.

https://t.me/learningnets

VSAN Tagging

- Traffic isolation:
  - Control over each incoming and outgoing port
- Each frame in the fabric is uniquely identified:
  - Labeled with a VSAN_ID header on the ingress port
  - VSAN_ID stripped away across E Ports
  - Maintained across TE Ports
- SAN and priority in header for QoS
- FC-ID can be reused across multiple VSANs



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

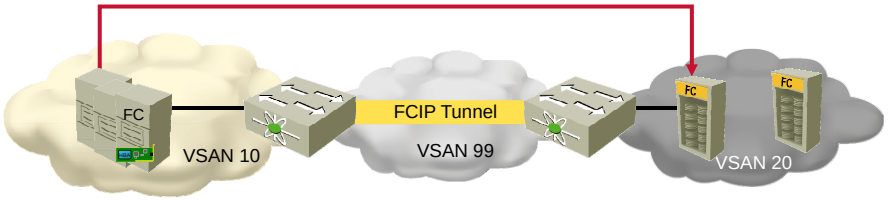
To uniquely identify each frame in the fabric, the frame is labeled with VSAN identification (VSAN_ID) on the ingress port; the VSAN_ID is stripped away across E Ports. Across TE Ports, the VSAN ID is still maintained. By carrying SAN and priority in the header, quality of service (QoS) can be properly applied. The VSAN_ID is always stripped away at the other edge of the fabric. If an expansion port (E Port) is capable of carrying multiple VSANs, it becomes a TE Port.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Inter VSAN Routing

- IVR allows selective routing between specific members from two or more VSANs.
- IVR process is stateful.
- Most Fibre Channel control traffic is blocked and cannot pass the VSAN boundary.
- Route (Domain ID) redistribution from one VSAN to the other enables IVR.



FC = Fibre Channel

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

Data traffic isolation between the VSANs inherently prevents sharing of resources that are attached to a VSAN, for example, robotic tape libraries. Using IVR, resources can be accessed from other VSANs without compromising the VSAN benefits.

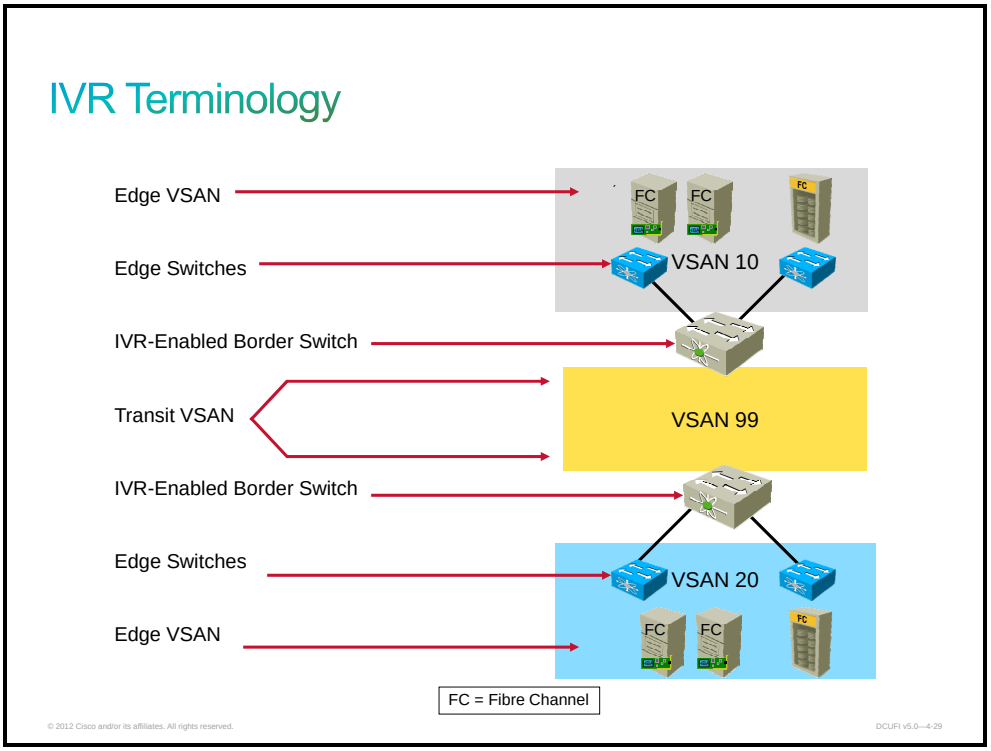
When implementing IVR, data traffic is transported between specific initiators and targets from different VSANs without merging them into a single logical fabric. Fibre Channel control traffic does not flow between VSANs, and initiators can access only designated resources across VSANs. With IVR, valuable resources like tape libraries can be shared across VSANs. IVR provides more efficient business continuity or disaster recovery solutions.

The IVR process is stateful. If the IVR process fails, it can be restarted without losing IVR process state information. Virtually all Fibre Channel control traffic is blocked and cannot pass the VSAN boundary. A small subset of control frames is still needed to allow communication between devices that are involved in IVR.

In the figure, IVR is used to allow a server in VSAN 10 to access selected logical unit numbers (LUNs) in a storage array in VSAN 20. The server-LUN traffic is transported across VSAN 99 over an FCIP link.

Note For example, in this graphic all traffic in VSAN 10 is isolated from VSAN 20, so that change notifications (adding or removing nodes, for example) are not sent across the FCIP Tunnel. Communication is still possible, however, by creating VSAN 99 and using IVR to route Fibre Channel traffic between VSAN 10 and VSAN 20 while reducing control-plane chatter.

https://t.me/learningnets



The following terms apply to IVR components and IVR configuration, as shown in the figure:

- **IVR zone:** A set of end devices that are allowed to communicate across VSANs within their interconnected SAN fabric. This definition is based on their port world wide names (pWWNs) and their native VSAN associations.
- **IVR zone set:** One or more IVR zones, which make up an IVR zone set. You can configure up to 32 IVR zone sets on any switch in the Cisco MDS 9000 Series. Only one IVR zone set can be active at any time.
- **IVR path:** An IVR path is a set of switches and interswitch links (ISL) through which a frame from an end device in one VSAN can reach another end device in a different VSAN. Multiple paths can exist between two such end devices.
- **IVR-enabled switch:** A switch in which the IVR feature is enabled.
- **Edge VSAN:** A VSAN that initiates (source edge VSAN) or terminates (destination edge VSAN) an IVR path. Edge VSANs can be next to each other, or they can be connected by one or more transit VSANs. In the figure, VSANs 10 and 20 are edge VSANs.

Note	An edge VSAN for one IVR path can be a transit VSAN for another IVR path.

- **Transit VSAN:** A VSAN that exists along an IVR path from the source edge VSAN of that path to the destination edge VSAN of that path is called a transit VSAN. In the figure, VSAN 99 is a transit VSAN.

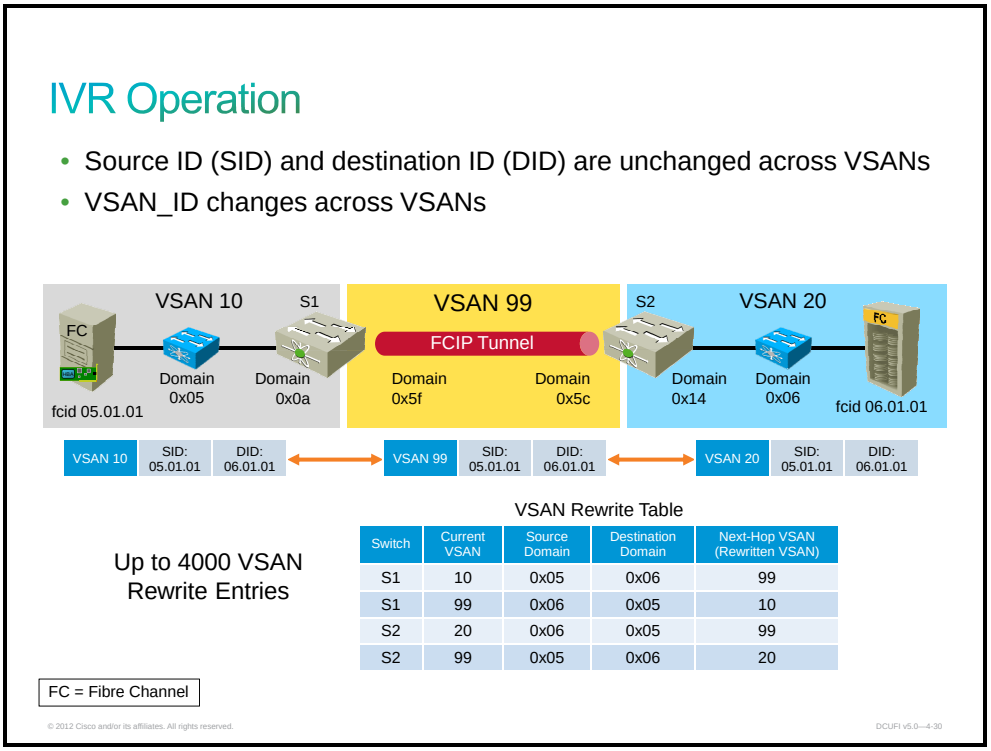
Note	When the source and destination edge VSANs are next to each other, a transit VSAN is not required between them.

© 2012 Cisco Systems, Inc.
https://t.me/learningnets

- **Border switch:** An IVR-enabled switch that is a member of two or more VSANs, such as switches S1 and S2 in the figure.
- **Edge switch:** A switch to which a member of an IVR zone has logged in. Edge switches are oblivious to the IVR configurations in the border switches. Edge switches do not need to be IVR-enabled. The edge switch can be a Cisco MDS 9000 Series or Nexus 5000 Series switch.

Note A single switch can be both a border switch and an edge switch.

https://t.me/learningnets



Whenever a frame enters a Fibre Channel switch, it is tagged with a VSAN header that indicates the native VSAN of the port. In the case of IVR, when the destination FCID resides in a different VSAN, the tag can be rewritten at the ingress port of the IVR border switch.

In the figure, assume that a frame is destined from FCID 5.1.1 in VSAN 10 to FCID 6.1.1 in VSAN 20. The left-most switch, with domain ID 5, applies a VSAN 10 ID tag. The next switch performs a VSAN rewrite, changing the VSAN tag to 99. The last switch changes the VSAN tag to 20. The process is reversed on the return path.

VSAN Rewrite Table

Each IVR-enabled switch maintains a copy of the *VSAN rewrite table*. The table can hold up to 4096 entries. Each entry includes the following information:

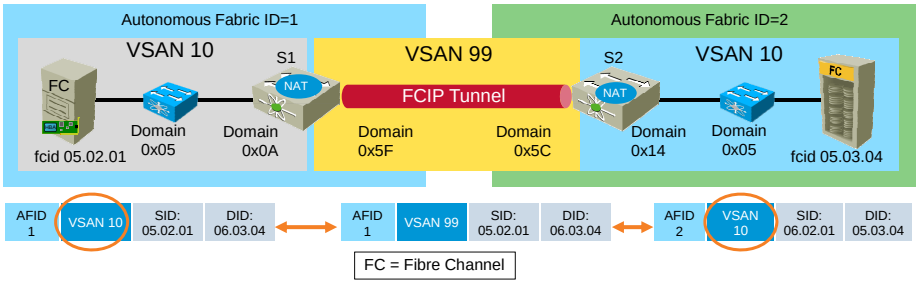
- Switch identifier
- Current VSAN_ID
- Source domain
- Destination domain
- Next-hop VSAN (rewritten VSAN)

In the figure, if a frame with source FCID 5.1.1 arrived at switch S1 with a destination FCID of 6.1.1, the VSAN tag is rewritten from 10 to 99. Notice that the table entries are per switch domain, not per end device.

© 2012 Cisco Systems, Inc.
https://t.me/learningnets

Routing Between Overlapping VSAN IDs

- IVR autonomous fabric ID (AFID):
  - Identifies IVR topology
  - Contains VSAN with unique IDs
  - Manually configured
  - Up to 64 separate IDs supported
- The same VSAN_ID can exist in different AFIDs.
- IVR NAT rewrites the SID or DID and routes frames.



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

The autonomous fabric ID (AFID) is used to distinguish segmented VSANs. Segmented VSANs are two VSANs that are logically and physically separate but have the same VSAN number. The AFID can be configured individually for each switch and list of VSANs, or the default AFID can be configured for each switch. Each VSAN in a single physical switch must belong to only one autonomous fabric. IVR must know about the topology of the IVR-enabled switches in the fabric to function properly.

You can specify the topology in two ways:

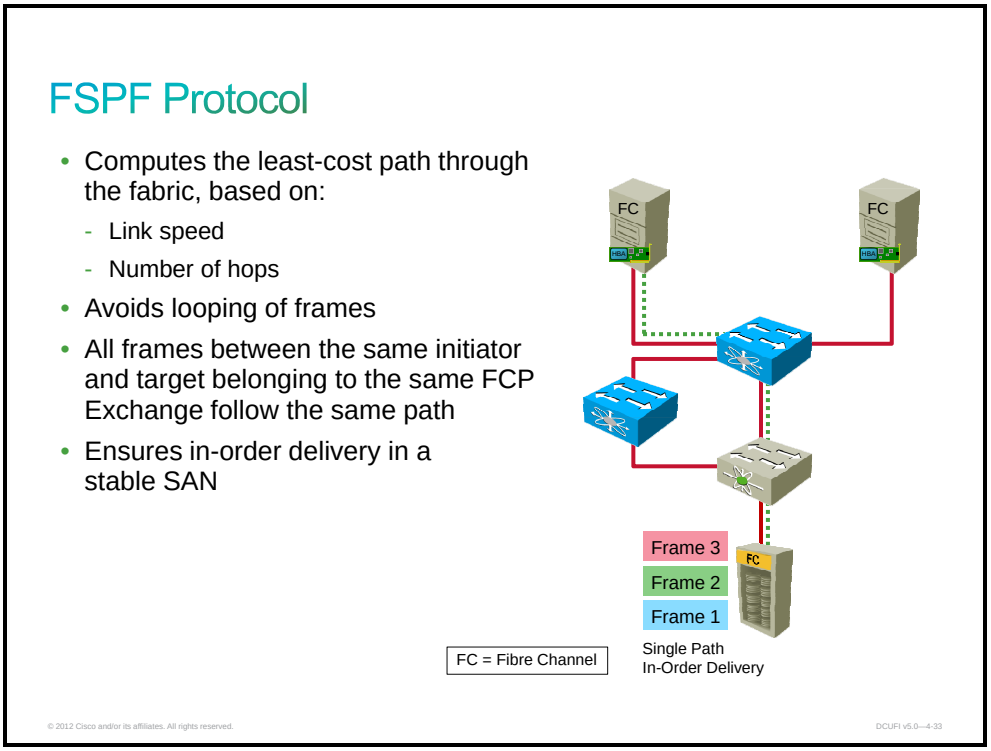
- **Manual configuration:** Configure the IVR topology manually on each IVR-enabled switch.
- **Automatic mode:** Uses Cisco Fabric Services configuration distribution to dynamically learn and maintain up-to-date information about the IVR topology.

In the figure, VSAN 10 on the left is joined to VSAN 10 on the right via a transit VSAN 99. This configuration would be illegal in a single autonomous fabric, so both sides are configured in separate autonomous fabrics 1 and 2. Notice that IVR Network Address Translation (NAT) now rewrites the AFID in the EISL frame header from AFID1:VSAN 10 to AFID2:VSAN 10 as it passes through the IVR edge switches.

https://t.me/learningnets

FSPF Protocol

This topic shows how the FSPF protocol is used on the SAN.



FSPF is the routing protocol that is used in Fibre Channel SAN fabrics. The figure shows that FSPF selects a single path for a given I/O transaction, avoiding looping and ensuring in-order delivery.

The FSPF algorithm is a cost-based routing algorithm that computes the most efficient path between two connected nodes. The cost of a given path is based on two factors:

- The speed of each of the ISLs along the path
- The number of hops on the path

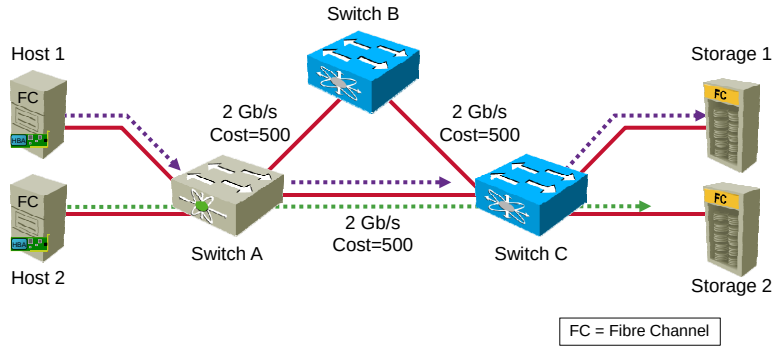
Routing using a single fixed path prevents looping of frames. In other words, all frames between the same initiator and target belonging to the same FCP Exchange follow the same path. However, if the least-cost route changes while a session is in progress, frames sent after the route change might take the new route.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

FSPF Path Selection

- FSPF algorithm does not account for traffic load.
- All frames in an exchange follow the same path.
- Path changes only in response to changes in the fabric topology.



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

The FSPF protocol supports load sharing, but it does not support load balancing. Load sharing is significantly different from load balancing. The distinction can have significant effects for fabric design, especially when you are tuning performance:

- Load sharing simply means that multiple paths can be used.
- Load balancing means that traffic load is balanced across multiple paths.

FSPF does not account for actual path utilization. In other words, an unused path with a cost of 1000 is disregarded in favor of an overutilized path with a cost of 500. All frames in an exchange must follow the same path, and paths are recomputed only when the physical ISL configuration changes.

The figure shows a simple SAN with two data paths:

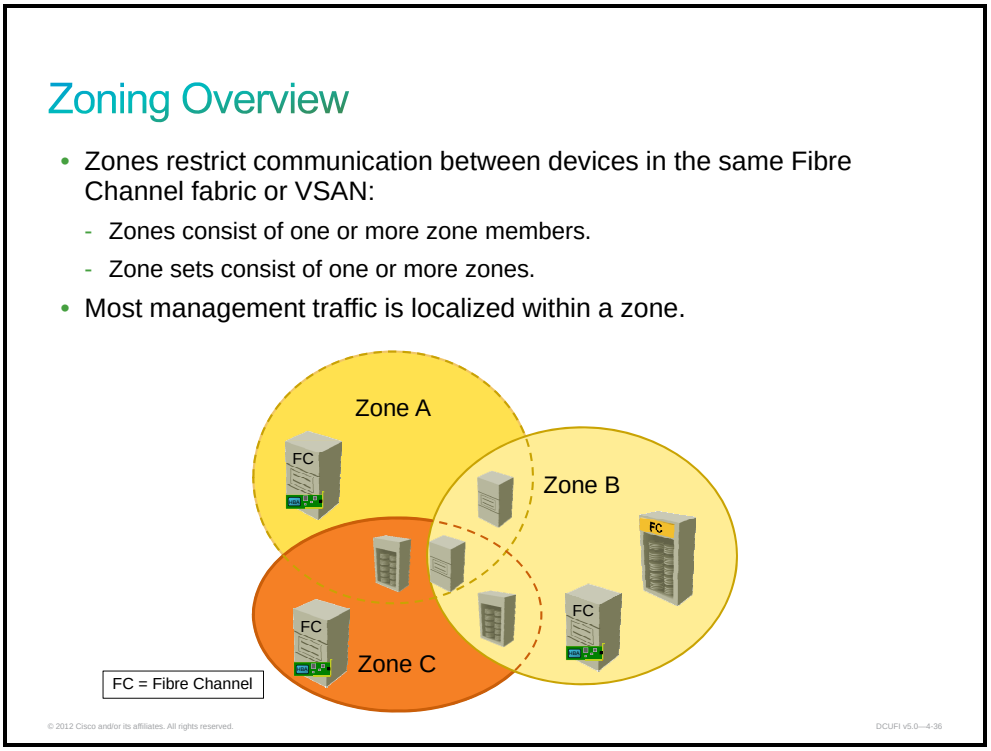
- Path A→C has a total cost of 500.
- Path A →B→C has a total cost of 1000.

FSPF never uses path A→B→C, even if path A→C is congested.

https://t.me/learningnets

Zoning

This topic shows how zoning is used within a Fibre Channel SAN.



Zoning is a security mechanism within Fibre Channel that is used to restrict communication between devices with the same Fibre Channel fabric. Zoning carves a Fibre Channel fabric into multiple partitions. Devices in one zone cannot learn of the existence of devices in other zones.

With many different types of servers and storage devices on the network, the need for security is critical. For example, if a host were to gain access to a disk that is being used by another host, potentially with a different operating system, the data on this disk could become corrupted. To avoid any compromise of critical data within the SAN, zoning allows the user to overlay a security map that dictates which devices, namely hosts, can see which targets, thus reducing the risk of data loss.

As shown in the figure, a zone set consists of one or more zones:

- A zone set can be activated or deactivated as a single entity across all switches in the fabric.
- Only one zone set can be activated at any time.
- A zone can be a member of more than one zone set.
- A zone consists of multiple zone members. Members in a zone can access each other; members in different zones cannot access each other.

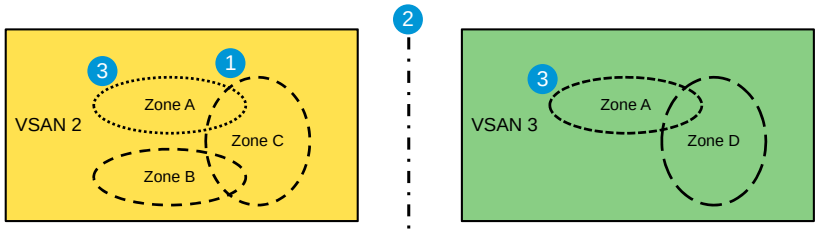
Typical uses for zoning include the following:

- Separating initiators from their targets. Frequently, each initiator port will belong in a separate zone with its targets.
- Separating devices that use different operating systems. This practice is useful to protect some operating systems from treating disks that are formatted by other operating systems as blank disks and potentially taking over and overwriting their storage.
- Separating devices that have no need to communicate with other devices in the fabric or that have classified data.

© 2012 Cisco Systems, Inc.
https://t.me/learningnets

Zoning Rules

1. Zones can overlap:
   - Permitted by Fibre Channel standards
2. Zones typically do not cross the VSAN boundary:
   - Zones completely contained within a VSAN
   - Exception: IVR zones can cross the VSAN boundary
3. Zones are per-VSAN significant:
   - Zone A in VSAN 2 is different and separate from zone A in VSAN 3.



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

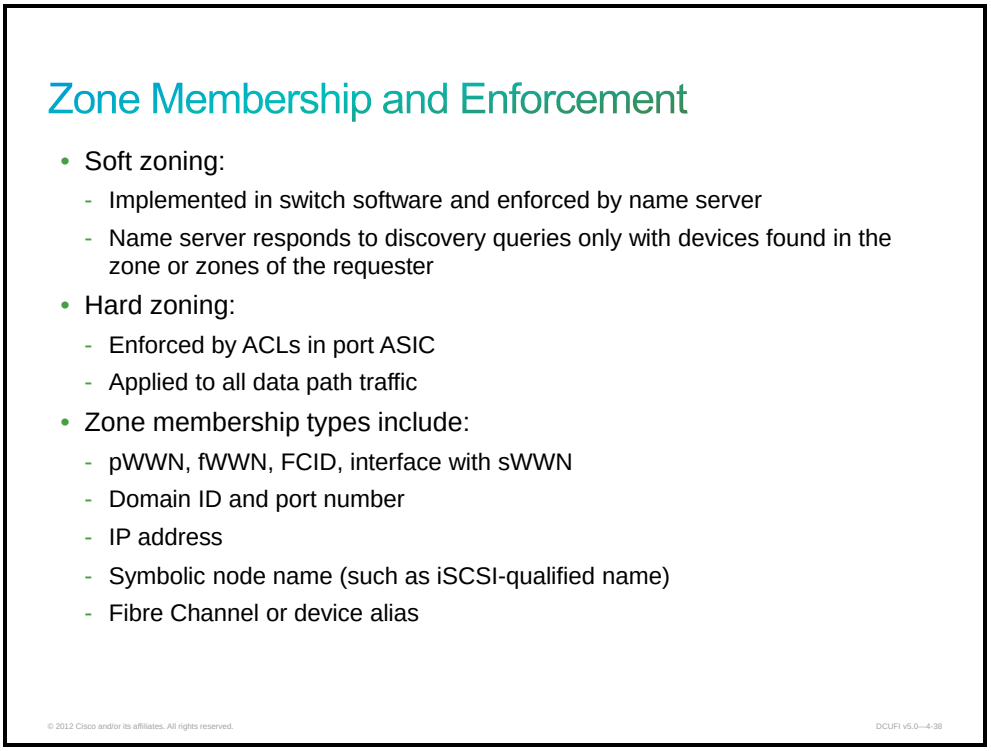
Several rules define the zone design and operation:

1. Zones can overlap. The overlapping configuration is permitted by the Fibre Channel standards.

2. Zones typically do not cross the VSAN boundary. Each zone is completely contained within a VSAN. An exception to this rule is the rule for IVR zones, which can cross the VSAN boundary.

3. Zones are per-VSAN significant. Two zones with identical names in two different VSANs are considered separate entities. In the figure, zone A in VSAN 2 is different and separate from zone A in VSAN 3.

https://t.me/learningnets



There are two basic ways to enforce zoning: software-based (soft) zoning and hardware-based (hard) zoning:

- **Soft zoning:** Enforced by the Fibre Channel Name Server (FCNS) service. When devices connect to a Fibre Channel fabric, they use the FCNS to correlate WWNs to FCIDs and to discover devices on the fabric. With soft zoning, a Fibre Channel switch responding to a name server query from a device will respond with a list of only those devices that are registered in the same zone or zones as that of the querying device.
- **Hard zoning:** Enforced through access control lists (ACLs) that are applied by the switch port ASIC to every Fibre Channel frame that is switched.

Because soft zoning does not enforce ACLs on a per-frame basis, soft zoning is not as secure as hard zoning. Soft zoning does not prevent a rogue device from attaching to the Cisco Fabric Manager Service and obtaining a list of all FCIDs.

Until recently, most Fibre Channel switches enforced interface-based zone membership in hardware, so the term “hard zoning” has often been used synonymously with “port zoning.” Similarly, WWN-based zone members were enforced only by the name server, so the term “soft zoning” has been used synonymously with “WWN zoning.” However, this is no longer the case.

A zone member is used to uniquely identify a device or devices that are to be included in a zone. Zone members can be identified by using any of nine possible methods: pWWN, fWWN, FCID, interface with switch world wide name (sWWN), domain ID, and port number, IP address, symbolic node name (such as Internet Small Computer Systems Interface [iSCSI]-qualified name), and Fibre Channel or device alias.

© 2012 Cisco Systems, Inc.
https://t.me/learningnets

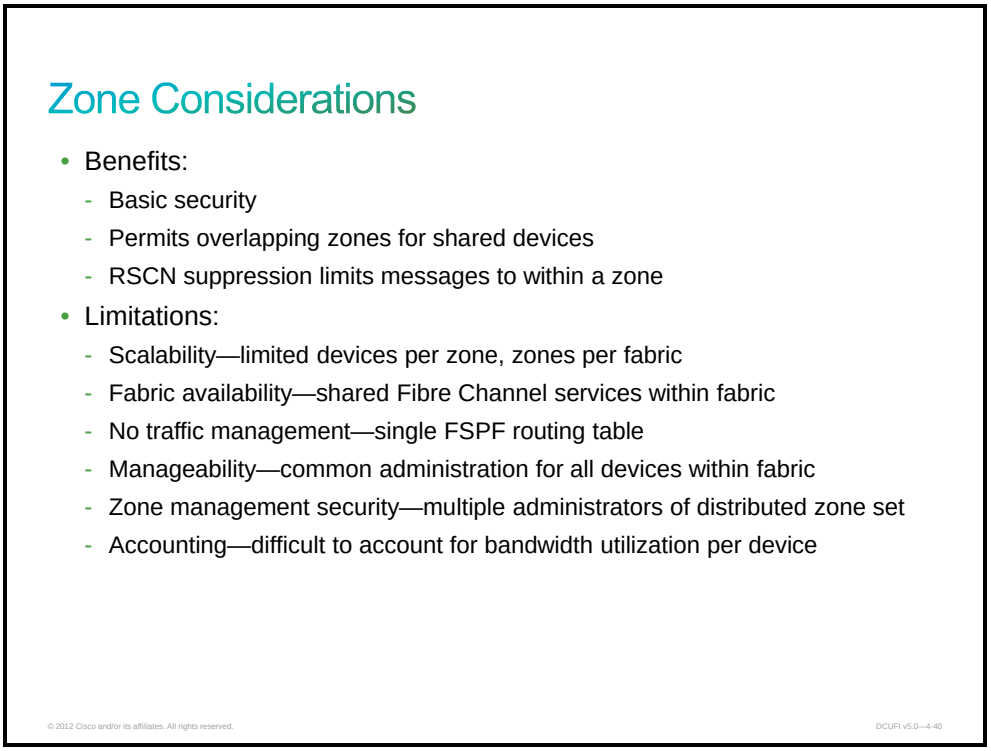
VSANs and Zoning Comparison

Characteristic	VSAN	Zone
Routing and namespace	Each VSAN uses its own separate routing and namespace	Only zones in the same VSAN are using the same routing database
Traffic flooding	Limit unicast, multicast, and broadcast traffic	Limit unicast traffic
Membership	Defined by physical port or pWWN	Defined by the pWWN
Endpoints	An HBA or a storage device can belong only to a single VSAN	An HBA or storage device can belong to multiple zones
Membership enforcement	Enforced at each E Port, source port, and destination port	Enforced only at the source and destination ports
Scope	Defined for larger environments (storage service providers)	Defined for a set of initiators and targets not visible outside the zone
Configuration change	Changes only when ports are needed per virtual fabric	Can change frequently (for example, backup)
Recommended use	Per application or per department	Single-initiator zoning

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

The table contrasts the features of VSANs and zones.

https://t.me/learningnets



Zoning was designed to do nothing more than prevent devices from communicating with other unauthorized devices. It is a distributed service that is common throughout the fabric. Any installed changes to a zoning configuration are therefore disruptive to the entire connected fabric.

Benefits of Zoning

Enhanced Device Security

By deploying zoning within a Fibre Channel fabric, device access is limited to devices within the zone. This limitation allows the user to segregate devices that are based on access to a particular storage device (target). It is generally an absolute requirement when you are dealing with multiple operating system environments that are accessing the same physical fabric.

RSCN suppression: By implementing zoning within a Fibre Channel network, RSCNs are isolated to the devices within the zone in which the state change happened. This method causes fewer disruptions to devices within the fabric, especially to devices that have no reliance or are not in the proximity of the actual state change. In some instances, however, such as a storage array interface that may reside in several zones simultaneously, zoning cannot provide complete protection.

Multiple zones per port: When you are using zoning to isolate device connectivity, a single shared device may have to exist in multiple zones at the same time. This situation typically occurs when multiple servers require access to the same disk subsystem interface. Each server usually resides in a separate zone with the disk subsystem interface. Hence, the disk subsystem interface resides in a zone with each server. This setup allows access to the common device, the storage subsystem interface, while preventing the servers themselves from communicating with each other.

© 2012 Cisco Systems, Inc.
https://t.me/learningnets

Limitations of Zoning

Scalability

Zoning does nothing to provide a higher level of scalability for the network. Even though devices may be in separate zones, such zones still reside in one common fabric that is bound by the addressing limitations and routing scalability within the fabric. In addition, as fabrics grow larger, so, too, will the number of deployed zones and the control plane bandwidth that is required to manage the increased number of zones.

Fabric Availability

While zoning, when enforced by hardware, can provide significant security benefits within a network, it does nothing to provide enhanced availability within the fabric itself. If the name server stops responding or FSPF begins to route erratically, this behavior will have a very disruptive effect on the entire storage network because common fabric services are shared across all zones within a fabric. In addition, because of the distributed nature of zoning, any changes that are made to an active zone set have fabric-wide impacts and potential disruptions as the zone set changes are installed.

Traffic Management

Zoning is very effective in managing connectivity among end devices within a storage network. However, zoning does not offer any capability to control the path selection and flow through a storage network among zoned devices. As long as two devices are allowed to communicate as dictated by the zoning configuration, their flows can traverse any path within the SAN as determined by routing protocols within the fabric. The requirement to engineer traffic flows becomes increasingly important as storage networking environments grow, to ensure that fabric bandwidth is utilized in an optimal manner.

Manageability

Because zoning is a common distributed service throughout the fabric, it is managed as a common service. Therefore, as organizations choose to collapse multiple applications on fewer larger SAN fabrics, zoning will still exist as a common distributed service. Zoning is not a service that was designed to have management that is partitioned among different management groups.

Zone Management Security

The zoning service is implemented as a distributed service within a Fibre Channel storage network. A distributed database is synchronized and maintained within all member switches of the storage network. As a distributed service, the zoning database can be updated by any member switch of the fabric. Therefore, although zoning is a security service, its very implementation is relatively insecure.

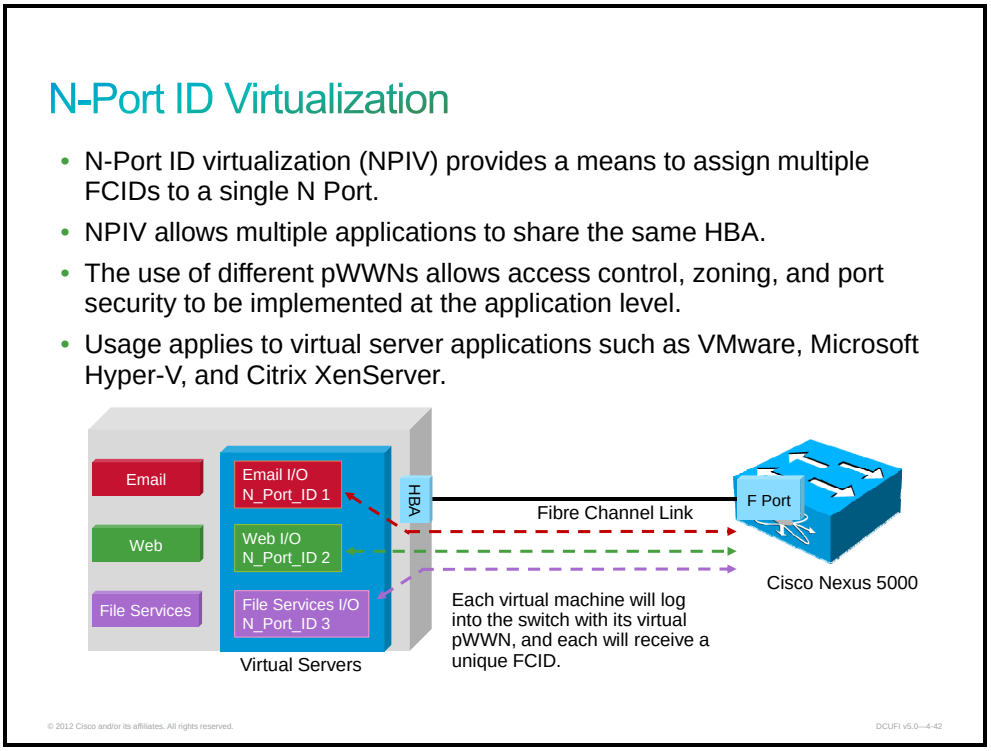
Accounting

Zoning is a relatively dynamic service that can be modified frequently for various application-oriented reasons. As such, it is exceedingly difficult to account for bandwidth utilization among application groups that may be defined by their zoning definition. Therefore, if a common fabric is built to contain multiple applications, each one divided by a zoning definition, it is generally impossible with current types of hardware to isolate and account for utilization on a per-zone basis.

https://t.me/learningnets

N-Port ID Virtualization

This topic shows when and how to use NPIV.



With the increased use of blade center deployments and top-of-rack (ToR) aggregation, the deployment and use of aggregation switches is becoming more widespread. A major concern is the total number of switches or domains that can exist in a physical fabric. As the edge switch population grows, the number of domain IDs becomes a concern. The domain is the address of a physical switch or logical virtual fabric; and is embedded in the FCID.

NPIV allows a Fibre Channel host connection (N Port), to be assigned multiple N Port IDs or FCIDs over a single link. All FCIDs assigned can be managed on a Fibre Channel fabric as unique entities on the same physical host. Different applications can be used with NPIV. In a virtual machine (VM) environment where many host operating systems or applications are running on a physical host, each VM can now be managed independently from zoning, aliasing, and security perspectives.

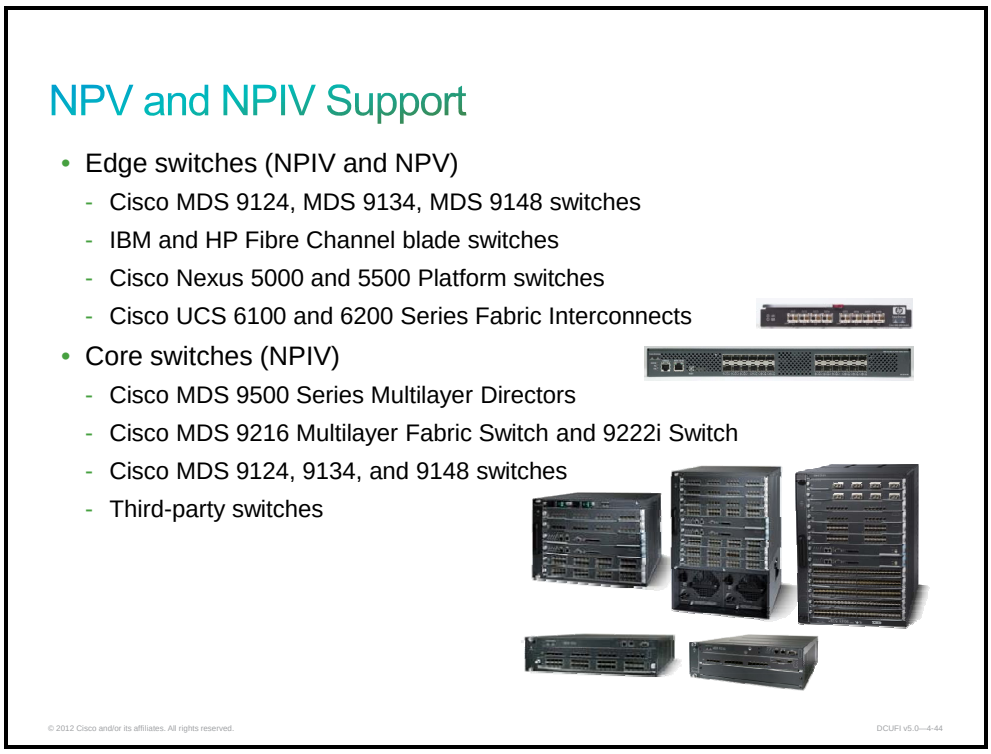
An HBA that supports the NPIV feature follows the standard login process. The initial connection and login to the fabric is performed through the standard fabric login (FLOGI) process. All subsequent logins for either VMs or logical part ions on a mainframe are transformed into fabric discover (FDISC) login commands. The FDISC logins follow the same standard process. Each subsequent FDISC login will be a part of the same VSAN as the original fabric login.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

N-Port Virtualization

This topic identifies when and how to use the NPV mode.



NPV is supported on the following NPV edge switches:

- Cisco MDS 9124, 9134, and 9148 Multilayer Fabric Switches
- IBM and HP Fibre Channel blade switches
- Cisco Nexus 5000 and 5500 Platform switches
- Cisco Unified Computing System 6100 and 6200 Series Fabric Interconnects (NPV is default)

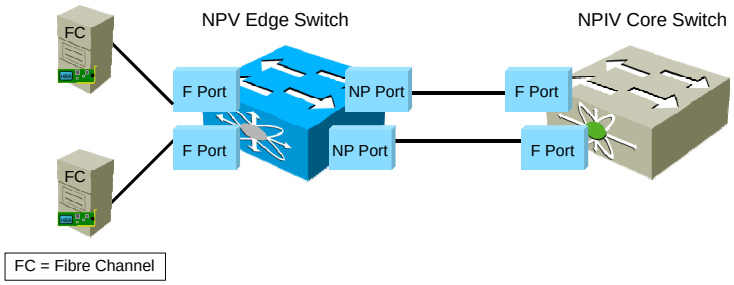
NPIV is supported on the following NPV core switches:

- Cisco MDS 9500 Series Multilayer Directors
- Cisco MDS 9222i Multiservice Modular Switch and 9216A and 9216 Multilayer Fabric Switches
- Cisco MDS 9124, 9134 and 9148 Multilayer Fabric Switches
- IBM and HP Fibre Channel blade switches
- Cisco Nexus 5000 Platform switches
- Third-party switches need to support NPIV when connecting an NPV-enabled switch

https://t.me/learningnets

N-Port Virtualization

- Extension to NPIV
- Allows the blade switch or ToR fabric device to behave as an NPIV-based HBA to the core Fibre Channel switch
- NPV device aggregates the locally connected host ports (N Ports) into one or more uplinks (pseudo-interswitch links) to the core switches



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

The NPV feature is an extension to NPIV. The NPV allows the blade switch or ToR fabric device to behave as an NPIV-based HBA to the core Fibre Channel switch. The device aggregates the locally connected host ports (N Ports) into one or more uplinks (pseudo-interswitch links) to the core switch.

The ports that connect to the upstream Fibre Channel switch are called border ports. The ports that connect to the host are called server ports. The interface that is connected to the host bus adapter (HBA) is the F Port. The interface that is connected to the Fibre Channel switch is the proxy N port (NP Port).

The login process for the NP Port uplink is the same as for an HBA that is NPIV enabled; the only requirement of the core switch is that it supports the NPIV feature. As end devices log into the NPV-enabled edge switches, the FCID addresses that are assigned use the domain of the core switch. Because the connection is treated as an N Port and not an E Port to the core switch, the edge switch shares the domain ID of the core switch as FCIDs are being allocated. The edge NPV-enabled switch no longer requires a separate domain ID to receive connectivity to the fabric.

When an uplink port is enabled on a Cisco N-Port Virtualizer (Cisco NPV) device, it must first perform a fabric login (FLOGI) and process login (PRLI) to the core director and register with the Fibre Channel name server. After this login process is completed for the physical uplink, the end devices can start their login processes. In the Cisco NPV process, an uplink port must first be logged into the core director for the end devices to start their login processes. As an end device initiates its login process into the edge NPV switch, it follows the standard login process with a FLOGI to the NPV edge switch. The edge switch will then convert the FLOGI to an FDISC login and allow the login to proceed down the NPV uplink. Several NPV uplinks can be connected to either a single or multiple core switches. The Cisco implementation allows connectivity from each edge switch to multiple core switches.

https://t.me/learningnets

NPV vs. Regular Switching

Feature	Regular switching	NPV
Fibre Channel services	• All Fibre Channel services are provided (FLOGI, name server, zoning, domain server, FSPF, and management). • FSPF, zoning, and name server databases are distributed among connected switches.	Most Fibre Channel services are switched off.
Switching operation	Local switching enabled	• Device acts as proxy • Subordinate to a Fibre Channel switch
Domain ID	Each switch consumes a domain ID.	• Does not use a domain ID • No domain ID limitation
Scalability and manageability	• ISL between switches becomes a path within the FSFP routing table. • Up to 16 ISLs may belong to a port channel. • Up to 239 switches can be in a fabric (theoretical limit),	• Eliminates the need for server administrators to manage the SAN • Third-party integration • VSAN scalability
QoS	Available	Not available

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

In NPV mode, a Cisco Nexus 5000 or 5500 Platform switch operates as a proxy N port (NP Port) mode. It does not do any Fibre Channel switching itself. There are no local switching and zoning checks on the Cisco Nexus switches. The Cisco Nexus switch appears as a host to the upper-layer switches and as a Fibre Channel switch to the server that is attached to it.

Because Cisco Nexus 5000 and 5500 switches do not provide Fibre Channel services, they do not need a Fibre Channel domain ID. As a result, the Cisco Nexus 5000 and 5500 switches increase the scalability and eliminate the switch-to-switch interoperability issues. They also make management easier because they do not get involved in the FSPF and Fibre Channel policy.

The regular switch mode has these characteristics:

- All Fibre Channel services are provided.
  - FLOGI, name server, zoning, domain server, FSPF, management.
  - FSPF, zoning, and name server databases are distributed among connected switches.
- Local switching is enabled.
- The ISL between switches becomes a path within the FSPF routing table.
- Up to 16 ISLs may belong to a single port channel.
- Each switch consumes a domain ID.

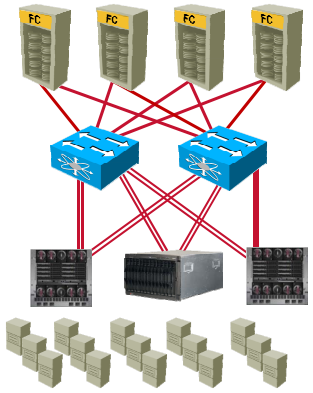
The NPV mode has these features:

- Most Fibre Channel services are switched off.
- The NPV-enabled switch now becomes a multiplexor.
- The NPV-enabled switch does not use a domain ID, which is why there is a domain ID limitation.
- There are a smaller number of switches to manage.
- It the need for server administrators to manage the SAN.

https://t.me/learningnets

NPV: Mitigation of Blade Switch Explosion

- Scalability
  - Each blade switch uses a single domain ID
  - Theoretical maximum number of 239 domain IDs per VSAN
  - Supported number of domains is smaller
- Manageability
  - More switches to manage
  - Shared management of blade switches between storage and server administrators



FC = Fibre Channel

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

This feature addresses the challenges of the blade switch explosion in the enterprise segment. The Gartner Group estimates that blade servers are the fastest-growing segment of the server market. The challenges include large increases in the number of domain IDs, increased server and network management complexity, and interoperability concerns in multivendor environments.

The Fibre Channel standards allow for up to 239 Fibre Channel domains per fabric or VSAN. However, storage vendors have only qualified up to 80 Fibre Channel domains per fabric or VSAN.

Each Fibre Channel switch is identified by a single domain ID. This effectively means that there can be no more than 40 switches that are connected together. Blade switches and ToR access layer switches will consume a domain ID, which will limit the number that can be deployed in data centers.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Summary

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

Summary

- The SCSI protocol performs the tasks of passing commands, status, and block data between hosts and storage devices.
- A Fiber Channel transaction takes place between an initiator and a target, and comprises a number of components including words, frames, sequences, and exchanges.
- Devices attaching to a Fibre Channel SAN must perform a FLOGI to be able to communicate with target devices.
- Fibre Channel addressing comprises WWNs and FCIDs.
- FSPF is the routing protocol used in the Fibre Channel SAN to route frames across multiple hops.
- VSANs are used to segment the physical fabric into multiple logical SANs.
- Zoning is used for access control within a single fabric.
- NPIV provides a means to assign multiple FCIDs to a single N Port.
- An NPV-enabled switch does not require a domain ID and is used at the access layer.

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

https://t.me/learningnets

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

Lesson 2

Describing FCoE Protocol

Overview

Unified fabric is the convergence of all the various data center I/O technologies over Ethernet. One of the most important technologies that is introduced in the Cisco Nexus Family of switches in this area is the ability to transport SAN traffic with Fibre Channel over Ethernet (FCoE). In a typical data center, the back-end servers are connected to a separate network infrastructure that is dedicated to storage traffic. Merging the devices that manage Fibre Channel and Ethernet traffic provides savings in terms of equipment purchase as well as space, power, and cooling costs. This lesson discusses the FCoE technology, the FCoE Initialization Protocol (FIP) and the related mechanisms.

Objectives

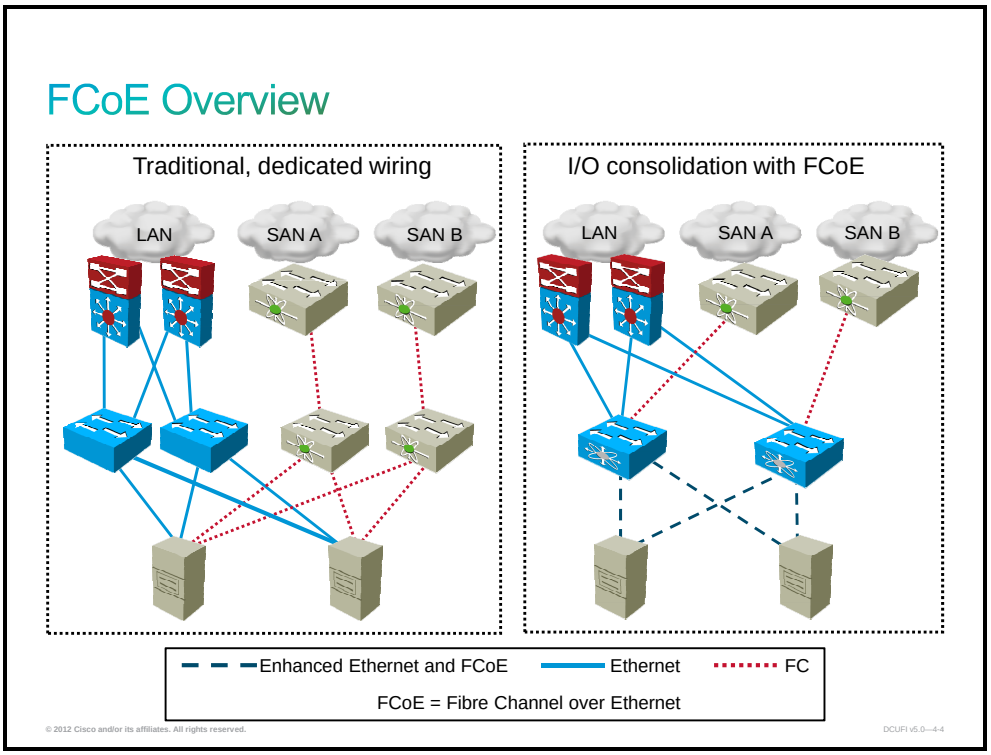
Upon completing this lesson, you will be able to explain the FCoE protocol, the adapters, the software stack, and the FIP and FCoE configuration. This ability includes being able to meet these objectives:

- Explain the need for and benefit of the FCoE protocol as well as its place in the data center
- Describe the architecture of the FCoE protocol
- Explain how the FIP is used
- Identify supported FCoE topologies
- Identify which adapters support the FCoE protocol and when to use the FCoE protocol software stack

https://t.me/learningnets

FCoE Essentials

This topic explains the need for and benefit of the FCoE protocol as well as its place in the data center.

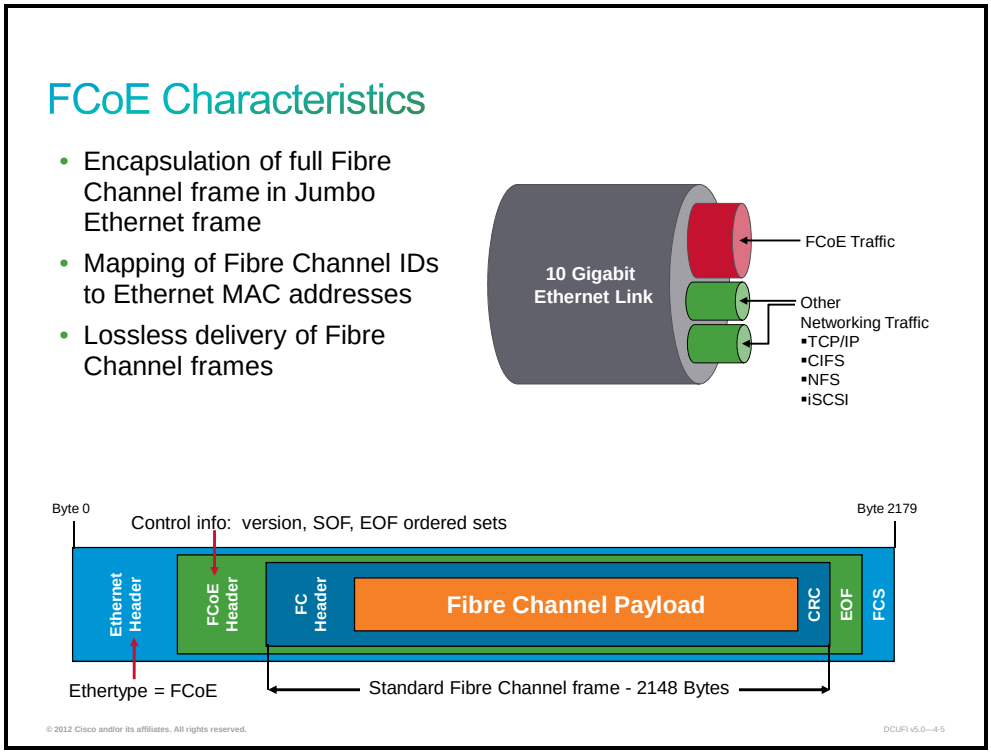


Data centers typically run multiple separate networks, including an Ethernet network for client-to-server and server-to-server communications and a Fibre Channel SAN. To support various types of networks, the traditional approach, which is shown on the left, uses separate redundant interface modules for each network: Ethernet network interface cards (NICs) and Fibre Channel interfaces in their servers, and redundant pairs of switches at each layer in the network architecture. Use of parallel infrastructures increases capital costs, makes data center management more difficult, and diminishes business flexibility.

The consolidation of I/O in the data center, shown on the right in the figure, brings the Fibre Channel and Ethernet networks into a single, integrated infrastructure. An important pillar of this consolidated approach is FCoE.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets



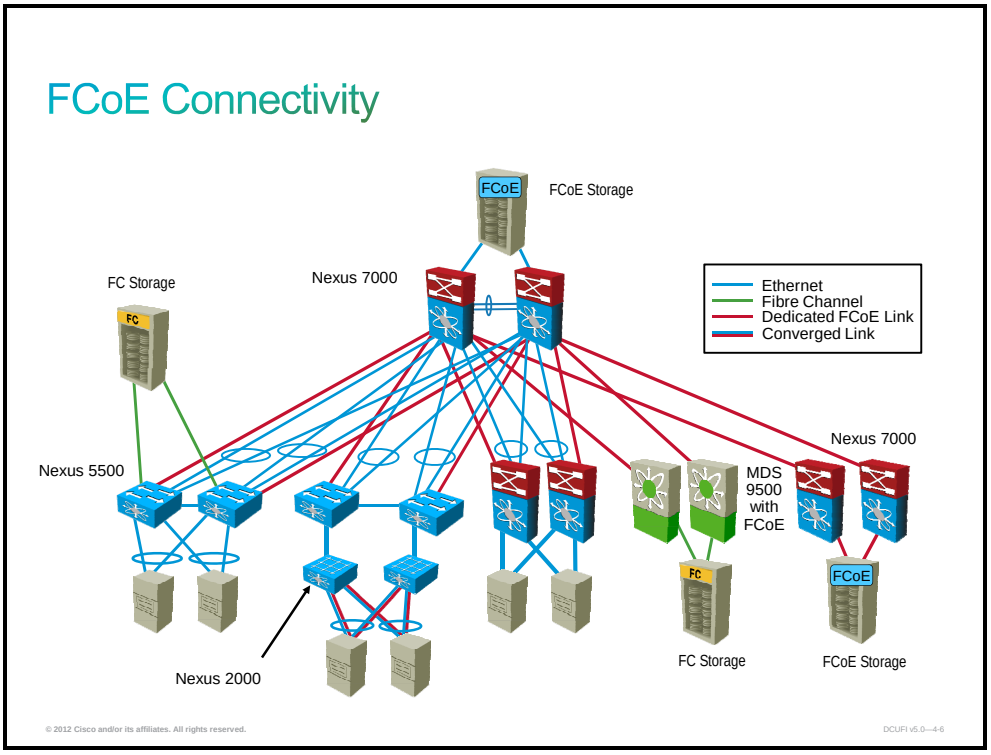
FCoE is a protocol that is based upon the Fibre Channel layers that are defined by the ANSI T11 committee. It replaces the lower layers of Fibre Channel with unified fabric I/O consolidation over Ethernet.

The minimum requirements for FCoE include the following:

- Jumbo frames. The entire Fibre Channel frame must be carried in the payload of a single Ethernet frame. (By default, jumbo frames are enabled for FCoE)
- Fibre Channel IDs must be mapped to Ethernet MAC addresses.
- Lossless delivery of Fibre Channel frames.
- A minimum of 10 Gb/s Ethernet platform.

FCoE traffic consists of a Fibre Channel frame that is encapsulated within an Ethernet frame. The Fibre Channel payload may in turn carry Small Computer Systems Interface (SCSI) messages and data.

https://t.me/learningnets



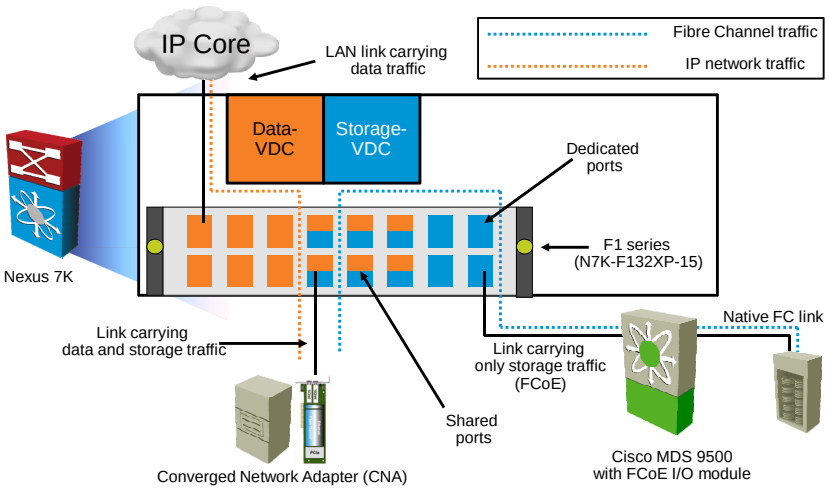
FCoE is supported by a wide range of Cisco devices, including Cisco Nexus 5000 and 5500 Platform switches, Cisco Nexus 7000 Series Switches, Cisco MDS 9000 Series Switches and 9500 Series Multilayer Directors, and Cisco Nexus 2000 Series Fabric Extenders with 10-Gb/s interfaces. FCoE endpoints can be attached using single wires or port channels. The network can consist of a mix of link types, such as dedicated FCoE interfaces and converged links.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

FCoE VDC on Nexus 7000

- Nexus 7000 support FCoE in storage VDC
- Nexus 7000 does not support FEX 2000 in FCoE mode



Cisco Nexus 7000 Series switches support FCoE in the Cisco Nexus Operating System (NX-OS) Software Release 5.2(1) and later and use a special virtual device context (VDC) type, called storage VDC, to provide the FCoE functionality. Only one storage VDC can exist on the system, and it must be a nondefault VDC. The storage should be dedicated to providing FCoE connectivity and should not fulfill unrelated tasks.

A storage VDC uses two types of interfaces; interfaces that are dedicated to it and interfaces that are shared by the storage VDC and one other VDC. The shared interfaces carry Ethernet and Fibre Channel traffic. They can be connected to hosts that are equipped with converged network adapters (CNAs). Traffic that is exchanged on shared ports is tagged. The tagging indicates the type of traffic and allows the switch to direct it to the appropriate VDC, storage or Ethernet.

Currently, FCoE is supported on Nexus 7000 F1 series modules and requires a feature license to be installed for each module.

The figure illustrates a Cisco Nexus 7000 switch that is equipped with a Cisco Nexus 7000 F1-Series Ethernet Module, such as the F1-Series 32-Port 1 and 10 Gigabit Ethernet Module (part number N7K-F132XP-15), and two nondefault VDCs, Data-VDC, and Storage-VDC. Storage-VDC is defined as the only storage VDC within the system. The server CNA is connected to a port that is shared between both VDCs. Based on the tagging information, the switch distinguishes the type of traffic and either directs the frames to Sales-VDC, and then, for example to the IP core, or via the Storage-VDC toward the MDS switch and the storage disks that are attached to it. The port connecting to the Cisco MDS switch is dedicated to the storage VDC.

https://t.me/learningnets

There are three types of links and traffic in this scenario:

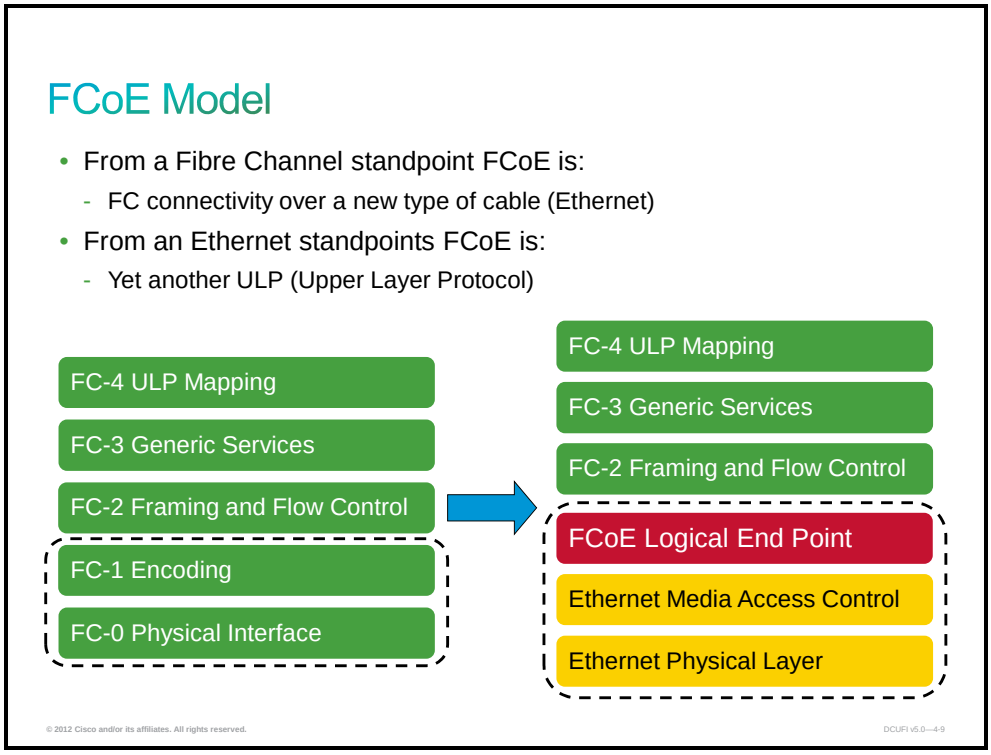
- The link between the server CNA and the F1 module port is an FCoE link and carries two types of traffic: network and Fibre Channel.
- The link between the F1 module port and the MDS switch is an FCoE link but carries only Fibre Channel traffic. Cisco Nexus 7000 does not provide any modules with native Fibre Channel ports. Such ports would present an alternative connection type to the Cisco MDS switch.
- The link between the Cisco MDS switch and the disk array is a native Fibre Channel link that carries only Fibre Channel traffic.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

FCoE Architecture

This topic shows how to understand the architecture of the FCoE protocol.



While Fibre Channel Protocols (FCPs) are designed to work only with point-to-point connectivity, FCoE switches introduce shared-media (point-multipoint) semantics. In other words, an FCoE pass-through switch is invisible to the Fibre Channel stack, but it acts as a concentrator of flows from multiple servers into the same port of a dual-stack switch, and thus it creates a shared-medium hanging off the dual-stack switch port.

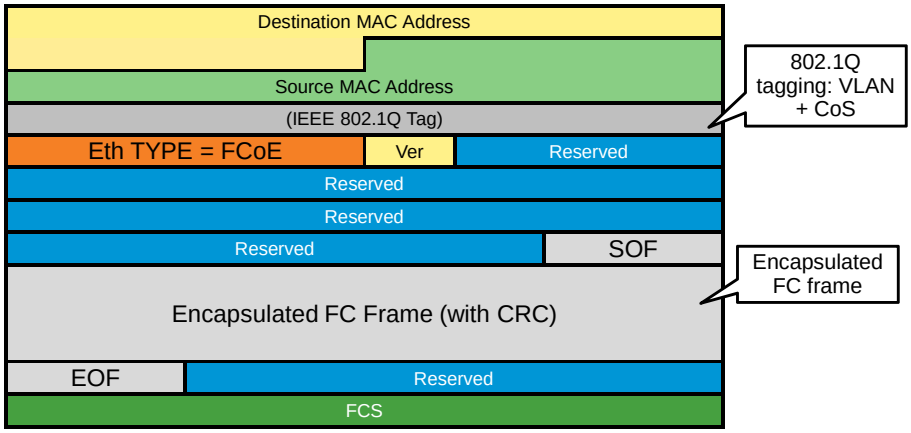
The solution in FCoE is the concept of virtual Fibre Channel (vFC) interfaces and the introduction of a small set of control plane functions to build virtual links among pairs of vFC interfaces. Multiple vFC interfaces can be presented on the same physical Ethernet interface of a dual-stack switch. After the virtual links are established among pairs of vFC interfaces, the result is presented to the Fibre Channel stack as a set of point-to-point connections that behave exactly like regular Fibre Channel connections, regardless of the number of FCoE pass-through switches that are physically inserted between the actual Fibre Channel endpoints.

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 includes the standard Fibre Channel layers starting with FC-2 and continuing up the FCP stack. The FCoE network appears to the higher-level system functions as a standard Fibre Channel network. The tools that are used in the native Fibre Channel can be thus used in an FCoE environment. Below the FCoE LEP layer you find the standard Ethernet media and physical layers for 10 Gigabit Ethernet. The Ethernet mechanisms have been enhanced to allow lossless Ethernet communication. The Ethernet standards allow the FCoE technology to take advantage of a significant amount of existing technology.

https://t.me/learningnets

FCoE Encapsulation

- EtherType 0x8906
- Maximum size 2148 bytes
- class-fcoe enables MTU of 2240 bytes



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

FCoE is encapsulated in an Ethernet frame with a dedicated EtherType of 0x8906. The packet has a 4-bit version field. All other header fields in the frame (the source and destination MAC addresses, VLAN tags, and frame markers) are standard Ethernet fields. Reserved bits pad the FCoE frame to the IEEE 802.3 minimum frame length of 64 bytes.

A Fibre Channel frame consists of 36 bytes of headers and up to 2112 bytes of data for a total maximum size of 2148 bytes. The encapsulated Fibre Channel frame has all the standard headers, which allows it to be passed to the storage network without further modification. To accommodate the maximum Fibre Channel frame in an FCoE frame, the class-fcoe is defined with a default maximum transmission unit (MTU) of 2240 bytes.

The Ethernet frames that are sent by the switch to the adapter include the IEEE 802.1Q tag. The 802.1Q tag also includes a VLAN field. It also includes a field for the class of service (CoS) value that is used by priority flow control (PFC).

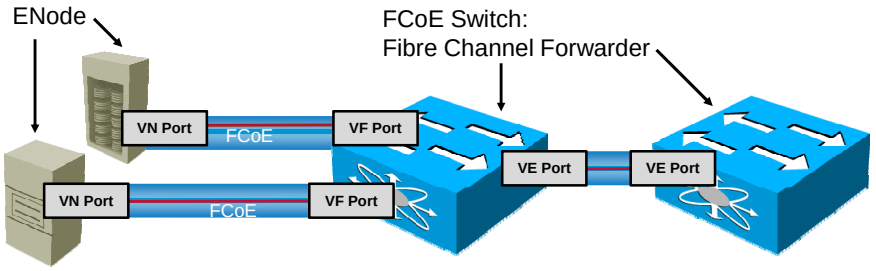
The Cisco Nexus 5000 and 5500 Platform switches expect frames from a FIP T11-compliant CNA to be tagged with the VLAN tag for the FCoE VLAN. Frames that are not correctly tagged are discarded.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

FCoE Elements and Ports

- FCoE end node (ENode)
  - FCoE endpoint
  - Has an adapter operating as virtual N Port (VN Port)
- Fibre Channel Forwarder (FCF)
  - FCoE switch
  - Connects to ENode via virtual F port (VF Port)
  - Connects to other FCFs via virtual expansion ports (VE Ports)



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

The FCoE standard defines two types of endpoints for the Ethernet encapsulation of Fibre Channel frames: FCoE end nodes (ENodes) and Fibre Channel forwarders.

End nodes are the combination of FCoE termination functions and Fibre Channel stack on the CNAs, and in that sense they are equivalent to host bus adapters (HBAs) in native Fibre Channel networks. FCFs are the combination of FCoE termination functions and Fibre Channel stack on Ethernet switches (dual-stack switches) and are therefore equivalent to Fibre Channel switches in native Fibre Channel networks.

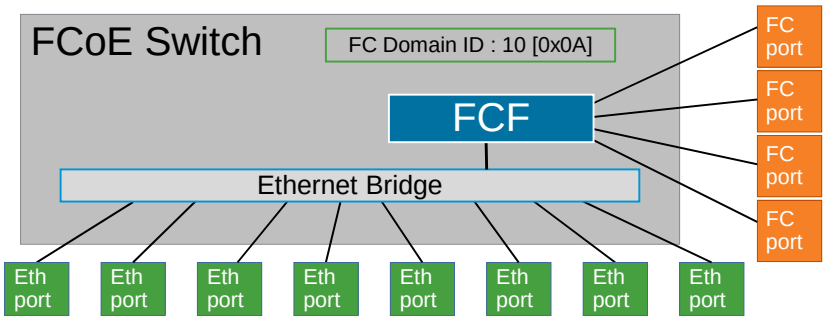
End nodes present vFC interfaces in the form of virtual node ports (VN Ports), which can establish FCoE virtual links using the virtual fabric ports (VF Ports) of the FCFs. The FCFs present vFC interfaces in the forms of VF Ports or virtual expansion ports (VE Ports); a VF Port establishes FCoE virtual links with the VN Port of a CNA, and VE Ports enable FCFs to establish FCoE virtual links with one another. These interface types have their equivalents in native Fibre Channel node ports (N Ports), fabric ports (F Ports), and expansion ports (E Ports).

Note	FCoE takes the same Fibre Channel elements and ports and virtualizes them for the use with consolidated I/O.

https://t.me/learningnets

Fibre Channel Forwarder

- FCF (Fibre Channel Forwarder) is the Fibre Channel switching element inside an FCoE switch
  - Fibre Channel logins (FLOGIs) occur at the FCF
  - Consumes a Domain ID
- FCoE encap/decap happens within the FCF
  - Forwarding based on FC information



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

A Fibre Channel Forwarder (FCF) is a switch that actually contains both a Fibre Channel switch and an Ethernet switch. One or more FCoE LEPs are used to attach the Fibre Channel switch to the Ethernet switch.

The FCF processes the Fibre Channel logins. It performs the encapsulation and de-encapsulation of the FCoE frames. It consumes one domain ID and forwards the FCoE traffic that is based on the Fibre Channel information that is contained in the frames.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

FCoE Initialization Protocol

This topic explains how the FIP is used.

FCoE and FIP

- FCoE is really two different protocols: FCoE and FIP
- Two different frame formats
- Both defined in FC-BB-5

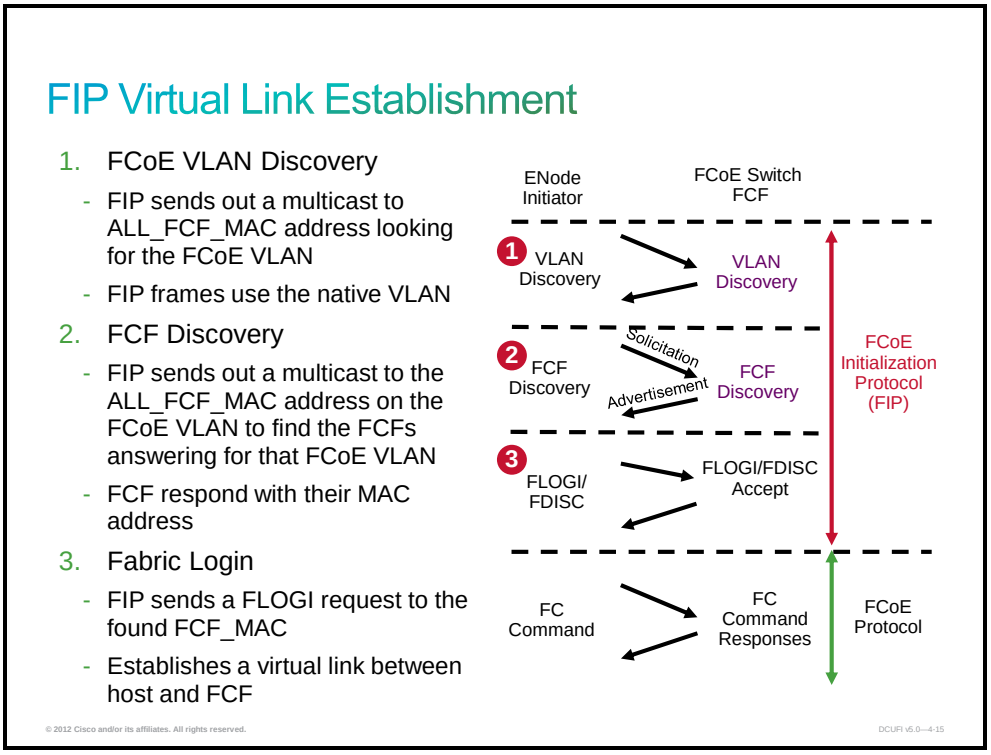
	FCoE	FCoE Initialization Protocol (FIP)
Plane	Data plane	Control plane
Purpose	• Transport most of the FC frames • Transport all SCSI traffic	• Discover FCFs connected to FCoE VLAN • Establishes a virtual link between adapter and FCF or between two FCFs • Login to and logout from FC fabric
EtherType	0x8906	0x8914

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

FIP is the FCoE control protocol that is responsible for establishing and maintaining Fibre Channel virtual links between pairs of FCoE devices (ENodes or FCFs). After the virtual link is established, Fibre Channel payloads can be exchanged on the virtual link, and FIP remains in the background to perform virtual link maintenance functions. It continuously verifies reachability between the two virtual Fibre Channel interfaces on the Ethernet network, and it offers primitives to delete the virtual link in response to administrative actions to that effect.

The FIP frames and the FCoE frames use different EtherTypes and encapsulations. FCoE frames encapsulate native Fibre Channel payloads. The table in the figure summarizes the differences between FCoE and FIP.

https://t.me/learningnets



The illustration in the 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.

During the virtual link establishment phase, FIP first discovers FCoE VLANs and remote vFC interfaces; then it performs virtual link initialization functions (fabric login [FLOGI] and fabric discovery [FDISC], or exchange link parameters [ELP]) similar to their native Fibre Channel equivalents.

Note As explained before, the FIP packets are built using a globally unique MAC address that is assigned to the CNA at manufacturing (a burned in address called server provided MAC address [SPMA]). 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. It is called fabric provided MAC address (FPMA).

FIP VLAN discovery discovers the FCoE VLAN that will be used 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 called All-FCF-MACs. All-FCF-MAC is a multicast MAC address to which all FCFs listen.

FIP FCF discovery is the protocol that is used by ENodes to discover FCFs that can accept logins. FCFs periodically send FIP FCF discovery advertisement messages on each configured FCoE VLAN; these messages are destined for the multicast MAC address All-ENode-MACs, a multicast MAC address to which all ENodes listen. The FIP FCF discovery advertisement is used by the FCF to inform any potential ENode in the VLAN that FCF VF Ports are available for virtual link establishment with the VN Ports of the ENodes. The advertisement includes the MAC address of the FCF as well as additional parameters.

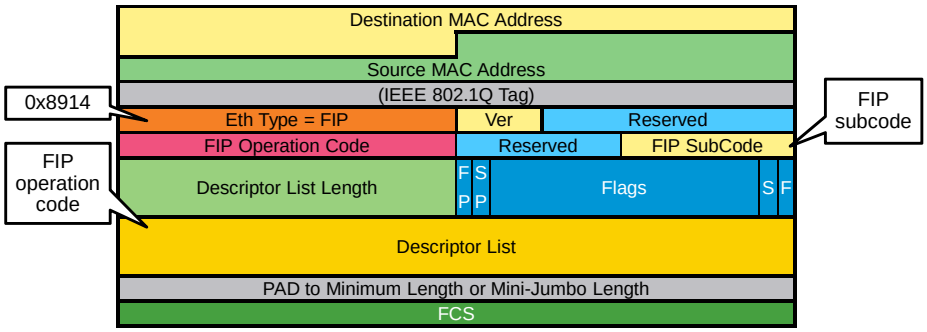
The third step in the FIP process includes the FLOGI and FDISC exchange.

© 2012 Cisco Systems, Inc.
https://t.me/learningnets

FIP Frame Format

FIP operation codes and subcodes identify the message types:

- Discovery Solicitation and Discovery Advertisement
- Virtual Link Instantiation Request and Virtual Link Instantiation Reply
- FIP Keep Alive and FIP Clear Virtual Links
- FIP VLAN Request and FIP VLAN Notification



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

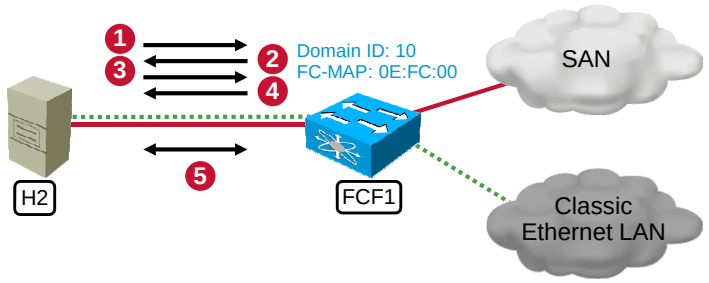
FIP is encapsulated in an Ethernet packet with a dedicated EtherType, 0x8914. The packet has a 4-bit version field, the source and destination MAC addresses, a FIP operation code, and a FIP operation subcode. The following list describes the FIP operation codes and subcodes:

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

https://t.me/learningnets

FIP Process

1. Host solicitation
2. Switch provides the fabric unique FC-MAP
3. Host performs Fabric Login (FLOGI)
4. FCF provides FC-ID
5. Host uses the FPMA for subsequent transmissions



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

The FIP communication can be broken down into four steps:

1. Host solicitation
2. Switch provides the fabric-unique FC-MAP
3. Host performs FLOGI
4. FCF provides Fibre Channel ID (FCID)

At this point, the host starts encapsulating SCSI commands and payloads using the fabric provided MAC address (FPMA) that was obtained from the FCF.

FIP VLAN discovery discovers the FCoE VLAN that will be used by all other FIP protocols as well as by the FCoE encapsulation for Fibre Channel payloads on the established virtual link. One of the goals of FC-BB-5 was to be as nonintrusive as possible on initiators and targets, and therefore FIP VLAN discovery occurs in the native VLAN that is used by the initiator or target to exchange Ethernet traffic. The FIP VLAN discovery protocol is the only FIP protocol running on the native VLAN; all other FIP protocols run on the discovered FCoE VLANs.

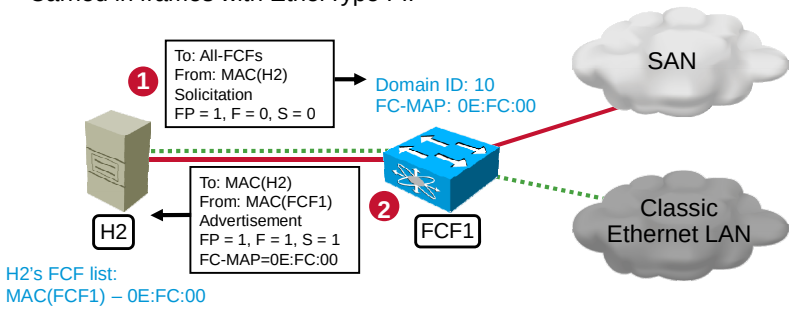
The ENode sends a FIP VLAN discovery request to a multicast MAC address called All-FCF-MACs, which is a multicast MAC address to which all FCFs listen. All FCFs that can be reached in the native VLAN of the ENode are expected to respond on the same VLAN with a response that lists one or more FCoE VLANs that are available for the VN Port login of the ENode. This protocol has the sole purpose of allowing the ENode to discover all the available FCoE VLANs, and it does not cause the ENode to select an FCF.

FIP VLAN discovery is an optional protocol in FC-BB-5. An ENode implementation can choose to offer only manual configuration for FCoE VLANs, and therefore choose not to perform FIP VLAN discovery.

https://t.me/learningnets

FIP Process: Solicitation

1. Host solicitation
   - Indicates the type of MAC supported by the CNA
   - Carried in frames with EtherType FIP
2. Switch provides the fabric unique FC-MAP
   - First 3 bytes of the FPMA
   - FP = 1: FPMA capable; F = 1: frame generated by an FCF
   - Carried in frames with EtherType FIP



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

FIP starts with a multicast solicitation from the host. The solicitation includes flags that indicate whether the host is capable of using FPMAs or SPMAs. The EtherType of this frame is FIP.

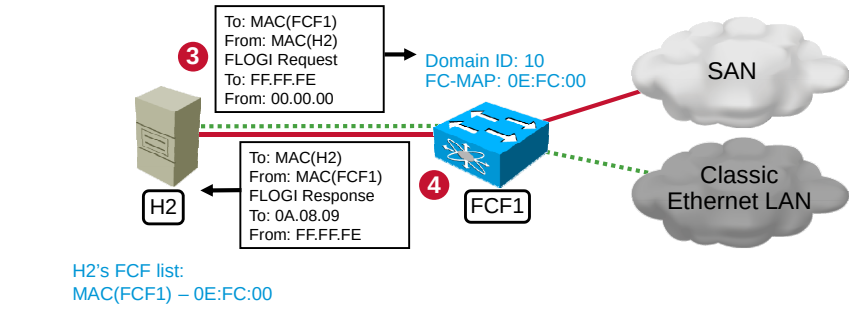
This is a frame from a host, so the F-bit is set to 0. This frame is not a solicitation response, so the S-bit is set to 0. This device is capable of using FPMAs, so the FP-bit is set to 1.

The FCF replies by sending an advertisement. The F-bit is set to 0, indicating that the FIP frame is from an FCF. The S-bit is set to 1, indicating that this is a response to a solicitation.

https://t.me/learningnets

FIP Process: FLOGI

3. Host continues FIP with FLOGI
   - At this point, host is still using the burned-in MAC address
   - Carried in frames with EtherType FIP
4. FCF provides an FCID
   - FCF replies using the host's burned-in MAC address
   - Carried in frames with EtherType FIP



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

Now that the host has an FC-MAP (a unique value per fabric MAC address prefix), it can continue with the FLOGI to finalize its FPMA. FLOGI frames are sent to the FCID of FF:FF:FE, which is the well-known address for FLOGI.

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.

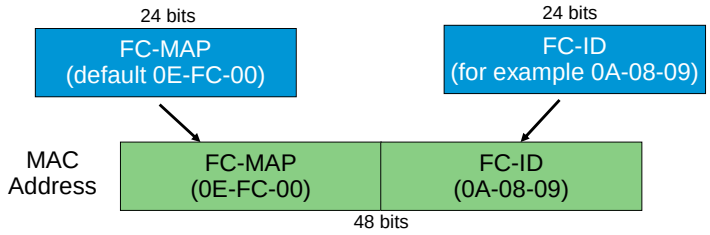
On any native Fibre Channel link between an N Port and an F Port, the first protocol exchange performed as part of activating the data-link layer is the FLOGI, which results in the assignment of an FCID to the N Port. In FIP, the logical step of FCF selection by an ENode has been merged with the native Fibre Channel FLOGI exchange. The result of this optimization is a single exchange for FCF selection, FLOGI, and FCID allocation.

FIP FLOGI and FDISC are unicast frames almost identical to the native Fibre Channel FLOGI and FDISC frames they replace. The VN Port sends an FLOGI or an FDISC request, followed by the corresponding FLOGI or "FDISC accept payload" from the FCF. Completion of this exchange terminates the FIP virtual link establishment phase.

https://t.me/learningnets

Fabric Provided MAC Address (FPMA)

- Consistent with the Fibre Channel model:
  - FCoE MAC address prefix (FC-MAP)
  - Fibre Channel ID (FCID)
- No table needed for MAC address mapping
- Multiple FC-MAPs supported
  - One per physical SAN
  - Used with N-Port ID Virtualization (NPIV)
- Must be supported by FCoE devices



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

FPMAs use the 24-bit-wide FCID that is assigned to the CNA during the FIP FLOGI and FDISC exchange, and therefore they cannot be available to the CNA before the FLOGI has occurred. The 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. Fibre Channel traffic can be encapsulated directly in FCoE frames with no table lookup because the FC-MAP is a known quantity.

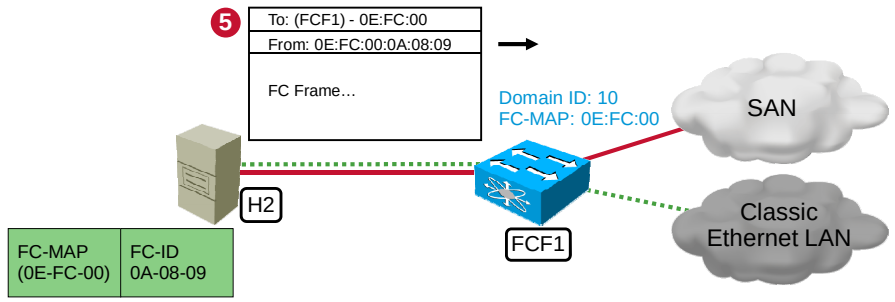
The FPMA is built by appending the FCID to a 24-bit quantity called the FCoE MAC address prefix (FC-MAP). FC-BB-5 defined a range of 256 FC-MAPs to facilitate FCoE deployments. Most users should find the default FC-MAP value 0E-FC-00 sufficient. The 256 different FC-MAPs make available to users up to 256 pools of locally unique MAC addresses. The pools are useful when the FCIDs are not unique on an Ethernet VLAN; for instance, when different Fibre Channel fabrics or different VSANs are encapsulated in the same Ethernet VLAN, the ranges of FCIDs assigned in each Fibre Channel fabric may overlap.

You should never attempt to map multiple Fibre Channel fabrics onto the same Ethernet VLAN. Most users will not ever need to map multiple Fibre Channel fabrics onto the same physical Ethernet network, but if such a need arises, each Fibre Channel fabric should be encapsulated in a separate VLAN.

https://t.me/learningnets

Subsequent FCoE Frames

5. Host uses the FPMA for subsequent transmissions
   - FPMA is the FCID obtained through FLOGI
   - FCID appended to the FC-MAP provided by the FCF
   - Carried in frames with EtherType FCoE



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

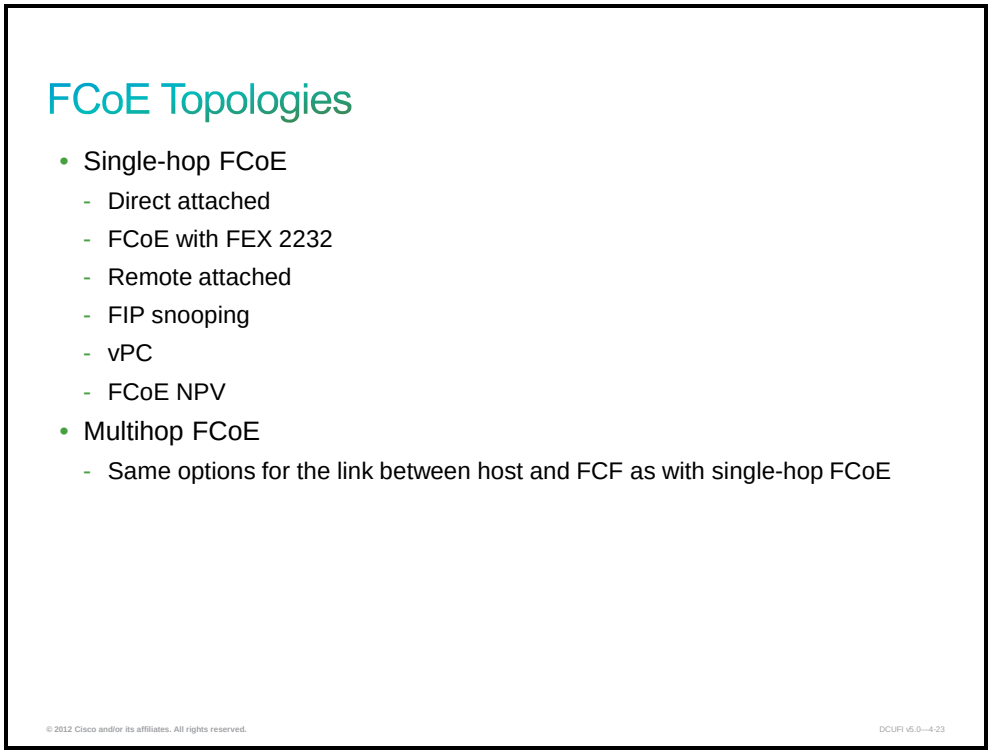
Now that the device has a complete FPMA (FC-MAP + FCID), it can communicate on the fabric using FCoE frames.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Supported FCoE Topologies

This topic identifies how FCoE works in various environments.



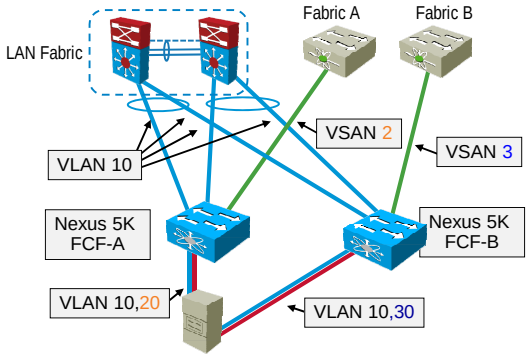
You can divide the FCoE topologies into single-hop and multihop environments. Each category supports a number of deployments between the host and the FCF:

- Direct attached
- FCoE with Cisco Nexus 2232 Fabric Extender
- Remote attached
- FIP snooping
- Virtual port channel (vPC)
- FCoE Cisco N-Port Virtualizer (Cisco NPV)

https://t.me/learningnets

Direct-Attached FCoE

- Different topology for each VLAN
  - FCoE Fabric ‘A’, FCoE Fabric ‘B’, and native Ethernet
- The topology of each VLAN is independent from other VLANs. In the figure the VLAN 10 is the LAN VLAN, VLAN 20 serves VSAN 2, and VLAN 30 serves VSAN 3.



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

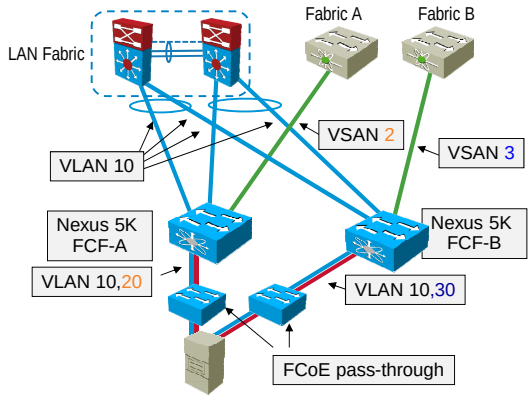
When an ENode is directly attached to an FCF, the FIP virtual link establishment that was described previously is a set of point-to-point exchanges that map one or more virtual links on top of a physical Ethernet link. In this simple scenario, a Fibre Channel stack exists at both ends of the physical wire. Although on the physical wire (at Layer 1) all packets are encoded using Ethernet encodings, the Fibre Channel stack assumes the responsibility of forwarding FCoE frames as if they were native Fibre Channel frames.

In the direct-attach model, Fibre Channel semantics apply fully to all FCoE frames from the data-link layer up.

https://t.me/learningnets

Remote-Attached FCoE

- Pass-through devices, such as pass through module
  - Do not have a Fibre Channel stack
  - Forwarding decisions based purely on Ethernet semantics
- FIP adds the most value to the FCoE solution



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

A blade switch is typically a simpler device than a traditional access-layer switch. The Ethernet blade switches behave as FCoE pass-through switches, rather than implement a full FCF. Under those conditions, users have no choice but to forego direct-attach FCoE and use remote-attach FCoE solutions. In a remote-attach FCoE solution, the virtual link that is established between a VN Port and a VF Port maps to an Ethernet path that includes one or more FCoE pass-through switches. Each FCoE pass-through switch does not have a Fibre Channel stack and therefore makes forwarding decisions that are based purely on Ethernet semantics. This type of environment is where FIP adds the most value to the FCoE solution.

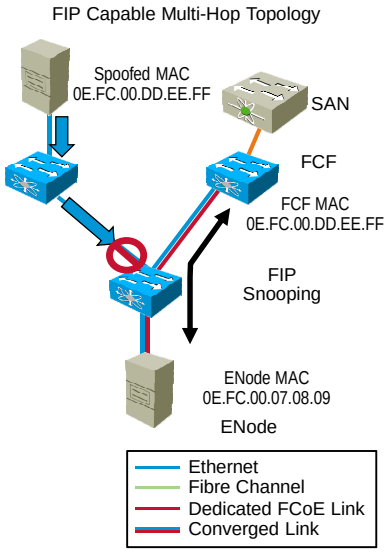
The Per VLAN Spanning Tree Plus (PVST+) mode creates an independent topology for each VLAN. This method effectively distributes the traffic load across the entire infrastructure but is also CPU- and memory-intensive because the state must be kept for each VLAN separately.

The Multiple Spanning Tree (MST) divides the VLANs into groups, which are configured through instances, and computes a separate spanning tree per group. MST achieves robust traffic load distribution while saving the CPU and memory resources on the switches.

https://t.me/learningnets

FIP Snooping

- Automatic configuration of ACLs
  - Locks down the forwarding path from CNA to FCF
  - Accomplished by snooping FIP packets
- Benefits
  - Prevention of MAC Address spoofing of FCoE end devices (ENodes)
- Fibre Channel links are Point-to-Point
- Alternative approach:
  - Manual ACLs to provide the equivalent point-to-point path control
- Support:
  - Nexus 4000 (Blade switch for IBM)



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

If a virtual link is built on top of one or more Ethernet devices that are transparent to the Fibre Channel fabric, the native Fibre Channel security model will not be sufficient to address security concerns on those devices. For this reason, FC-BB-5 includes an informative addition, Annex D, to offer a solution to the potential Fibre Channel security threats on FCoE pass-through switches.

In the industry, FCoE pass-through switches that implement either Recommendation D.4.1 or D.4.2 are commonly referred to as FIP snooping switches, to distinguish them from FCoE pass-through switches that do not offer any additional FCoE security capability on top of that.

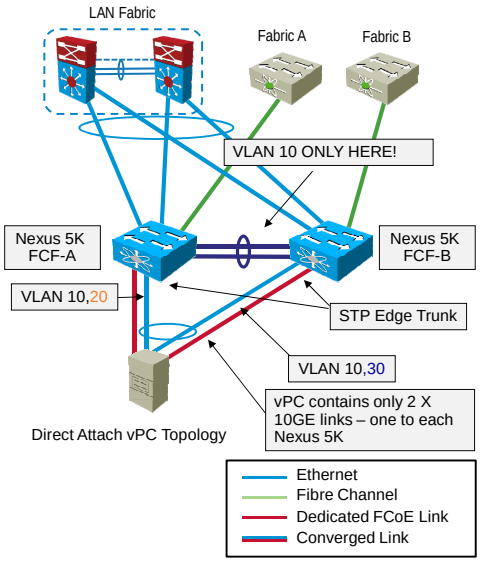
FIP snooping locks down the forwarding path from CNA to FCF by using an automatic configuration of access control lists (ACLs). This security precaution mitigates the spoofing of FCoE end devices, such as the node with MAC address 0E.FC.00.DD.EE.FF that is shown in the figure. Without FIP spoofing, such ACLs could be configured manually, but FIP spoofing automates the ACL provisioning.

FIP spoofing is supported on Cisco Nexus 4000 Series Switches.

https://t.me/learningnets

FCoE with vPC

- FCoE supported on vPC edge links
- Special considerations for FCoE traffic:
  - FCoE VLANs are ‘not’ carried on the vPC peer-link
  - FCoE and FIP EtherTypes are ‘not’ forwarded over the vPC peer link



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

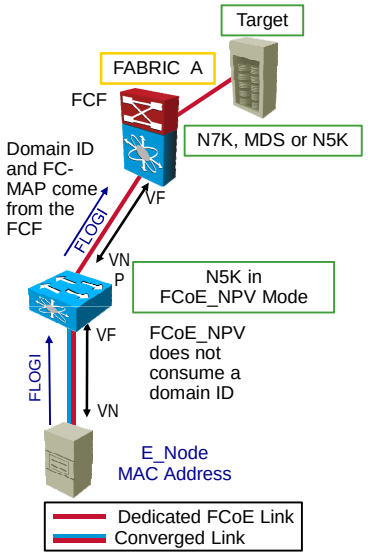
FCoE traffic can be carried over vPCs instead of single links. There are two issues that must be considered when FCoE over vPC is deployed:

- FCoE VLANs are *not* carried on the vPC peer-link. In this example, VLANs 20 and 30 are FCoE VLANs. They are not transported over the vPC peer link between FCF-A and FCF-B.
- FCoE and FIP EtherTypes are *not* forwarded over the vPC peer link. Effectively the vPC peer link is used to transport the LAN fabric traffic from and to orphan ports.

https://t.me/learningnets

FCoE NPV

- Similar to NPV in native Fibre Channel
  - No domain ID consumption
  - Consistent management
- NPV switch:
  - Pass-through device
  - FCoE VLAN assignment
  - FCF assignment
  - Login load balancing across the available FCF uplink ports
  - Takes VSANs into account when mapping or 'pinning' logins to an FCF uplink
- Upstream FCF:
  - Handles all Fibre Channel Services



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

FCoE deployments support the Cisco FCoE NPV feature that offers a similar functionality to Cisco NPV in native Fibre Channel environments. When you use the NPV feature, the access switch does not consume an FCID. The benefits include cost-effectiveness and consistent management.

The FCoE NPV switch operates as a pass-through device with several proxy functions. It assigns the FCoE VLAN and the appropriate FCF to the host. The NPV switch load balances the logins across the available FCF uplink ports, and while doing so, it takes VSANs into account when mapping or "pinning" logins to an FCF uplink.

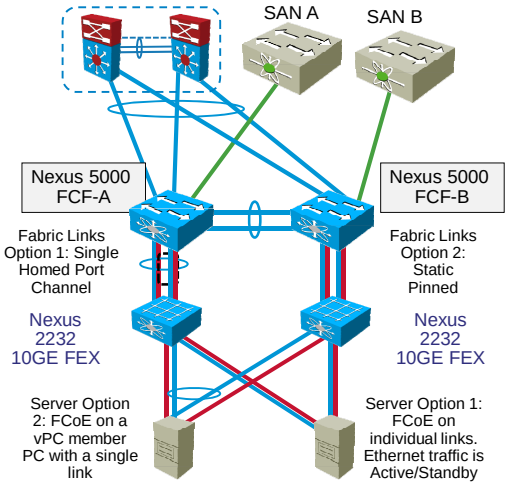
The upstream FCF passes out the FPMA addressing, consisting of the FC-MAP and FCID values and performs FCoE switching.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

FCoE with FEX 2232

- FCoE supported on FEX 2232 (10 Gigabit Ethernet)
- NIC teaming or vPC
- Only with Nexus 5000/5500 as managing switch
- FEX must be single homed
- Two options for upstream links:
  - Static pinning
  - Port channel
- Hosts require FIP Generation-2 CNAs



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

FCoE is supported on the Cisco Nexus 2232 Fabric Extender, which supports 10 Gigabit Ethernet and unified fabric. Supported downlink configurations are NIC-teaming and vPC.

Unified fabric is available only when you use a Cisco Nexus 5000 or 5500 Platform switch as a managing switch for the fabric extender (FEX). The FEX must be single-homed.

Two options exist for upstream links of the FEX:

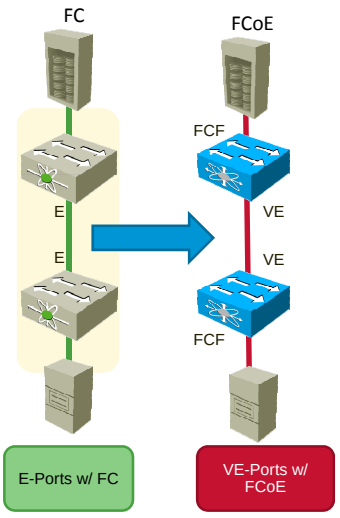
- Static pinning
- Port channel

Hosts connecting to the FCF through an FEX require a Generation-2 CNA, which is FIP-capable.

https://t.me/learningnets

Multihop FCoE

- Multihop refers to FCoE extending beyond a single “hop” or “access” switch
- Switching of FCoE frames across multiple hops
- Standards-based solution
- Fully interoperable with FC SANs
- Maintains SAN fabric separation



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

Multihop FCoE is a standards-based solution that extends the FCoE reach beyond a single “hop” or “access” switch. The solution enables the switching of FCoE frames across multiple hops.

In the multihop FCoE topology, the FCoE traffic is sent across a virtual expansion port (VE Port) that has been established between two FCFs.

A VE Port is a port that emulates an E Port over a non-Fibre Channel link. It is supported over point-to-point links between FCFs. These links can be Ethernet interfaces or members of an Ethernet port channel interface. For each of the FCF-connected Ethernet interfaces, you must create and bind a vFC interface to the Ethernet interface.

VE Ports have the following guidelines:

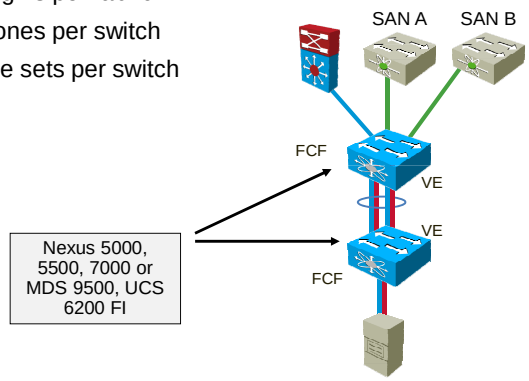
- Auto mode on the vFC interface is not supported.
- VE Port trunking is supported over FCoE-enabled VLANs.
- VE Port interface binding to MAC addresses is not supported.
- A VE Port is enabled for trunk mode by default.

The Spanning Tree Protocol (STP) is disabled on the FCoE VLANs on any interface that a vFC is bound to. This includes the interfaces that the VE Ports are bound to.

https://t.me/learningnets

Multihop FCoE Features

- Can be designed with Port Channels
- FCF supports all FC functionalities:
  - Support up to 7 hops
  - Support up to 10,000 logins per fabric
  - Supports up to 8,000 zones per switch
  - Supports up to 500 zone sets per switch
- Support:
  - Nexus 5000
  - Nexus 5500
  - Nexus 7000
  - MDS 9500
  - Cisco UCS 6200 FI



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

Multihop topologies can be deployed using single links or PortChannel connections between the FCFs.

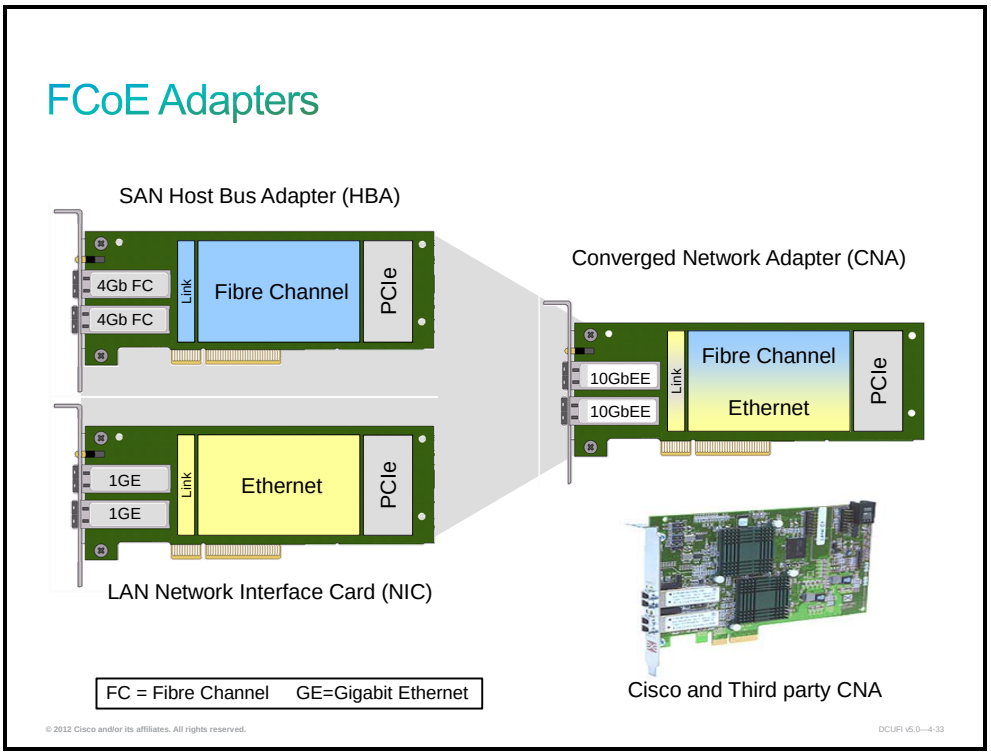
The multihop FCoE solution offers the equivalent Fibre Channel functions, such as a 7-hop network diameter, support for 10,000 logins per fabric, as well as 8,000 zones and 500 zone sets per switch.

The multihop solution is supported on Cisco Nexus 5000 and 5500 Platform switches, Cisco Nexus 7000 Series Switches, Cisco MDS 9500 Series switches, and on the Cisco Unified Computing System (UCS) 6200 Series Fabric Interconnects.

https://t.me/learningnets

FCoE Hardware

This topic identifies which adapters support the FCoE protocol and when to use the FCoE protocol software stack.



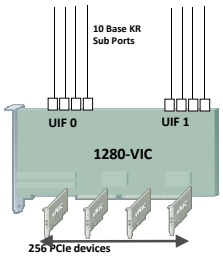
Traditionally HBAs were deployed in a Fibre Channel SAN environment, while devices in an Ethernet network environment have used NICs. To support the encapsulation and de-encapsulation of the FCoE traffic, the CNAs have been developed. A CNA combines a NIC and a Fibre Channel HBA into one card. Traditional HBA manufacturers, such as QLogic and Emulex, now also make CNAs.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Cisco UCS 1280 VIC Adapter

- New Generation VIC
- Dual 4x10 Gbps connectivity into fabric
- PCIe x16 Gen2 host interface
- Capable of 256 PCIe devices (OS dependent)
- Same host side drivers as VIC (M81KR)
- Retains 1st gen. VIC features with enhancements
- Fabric Failover capability
- SRIOV “capable” device



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

A Cisco innovation, the Cisco UCS Virtual Interface Card (VIC) 1280 is an eight-port 10 Gigabit Ethernet, FCoE-capable mezzanine card that is designed exclusively for Cisco UCS B-Series Blade Servers. The card enables a policy-based, stateless, agile server infrastructure that can present up to 256 PCI Express (PCIe) standards-compliant interfaces to the host that can be dynamically configured as either NICs or HBAs. In addition, the Cisco UCS VIC 1280 supports Cisco Virtual Machine Fabric Extender (VM-FEX) technology, which extends the Cisco UCS fabric interconnect ports to virtual machines (VMs), simplifying server virtualization deployment.

https://t.me/learningnets

FCoE Support on Nexus 7000 and MDS

Line card	N7K-F132XP-15 (F1 Series)	DS-X9708-K9	N7K-F248XP-25 (F2 Series)
Platform	Nexus 7000	MDS	Nexus 7000
Features	• 32 ports at 1/10 GbE • Lossless Ethernet Capable (PFC, ETS, DCBX) • 512 ports per system (Nexus 7018)	• 8 ports at 10 GbE (line-rate) • Lossless Ethernet Capable (PFC, ETS, DCBX)	• 48 port at 1/10 GbE • Lossless Ethernet Capable (PFC, ETS, DCBX) • 768 ports per system (Nexus 7018)
Requirements	• Storage VDC • FCoE License per Line card	• Supervisor 2A Module and Fabric2 • FCoE only ports • FCoE license	• Storage VDC • FCoE license per Line Card • Supervisor 2/2E • Fabric2
Capabilities	• 32 ports at full rate		• 48 ports at full rate

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

Cisco Nexus 7000 Series Switches and Cisco MDS switches support FCoE on certain modules only.

Cisco Nexus 7000 Series switches can run FCoE on N7K-F132XP-15 modules that offer 32 ports with speeds of 1 and 10 Gigabit Ethernet, and lossless Ethernet capabilities. The total number of FCoE ports per system cannot exceed 512. FCoE is supported only in the storage VDC and requires a separate FCoE license for each configured line card.

Cisco MDS switches support FCoE on the DS-X9708-K9 modules, which provide eight 10 Gigabit Ethernet ports with lossless Ethernet capabilities. The switch must have a Cisco Nexus 7000 Series Supervisor 2 Module and Cisco Nexus 7000 Fabric-2 Module installed in the system. FCoE-enabled ports can carry only FCoE traffic. Cisco MDS switches require a per-system FCoE license.

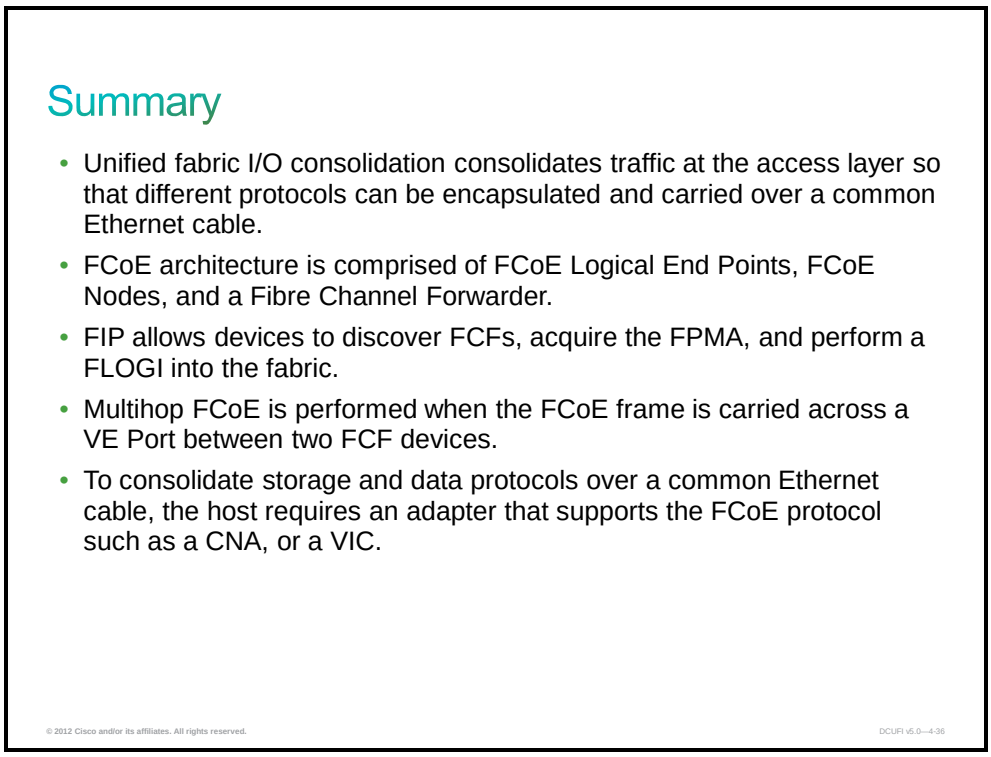
The Cisco Nexus 7000 F2-Series 48-Port 1 and 10 Gigabit Ethernet Module offers outstanding flexibility and wire-rate performance on each port. The module enables the deployment of high-density, low-latency, scalable data center architectures.

The Cisco Nexus 7000 F2-Series module is a low-latency, high-performance, high-density 10 Gigabit Ethernet module that is designed for mission-critical data center networks. Up to 768 wire-rate 10 Gigabit Ethernet ports are supported in a single system by using the Cisco Nexus 7000 18-Slot Switch chassis, providing the highest-density wire-rate 10 Gigabit Ethernet ports on the market. Populating the Cisco Nexus 7000 10-Slot and 9-Slot Switch chassis with this module delivers, respectively, up to 384 and 336 ports of 10 Gigabit Ethernet in a single chassis.

https://t.me/learningnets

Summary

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



References

For additional information, refer to these resources:

- To learn more about FCoE refer to *Fibre Channel over Ethernet Storage Networking Evolution* at: http://www.cisco.com/en/US/prod/collateral/switches/ps5718/ps6021/white_paper_c11-472771.html
- To learn more about FCoE Initiation Protocol refer to *FCoE Initiation Protocol* at : http://www.cisco.com/en/US/prod/collateral/switches/ps9441/ps9670/white_paper_c11-560403.html

https://t.me/learningnets

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

Lesson 3

Identifying DCB Enhancements

Overview

Data Center Bridging (DCB) enhancements are a required element with unified fabric. In this lesson, you will learn how DBC are used to provide needed non-drop traffic paths, QoS policies and other enhancements.

Objectives

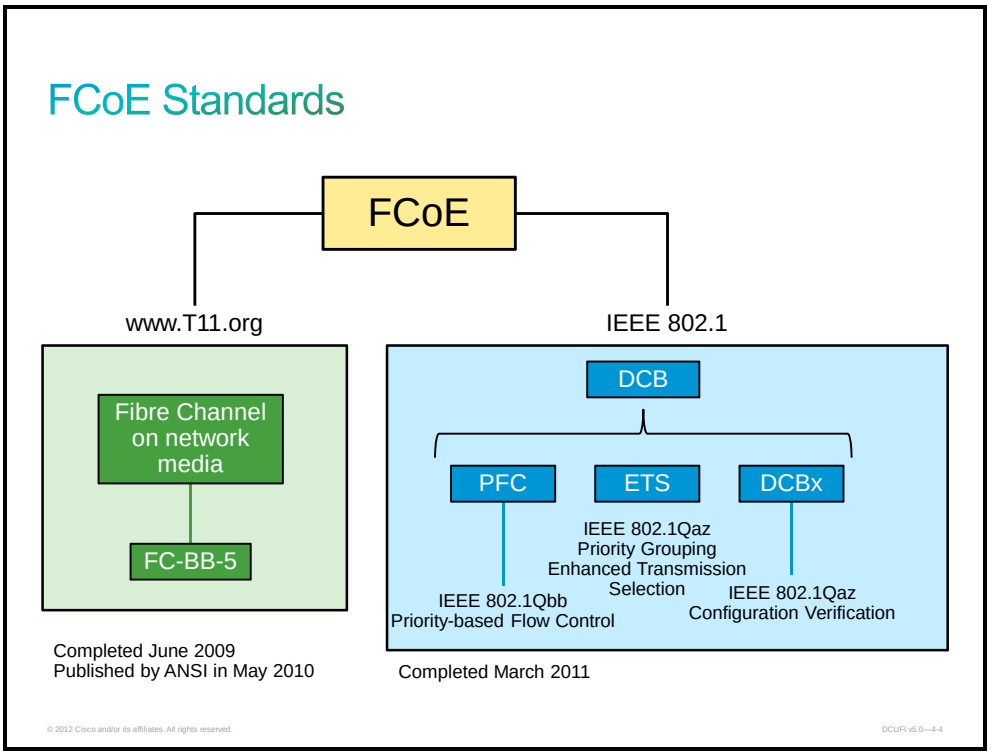
Upon completing this lesson, you will be able to identify the relevant DCB enhancements that are required to support I/O consolidation in the data center. This ability includes being able to meet these objectives:

- Identify the purpose of DCB enhancements for unified fabric I/O consolidation
- Identify the purpose of PFC
- Identify the purpose of the ETS feature
- Identify the purpose of the DCBX protocol

https://t.me/learningnets

Data Center Bridging

This topic identifies the purpose of DCB enhancements for Cisco Unified Fabric I/O consolidation.



The standardization efforts for the Fibre Channel over Ethernet (FCoE) technology encompass not only the FCoE protocol and FCoE Initialization Protocol (FIP), but also other Fibre Channel functions over a classical Ethernet infrastructure.

The Technical Committee T11 is a member of the International Committee for Information Technology Standards (INCITS) that is responsible for Fibre Channel Interfaces. The T11 committee has been producing interface standards for high-performance and mass storage applications since the 1970s. INCITS is the primary U.S. focus of standardization in the field of Information and Communications Technologies (ICT), encompassing storage, processing, transfer, display, management, organization, and retrieval of information. INCITS serves as the ANSI Technical Advisory Group for ISO/IEC Joint Technical Committee 1.

The other main standardization body is the IEEE. The 802.1 working group has four active task groups and one of them is focused on DCB.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

DCB Overview

- IEEE-based enhancements to classical Ethernet
- Ethernet enhancements
  - Priority groups: virtualizes links and allocates resources per traffic classes
  - Priority flow control by traffic class
  - End-to-end congestion management and notification
  - Layer 2 multipathing
- Benefits
  - Eliminates transient and persistent congestion
  - Lossless fabric: no-drop storage links
  - Deterministic latency for HPC clusters
  - Enables a converged Ethernet fabric for reduced cost and complexity

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

The DCB task force within the IEEE 802.1 working group has addressed the quality of service (QoS)-related challenges of Fibre Channel deployments over classical Ethernet.

Unified fabric I/O consolidation requires intermediate devices to be aware of the different classes of traffic so that appropriate levels of service can be applied. DCB incorporates the Ethernet enhancements that are required to support unified fabric I/O consolidation.

The following information summarizes what enhancements are being proposed and implemented currently:

- Priority groups
- PFC
- End-to-end congestion management
- Layer 2 multipathing

https://t.me/learningnets

DCB Standards

Technology	IEEE Standard	Description
Priority Flow Control (PFC)	802.1Qbb	Provides lossless delivery for selected classes of service (CoS)
Enhanced Transmission Selection (ETS)	802.1Qaz	Bandwidth management and priority selection
Quantized Congestion Notification (QCN)	802.1Qau	Congestion awareness and avoidance (optional)
Data Center Bridging Exchange (DCBX)	802.1Qaz	Protocol for exchanging parameters between DCB devices Leverages functions provided by Link Layer Discovery Protocol (LLDP)

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

DCB is a collection of IEEE standards that provides the foundation for a unified fabric supporting I/O consolidation.

The standards being proposed and implemented include the following:

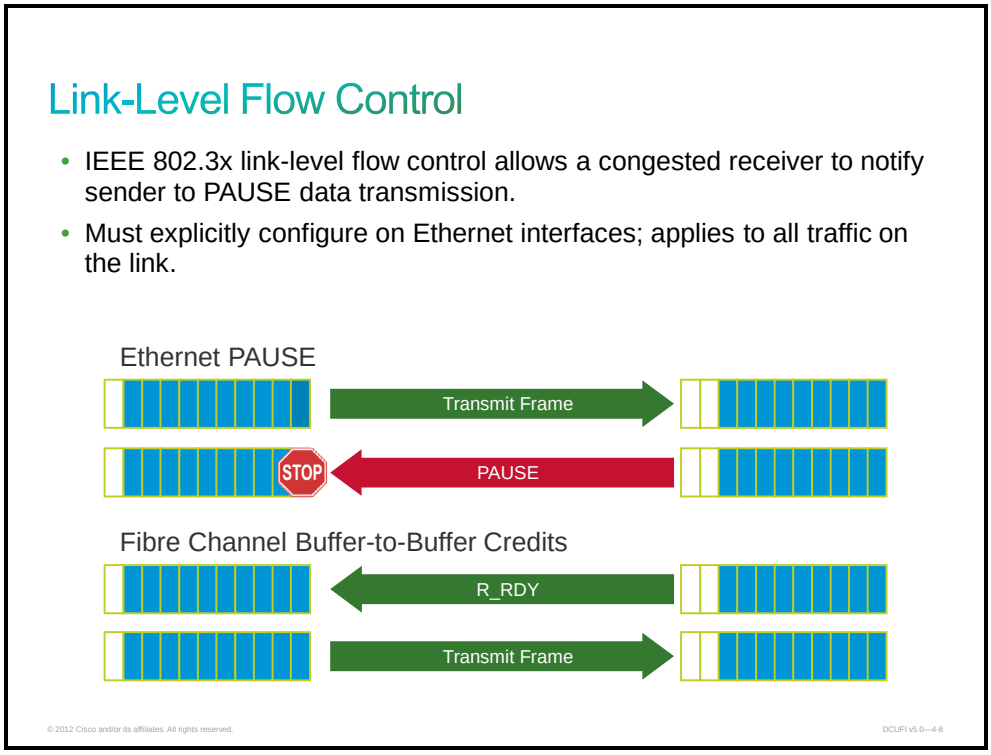
- **Priority flow control (PFC):** This is the 802.1Qbb standard providing lossless delivery for selected classes of service.
- **Enhanced Transmission Selection (ETS):** This is the 802.1Qaz standard providing bandwidth management and priority selection.
- **Quantized Congestion Notification (QCN):** This is the 802.1Qau standard providing congestion awareness and avoidance. This is an optional standard that vendors can choose to implement.
- **Data Center Bridging Exchange (DCBX):** This is the 802.1Qaz standard. It provides a protocol for exchanging parameters between DCB-capable devices and taking advantage of functions that are provided by the Link Layer Discovery Protocol (LLDP).

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Priority Flow Control

This topic identifies the purpose of PFC.



In a unified fabric providing I/O consolidation where different classes of traffic are being carried across the same physical link, it is important to be able to provide prioritization for those classes of traffic requiring it.

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 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.

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

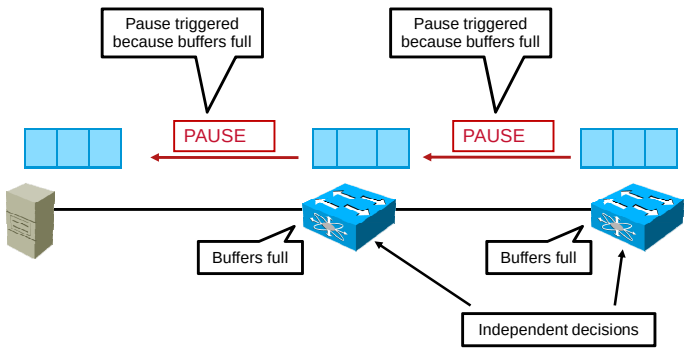
On the Cisco Nexus 5000 and 5500 Platform switches, Ethernet interfaces do not automatically detect the link-level flow control capability. You must explicitly configure the capability on the Ethernet interfaces.

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

https://t.me/learningnets

Pause Propagation

- PAUSE is a hop-by-hop mechanism
- Each hop accepts pause frames and sends pause frames independently
- Decision 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.

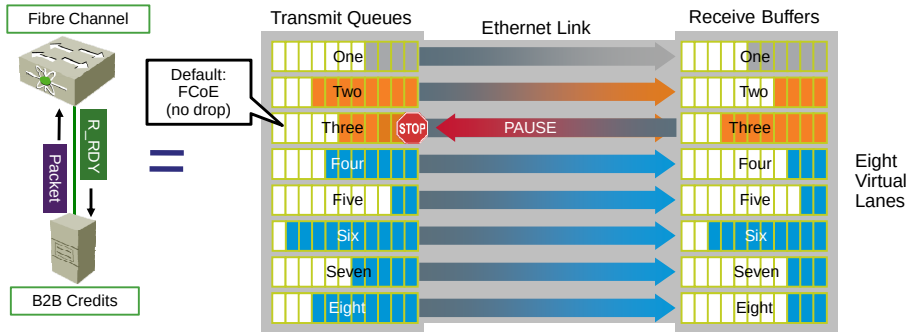
The goal is to suspend transmission of frames so that the receiver will not drop them because of congestion. When the receiving queue reaches the threshold, the switch sends a pause message back to the sender. The pause is propagated back to the servers using the same congested link.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Priority Flow Control

- Defined as 802.1Qbb for FCoE
- Enables lossless Ethernet using PAUSE based on 802.1p COS
- When link is congested, CoS assigned to “no-drop” will be PAUSED
- Other traffic continues and relies on upper layer protocol retransmission
- Solution not limited to FCoE traffic



© 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 CoS value. When the switch enables PFC, it communicates to the adapter which class of service (CoS) values to apply to 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 IEEE 802.1p CoS value that is assigned by default or 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.

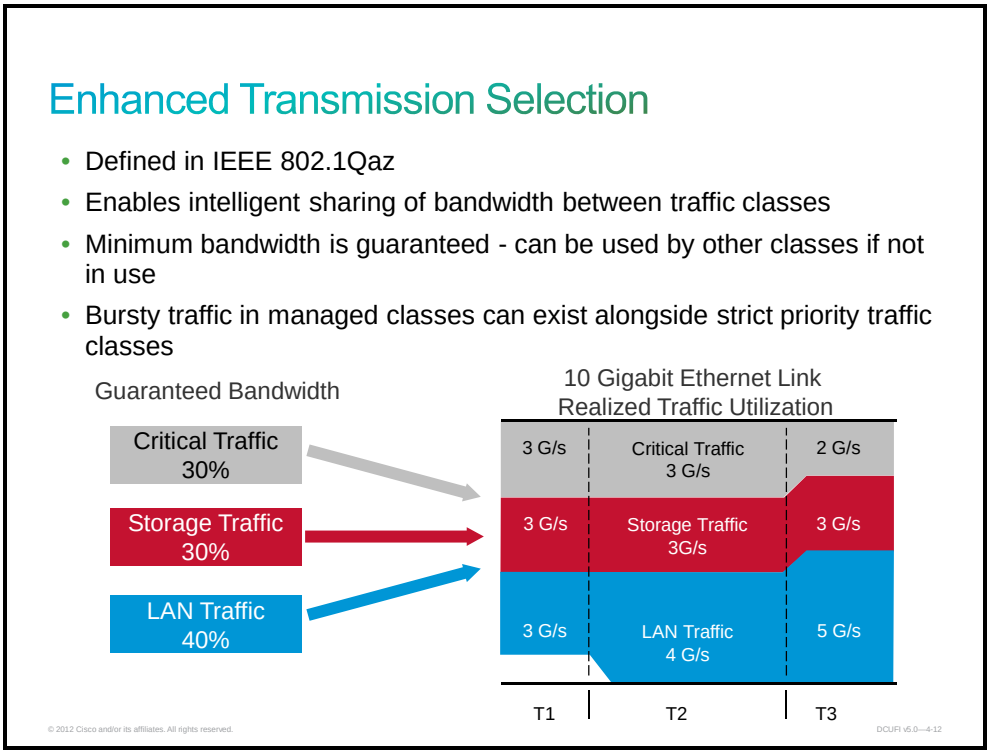
If you do not enable PFC on an interface, you can enable IEEE 802.3x link-level pause, which by default is disabled.

The PFC functionality provides an equivalent flow control mechanism to the credit-based solution in pure Fibre Channel environments.

https://t.me/learningnets

Enhanced Transmission Selection

This topic identifies the purpose of the Enhanced Transmission Selection (ETS) feature.



Bandwidth management is an important requirement when you are 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 will 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.

ETS provides advanced traffic scheduling. ETS includes features such as guaranteed minimum bandwidth for certain traffic classes like storage or high-performance computing (HPC).

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

- At T1, critical, storage, and LAN traffic use 3 Gb/s. At this point, each 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 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.

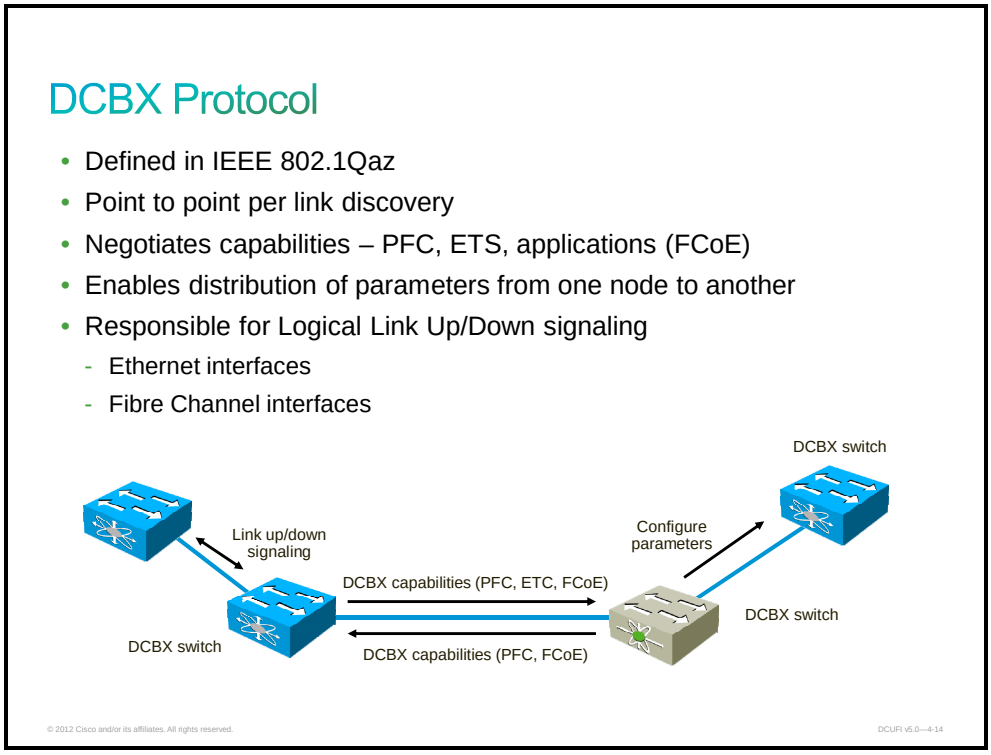
The reserved bandwidth implementation is based on deficit weighted round robin (DWRR) and strict priority could also be configured for CoS.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

DCBX Protocol

This topic identifies the purpose of the Data Center Bridging Exchange (DCBX) protocol.



The DCBX protocol, which is defined in IEEE 802.1Qaz, is a point-to-point discovery and capability exchange protocol that is used to discover peers and exchange configuration information between DCB-compliant bridges.

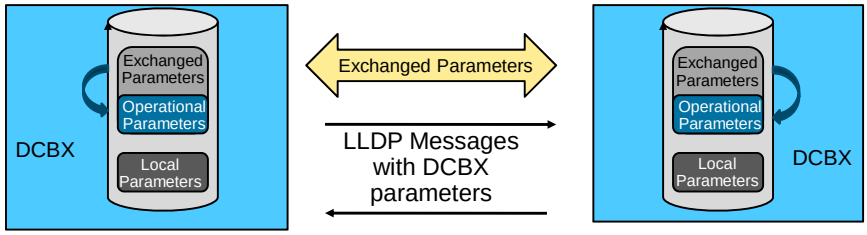
The enhancements that have been presented so far enable I/O consolidation on an Ethernet network. For these enhancements to work seamlessly with existing Ethernet deployments, a management protocol that dynamically discovers the capabilities of its peers is needed. The DCBX protocol provides these main functions:

- Capability negotiation for features, such as PFC, ETS, applications (FCoE)
- Distribution of parameters from one node to another
- Logical link up and down signaling, for both Ethernet and Fibre Channel interfaces

https://t.me/learningnets

DCBX Negotiation

- Discovery of peer DCB capabilities
- Misconfiguration detection
- Peer configuration:
  - Administered parameters – provisioned to peer device
  - Operational parameters – informational purposes only
  - Local parameters – not exchanged
- DCBX negotiation failures result in:
  - per-priority-pause not enabled on CoS values
  - vfc not coming up – when DCBX is being used in FCoE environment



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

The IEEE DCBX messages are transported over the LLDP. These are the goals of DCBX:

- **Discovery of DCB capabilities in a peer:** DCBX is used to identify the capabilities of peer devices. It provides a mechanism where a device can know if its peer device supports a particular feature such as Priority Groups or PFC.
- **Misconfiguration detection:** DCBX can be used to detect misconfiguration of a feature between the peers on a link. The misconfiguration detection is feature-specific because some features permit asymmetric configuration.
- **Peer configuration:** DCBX can be used by a device to perform configuration of DCB features in its link peer. In the initial version of DCBX, the purpose was to provide basic peer-to-peer configuration. As DCBX evolves, higher-level applications can build on it to provide more complex configuration distribution mechanisms. Each DCB feature has a set of parameters. These parameters are classified into two broad categories:
  - **Exchange parameters:** Exchange parameters are sent to the peer, and within these parameters there are two subgroups:
    - **Administered parameters:** These are configured parameters.
    - **Operational parameters:** These parameters provide the operational state of the related administered parameters. The operational state may be different from the administrative state, primarily as a result of the DCBX exchange with the peer. The operational parameters are included only for information purposes.
  - **Local parameters:** Local parameters are not exchanged in LLDP messages.

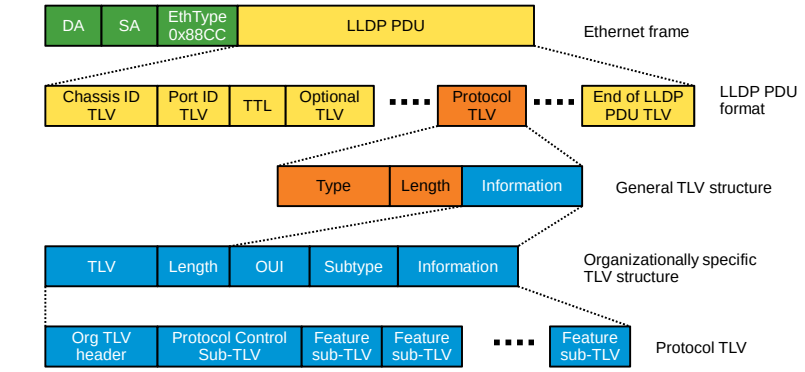
DCBX negotiation failures can lead to various states, such as:

- Per-priority-pause not enabled on CoS values
- The vFC not coming up—when DCBX is being used in FCoE environment

https://t.me/learningnets

DCBX Packets

- Transported in Link Layer Discovery Protocol (LLDP) PDUs
- Easily extendable by Type Length Values (TLV)



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

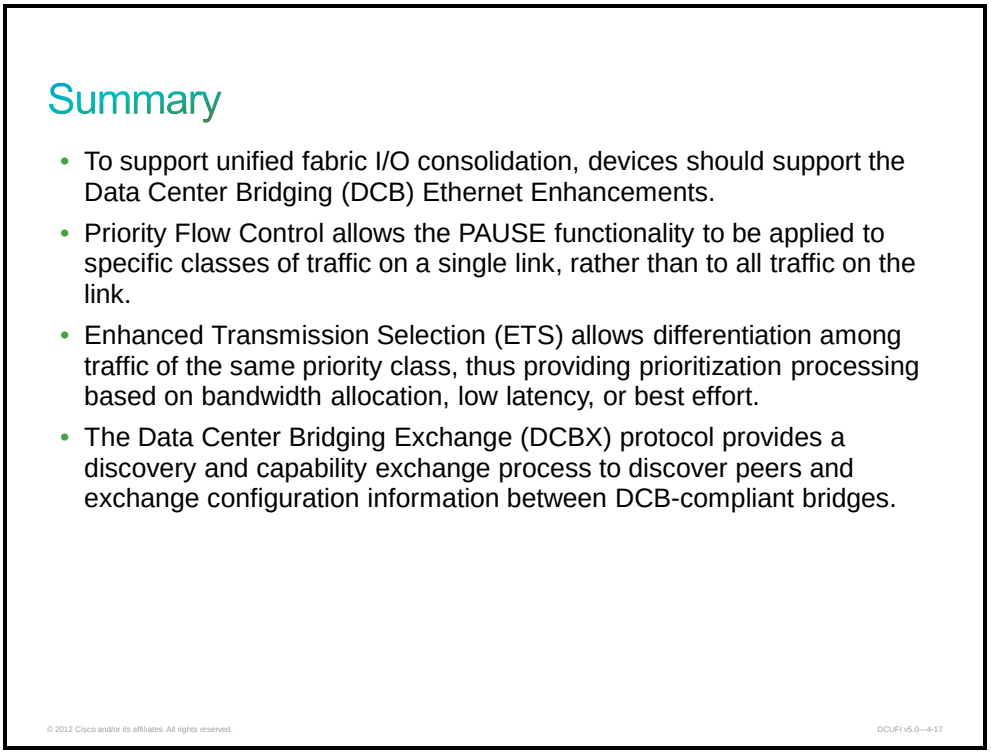
The DCBX protocol is an extension of the LLDP. DCBX endpoints exchange request and acknowledgment messages. For flexibility, parameters are encoded in a type, length, value (TLV) format.

DCBX uses LLDP IEEE 802.1AB, and defines new TLVs for capability exchange parameters. The protocol state machines manage local operational configuration for each feature by comparing and synchronizing with the feature settings of the peer.

https://t.me/learningnets

Summary

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



© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Lesson 4

Configuring FCoE

Overview

In this lesson, you will learn how to configure Fibre Channel over Ethernet (FCoE) features on the Cisco Nexus 7000 Series Switches and the switches in the Cisco Nexus 5000 and 5500 Platforms (in the Cisco Nexus 5000 Series Switches). Apart from a generic FCoE implementation, you will learn how to deploy FCoE with enhanced virtual port channel (vPC) and Cisco Adapter Fabric Extender (FEX).

Objectives

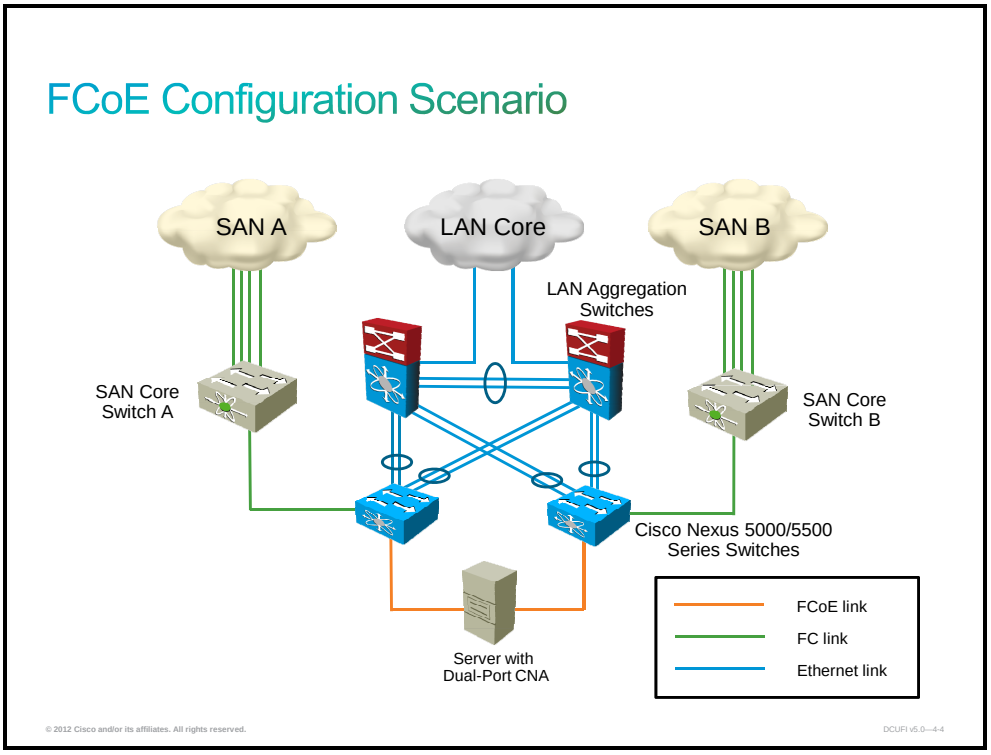
Upon completing this lesson, you will be able to configure the FCoE protocol on a Cisco Nexus switch. This ability includes being able to meet these objectives:

- Configure FCoE on a Cisco Nexus switch
- Configure FCoE VLANs and vFC interfaces
- Configure FCoE with Enhanced vPC and Cisco Adapter FEX

https://t.me/learningnets

FCoE Configuration

This topic describes how to configure FCoE on a Cisco Nexus switch.



The illustration shows a Cisco Nexus 5000 or 5500 Platform switch being deployed as a Fibre Channel forwarder (FCF) with directly-connected converged network adapters (CNAs).

If an FCF is only reachable through a transit Cisco Nexus 5000 Platform switch FCF, the CNA cannot discover or log into the FCF. Because of a hardware limitation, the Cisco Nexus 5000 Platform switch cannot perform the FCoE transit function between a CNA and another FCF.

Therefore, you must design your network topology so that the active Spanning Tree Protocol (STP) path of the FCoE VLAN is always over the directly-connected links between the CNA and the FCF. The FCoE VLAN should be configured on the directly connected links only.

The illustration shows a best practice topology for an access network using directly connected CNAs with the Cisco Nexus 5000 and 5500 Platform switches.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Cisco Nexus 5000 FCoE Configuration Procedure

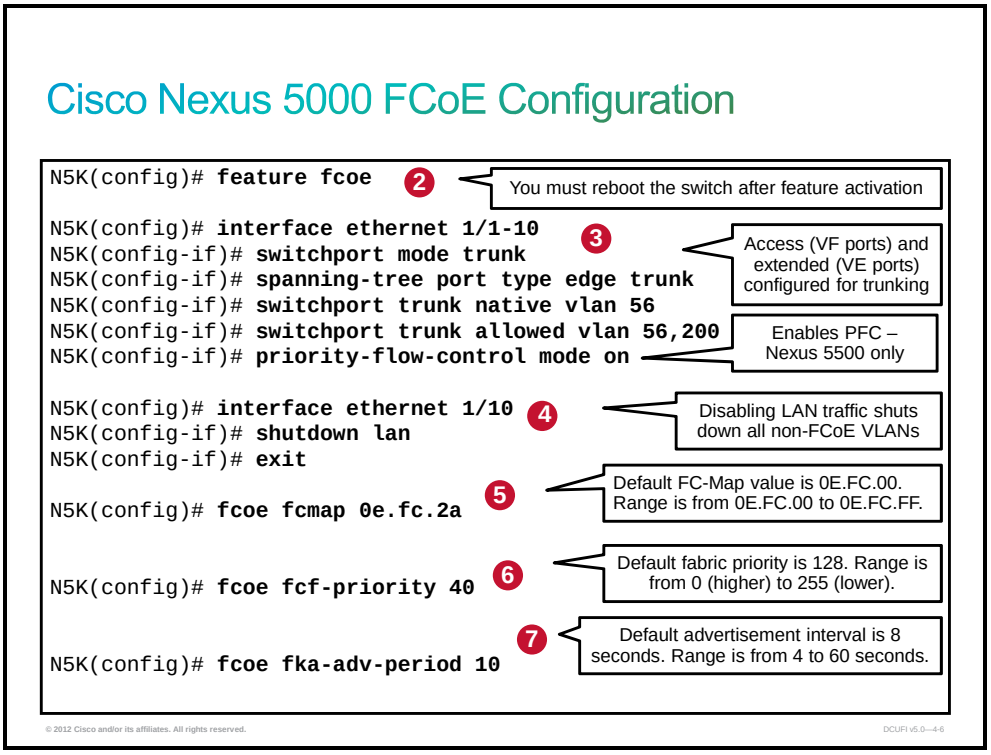
1. Make sure that correct license is installed
2. Enable FCoE feature
3. Configure FCoE interfaces for trunking and flow control (Nexus 5500 only)
   - PortFast (port type edge trunk)
   - FCoE VLANs allowed
4. Disable LAN traffic on an FCoE link (optional)
   - Switch sends a DCBX LAN Logical Link Status (LLS) message to CNA
   - Brings down all VLANs on the interface that are not enabled for FCoE
5. Configure the FCoE MAC Address Prefix (FC-MAP) (optional)
   - Provided to hosts in the FIP advertisement
   - Switch discards the MAC addresses that are not part of the current fabric
6. Configure the fabric priority advertised by the switch (optional)
   - Used by CNAs to determine the best switch to connect to
7. Set the advertisement interval (optional)

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

Follow this configuration procedure to configure general FCoE on the Cisco Nexus 5000 and 5500 Platform switches:

1. Ensure that the correct license is installed.
2. Enable the FCoE feature.
3. Configure FCoE interfaces trunk ports. The trunk ports must be configured as PortFast interfaces using the **switchport port type edge trunk** command. Ensure that the FCoE VLANs are allowed on the trunk ports.
4. Optionally, disable LAN traffic on an FCoE link. When this feature is enabled, the switch sends a Data Center Bridging Exchange (DCBX) LAN Logical Link Status (LLS) message to the attached CNA. In addition, the switch brings down all VLANs on the interface that are not enabled for FCoE. The SAN traffic is transported without any interruption.
5. Optionally, configure the FCoE MAC address prefix (FC-MAP). The FC-MAP is provided to end nodes in the FCoE Initialization Protocol (FIP) advertisements. The switch uses the FC-MAP to discard all frames from the MAC addresses that are not part of the current fabric.
6. Optionally, configure the fabric priority that is advertised by the switch. The fabric priority is used by CNAs to determine the best switch to connect to.
7. Optionally, tune the advertisement interval.

https://t.me/learningnets

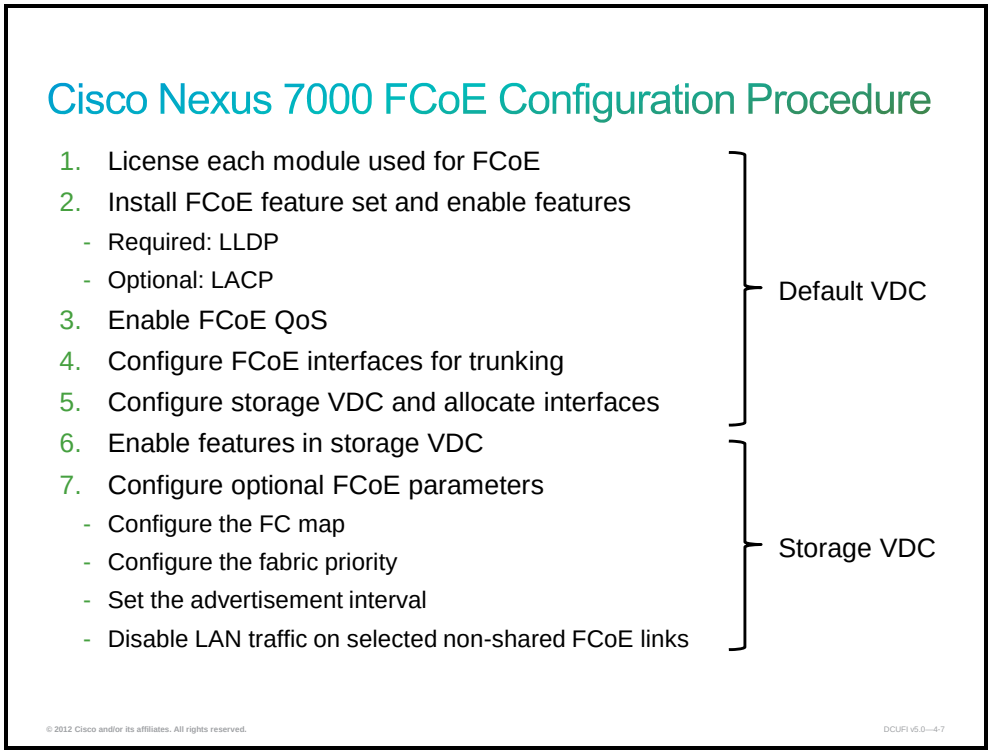


The sample configuration presents general FCoE settings that are configured on the Cisco Nexus 5000 and 5500 Platform switches. All FCoE interfaces, including the virtual fabric ports (VF Ports) and virtual expansion ports (VE Ports) are configured for trunking with the PortFast feature.

VLAN 200 will be used in this scenario to transport FCoE traffic and is therefore allowed on the trunk port. FCoE VLAN 200 must not be configured as native VLAN. In this example, VLAN 56 is configured as native VLAN. VLAN 56 is not used for FCoE.

The remaining optional configuration disables LAN traffic on Ethernet 1/10, changes the default Fibre Channel map value 0E.FC.00 to 0E.FC.2A, lowers the fabric priority from the default value 128 to 40, and extends the advertisement interval from the default value of 8 seconds to 10 seconds.

© 2012 Cisco Systems, Inc.
https://t.me/learningnets



The FCoE configuration procedure on the Cisco Nexus 7000 Series Switches differs slightly from the procedure that is performed on the Cisco Nexus 5000 and 5500 Platform switches, because of the use of the storage virtual device context (VDC).

The initial configuration is performed in the default VDC, and includes the licensing of each Cisco Nexus 7000 F Series module that is configured for FCoE, installing FCoE feature set, enabling features, quality of service (QoS), configuring FCoE interfaces, and allocating them to the storage VDC.

LLDP and LACP are Cisco Nexus 7000 Series feature names. The commands are:

- Feature lldp
- Feature lacp

So during FCOE configuration those two commands have to be typed.

The **Link Layer Discovery Protocol (LLDP)** is an IEEE standardized protocol used by network devices for advertising their identity, capabilities, and neighbors local area network.

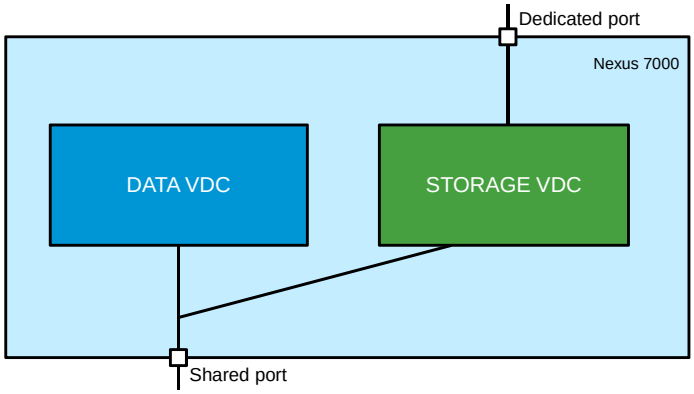
Link Aggregation Control Protocol (LACP) is part of an IEEE specification (802.3ad) that allows you to bundle several physical ports together to form a single logical channel. LACP allows a switch to negotiate an automatic bundle by sending LACP packets to the peer. It performs a similar function as Port Aggregation Protocol (PAgP) with Cisco EtherChannel.

The configuration of the storage VDC is equivalent to the FCoE settings on the Cisco Nexus 5000 and 5500 Platform switches and includes the enabling of required and optional features, and setting optional FCoE parameters.

https://t.me/learningnets

Cisco Nexus 7000 FCoE Configuration

Shared and Dedicated Ports



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

FCoE traffic can arrive at the Cisco Nexus 7000 switch over either a dedicated or a shared port.

Cisco Nexus 7000 FCoE Configuration

```
N7K(config)# license fcoe module 2                                  (1) License module 2

N7K(config)# install feature-set fcoe                               (2) Install and enable features
N7K(config)# feature lldp

N7K(config)# system qos                                             (3) QoS policy provides preferred treatment to FCoE
N7K(config-sys-qos)# service-policy type network-qos default-nq-7e-policy

N7K(config)# interface ethernet 2/1-10                              (4) FCoE interfaces in trunk mode as
N7K(config-if)# switchport mode trunk                                   STP edge ports
N7K(config-if)# spanning-tree port type edge trunk

N7K(config)# vdc fcoe_vdc type storage                              (5) Configure storage
N7K(config-vdc)# allocate fcoe-vlan-range 10-20 from vdc RED            VDC, allocate
N7K(config-vdc)# allocate interface ethernet 2/1-5                      interfaces and
N7K(config-vdc)# allocate shared interface ethernet 2/6-10              VLANs

N7K(config-vdc)# switchto vdc fcoe_vdc                              (6) LLDP feature required in storage VDC
N7K-fcoe_vdc# configure terminal
N7K-fcoe_vdc(config)# feature lldp

N7K-fcoe_vdc(config)# fcoe fcmap 0e.fc.2a                           (7) Optional FCoE features
N7K-fcoe_vdc(config)# fcoe fcf-priority 40
N7K-fcoe_vdc(config)# fcoe fka-adv-period 10
N7K-fcoe_vdc(config)# interface ethernet 1/1-5
N7K-fcoe_vdc(config-if)# shutdown lan
```

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

This configuration example presents the FCoE configuration in a storage VDC. There is a QoS policy configuration and PortFast configuration for the interfaces Ethernet 2/1-10. The QoS configuration enables a predefined QoS policy that grants the FCoE traffic the required preferred treatment.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

The interfaces Ethernet 2/1-10 are put into trunk mode and configured as STP-type edge ports to support STP Lite for loop prevention. Ports 2/1-5 are dedicated to storage VDC, while ports 2/6-10 are shared between the storage VDC and another VDC (RED in this example).

Apart from interface allocation, the storage VDC also needs VLANs allocated to it. This is done using the **allocate fcoe-vlan-range** command. The VLANs can be shared with another VDC (RED in this example).

Additional settings, equivalent to the configuration of the Cisco Nexus 5000 and 5500 Platform switches are entered in the storage VDC configuration mode.

FCoE Verification

- Same verification on Nexus 5000/5500 and 7000 switches
- Displays global FCoE parameters
- Run in storage VDC on Nexus 7000

```
switch(config)# show fcoe
Global FCF details
       FCF-MAC is 00:0d:ec:6d:95:00
       FC-MAP is 0e.fc.2a
       FCF Priority is 40
       FKA Advertisement period for FCF is 10 seconds
```

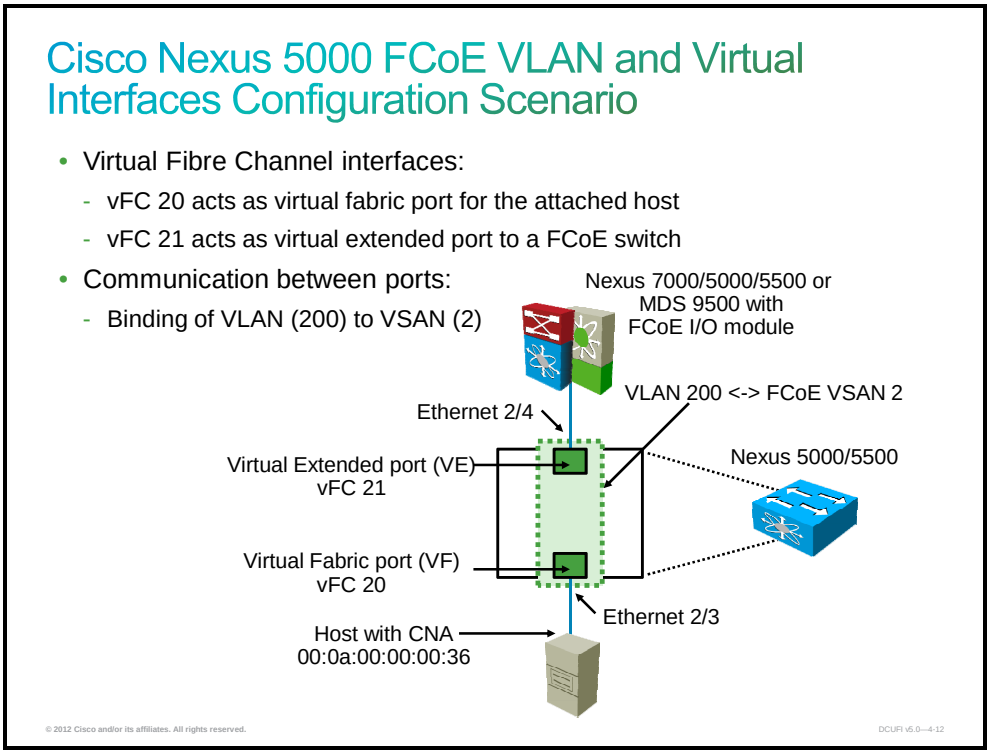
© 2012 Cisco and/or its affiliates. All rights reserved.

You can use the **show fcoe** command to display the status of FCoE parameters on the switch. You can see the values that are configured in the preceding example. The same command is available on the Cisco Nexus 5000 and 5500 Platform switches, and on the Cisco Nexus 7000 switches. On the Cisco Nexus 7000 Series switches, you should run this command in the storage VDC.

https://t.me/learningnets

FCoE VLANs and Virtual Interfaces

This topic explains how to implement FCoE VLANs and virtual interfaces.



This configuration scenario is based on a Cisco Nexus 5000 or 5500 Platform switch. The switch has been enabled for FCoE. A host with appropriate CNA is connected to Ethernet 2/3 interface. The switch is configured with two virtual Fibre Channel (vFC) interfaces: vFC 20 and vFC 21. The vFC 20 interface acts a VF Port to the attached host. The vFC 21 interface acts a VE Port to another FCoE switch. The FCoE connectivity between the vFC interfaces is provided by a FCoE VLAN (200) mapped to a FCoE VSAN (2).

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Cisco Nexus 5000 FCoE VLAN and Virtual Interfaces Configuration Procedure

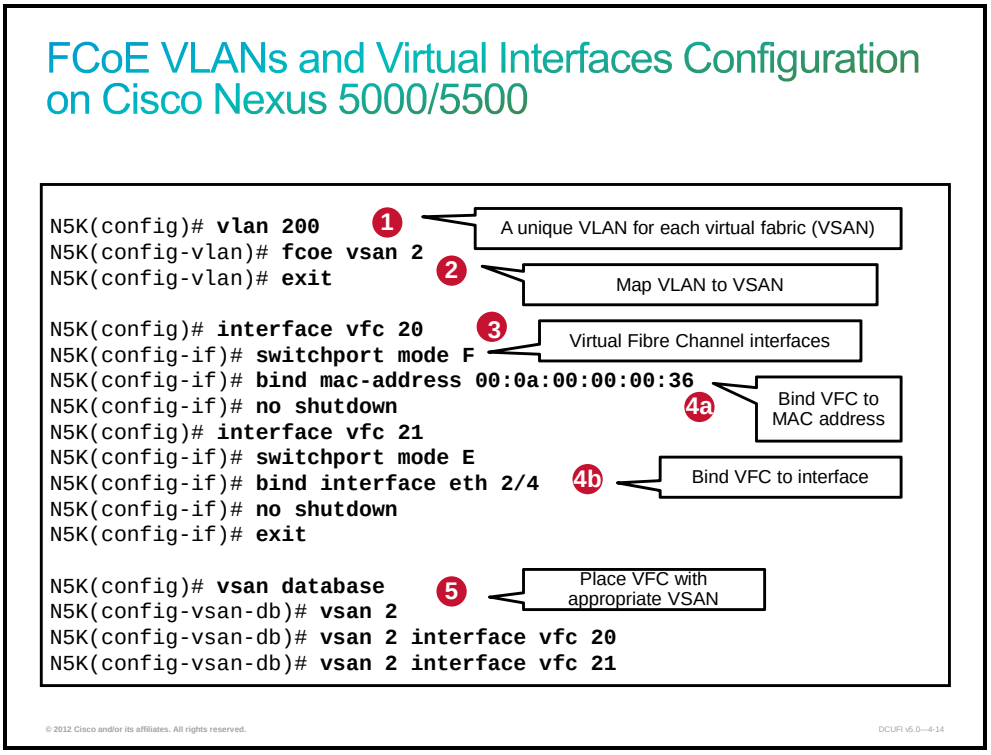
1. Configure a dedicated VLAN for each Virtual Fabric (VSAN)
2. Map the VLAN to the specified VSAN
3. Configure a virtual Fibre Channel interface
   - Switchport modes: virtual F (VF), virtual E (VE), N-Port Virtualizer (NPV)
   - Default mode: virtual fabric (VF)
4. Bind the virtual Fibre Channel interface to physical interface
   a. To interface
   b. To MAC address – for use of FIP devices
5. Associate the virtual Fibre Channel interface to an appropriate VSAN

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

Follow this procedure to configure FCoE VLANs and vFC interfaces on Cisco Nexus 5000 and 5500 Platform switches:

1. Configure a dedicated, unique VLAN for each virtual fabric (VSAN)
2. Map the VLAN to the specified VSAN. This procedure is done in VLAN configuration mode.
3. Configure a vFC interface. A vFC interface can be configured as VF Port, VE Port, or N-Port Virtualization (NPV) port. The default mode is virtual fabric (VF) mode.
4. Bind the vFC interface to a physical interface. You can bind the vFC interface to a MAC address, or to the switch interface.
5. Associate the vFC interface to the appropriate VSAN

https://t.me/learningnets



The configuration shows the necessary commands to implement the scenario that is based on a Cisco Nexus 5000 and 5500 Platform switch.

The FCoE VLAN 200 is mapped to VSAN 2 and provides connectivity between the vFC interfaces 20 and 21 that are associated with the VSAN 2. The vFC interfaces 20 and 21, configured as VF and VE Ports, are bound to a host MAC address, and to a switch interface, respectively.

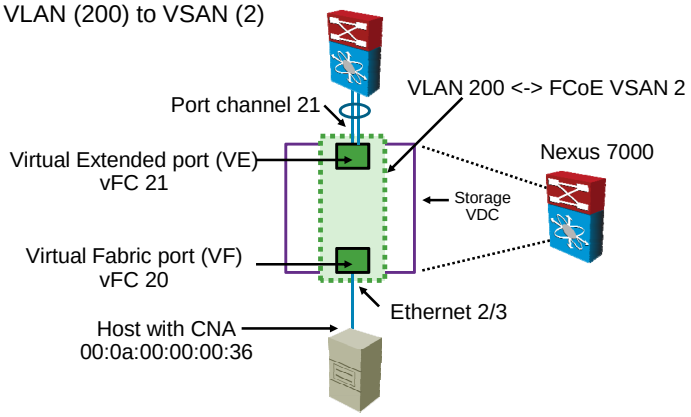
Note	As a best practice, it is good to have the vFC defined to a particular VSAN (that is, VSAN port membership).

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Cisco Nexus 7000 FCoE VLAN and Virtual Interfaces Configuration Scenario

- Virtual Fibre Channel interfaces:
  - vFC 20 acts as virtual fabric port for the host
  - vFC 21 acts as virtual extended port to a FCoE switch
- Communication between ports:
  - Binding of VLAN (200) to VSAN (2)



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

The configuration scenario that is based on a Cisco Nexus 7000 Series switch uses the same FCoE VLANs and vFC interfaces. Most settings are, however, configured in the storage VDC. The VE interface connecting to the other FCoE switch is a port channel interface.

Cisco Nexus 7000 FCoE VLAN and Virtual Interfaces Configuration Procedure

1. Make sure that the configuration is in storage VDC
2. Configure a dedicated VLAN for each Virtual Fabric (VSAN)
3. Map the VLAN to the specified VSAN
4. Configure a virtual Fibre Channel interface
5. Bind the virtual Fibre Channel interface to physical interface
6. Configure virtual Fibre Channel port channel (optional)
7. Place the virtual Fibre Channel interface to VSAN
8. Place the virtual Fibre Channel port channel to VSAN (optional)
9. Configure VE loopback (optional)
   - By default VFID check verifies that the VSAN configuration is correct on both ends of a VE link
   - VE loopback turns off the VFID check for VE ports

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

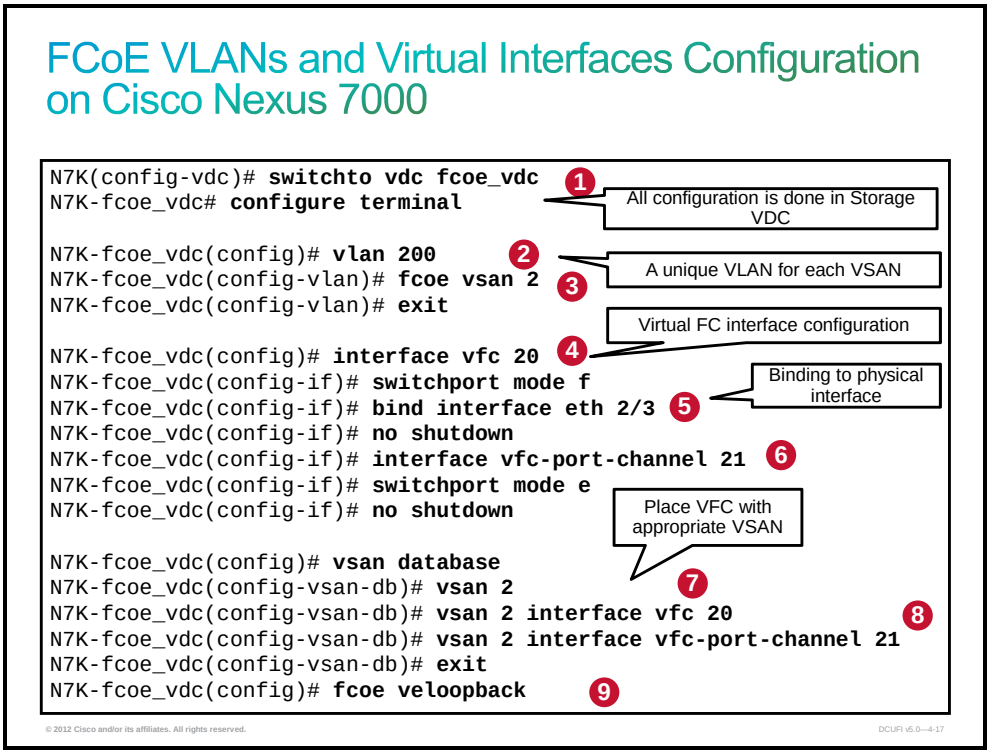
The Cisco Nexus 7000 configuration procedure is very similar to the one that is based on the Cisco Nexus 5000 Platform switch configuration procedure. There are a couple of differences.

You configure most FCoE settings in the storage VDC.

You can configure vFC port channel interfaces and assign them to a VSAN.

https://t.me/learningnets

You can configure the VE loopback feature. By default the virtual fabric ID (VFID) check is used to make sure that the VSAN configuration is identical on both ends of a VE link. If you enable the VE loopback feature, the VFID check is turned off for VE ports.



The configuration is performed in the storage VDC. The vFC port channel interface that is configured in Step 5 is automatically bound to the port channel with the same number. The vFC port channel interface is associated with the VSAN 2 in Step 8. In Step 9, you disable the VFID check and thus allow different VSAN configurations on both ends of a VE link. The VE link is the vFC port channel interface, which is configured in Step 6.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Verifying VLANs, VSANs, and Interfaces

```
switch(config)# show vlan fcoe
VLAN      VSAN      Status
--------  --------  --------
200       2         Operational


switch# show vsan membership

vsan 2 interfaces:
        vfc20 vfc21


switch# show interface brief
<output about physical omitted>
Interface  Vsan   Admin  Admin   Status          SFP    Oper  Oper   Port
                  Mode   Trunk                          Mode  Speed  Channel
                         Mode                           (Gbps)
--------------------------------------------------------------------------------
vfc 20        2      F      --      up              --    F      --      --
vfc 21        2      E      --      up              --    E      --      --
...
```

VLANs enabled for FCoE and their mapping to VSANs

Interface members in a VSAN

Brief interface output listsVFC interfaces

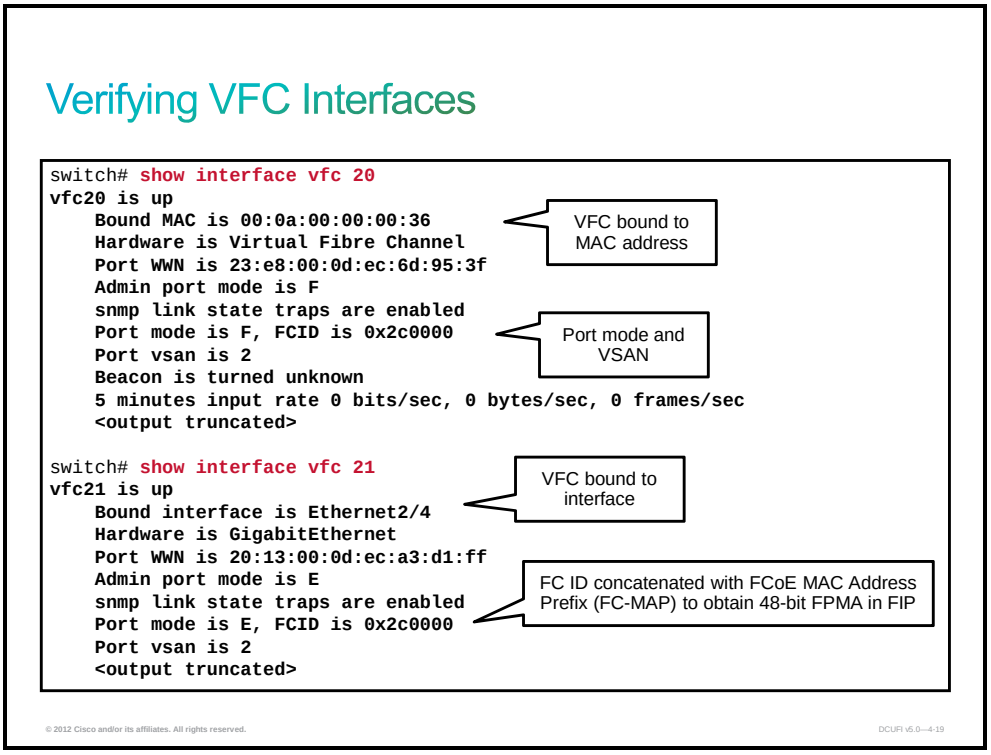
© 2012 Cisco and/or its affiliates. All rights reserved.

The **show vlan fcoe** command is used to validate the mapping between FCoE VLANs and VSANs , in this example VLAN 200 and the FCoE VSAN 2.

The **show vsan membership** command is used to determine which interfaces are assigned to various VSANs. The vFC interfaces vcf20 and vfc21 are members of VSAN 2.

You can use the **show interface brief** command to display all interfaces that are available in the system. The vFC interfaces are listed in the output, along with the associated VSAN ID, switchport mode, and status.

https://t.me/learningnets



You can use the **show interface** command to display more detailed information about the vFC interfaces.

The first example shows the parameters of the vFC interface bound to a host MAC address and operating in the F mode. The second example shows the parameters of the vFC interface bound to a switch interface and operating in the E mode.

Both interfaces have a unique port world wide name (pWWN) shown as Port WWN in the figure. The Fibre Channel ID (FCID in the figure) is 0x2c0000. The FCID is concatenated with the FCoE MAC address prefix (FC-MAP) to obtain a 48-bit fabric provided MAC address (FPMA). The FC-MAP and FCID are communicated to the host CNA using FIP.

https://t.me/learningnets

Verifying FCoE Connectivity

Connectivity verification through the VSAN

```
switch# fcping fcid 0x4a00dc vsan 2
28 bytes from  0x5900dc  time = 906 usec
28 bytes from  0x5900dc  time = 899 usec
28 bytes from  0x5900dc  time = 899 usec
28 bytes from  0x5900dc  time = 899 usec
28 bytes from  0x5900dc  time = 899 usec

5 frames sent, 5 frames received, 0 timeouts
Round-trip min/avg/max = 899/900/906 usec

switch# fcping fcid 0x4a00e0 vsan 2
28 bytes from  0x5900e0  time = 907 usec
28 bytes from  0x5900e0  time = 899 usec
28 bytes from  0x5900e0  time = 898 usec
28 bytes from  0x5900e0  time = 896 usec
28 bytes from  0x5900e0  time = 924 usec

5 frames sent, 5 frames received, 0 timeouts
Round-trip min/avg/max = 896/904/924 usec
```

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

The **fcping** command that is available on Cisco MDS switches is analogous to the Internet Control Message Protocol (ICMP) **ping** command. It is used to validate the connectivity of Fibre Channel interfaces.

The command output shows five separate frames that have been sent from two different Fibre Channel source addresses and the round-trip time.

https://t.me/learningnets

General FCoE Guidelines

- FCoE interfaces:
  - Can be Ethernet or port channel interfaces
  - Can be connected to FIP snooping bridges
  - Must be trunk ports
  - Must be configured as PortFast
- FCoE VLAN mapped to VSAN:
  - Must be in the allowed VLAN list
  - Cannot be the native VLAN of the trunk port
  - Should carry FCoE traffic only
  - Should not be default VLAN (VLAN1)
- FCoE is not supported on private VLANs

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

You should follow several guidelines when implementing FCoE on any Cisco Nexus switches.

The FCoE interfaces can be either Ethernet or port channel interfaces. The FCoE ports can be connected to FIP snooping bridges. They must be Layer 2 trunk ports that are configured with the PortFast feature.

You should make sure that the FCoE VLANs that are mapped to the appropriate VSANs are allowed on the trunk ports. This result is achieved either by putting the respective VLAN IDs in the allowed VLAN list or not configuring any VLAN restrictions.

The FCoE VLAN cannot be the native VLAN of the trunk port. It should carry only FCoE traffic and should not be the default VLAN (VLAN1).

FCoE is not supported on private VLANs.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Cisco Nexus 7000 FCoE Guidelines

- Cisco Nexus 7000 supports Gen-2 or newer CNAs only
- QoS policy must be the same on all Cisco FCoE switches in the network
- Storage VDC:
  - Should provide only storage-related features
  - FCoE feature set can be enabled in only one VDC
  - FCoE VLANs configured in the FCoE allocated VLAN range
  - Uses resources from an F-series module
  - Does not support rollback
- Shared interfaces:
  - Can be shared with only one other VDC
  - Do not support certain features, such as SPAN, private VLANs, port channels, access mode, mac-packet-classify

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

You should consider several additional guidelines when implementing FCoE on a Cisco Nexus 7000 Series switch. Cisco Nexus 7000 switches support Generation-2 or newer CNAs only. The QoS policy must be the same on all Cisco FCoE switches in the network.

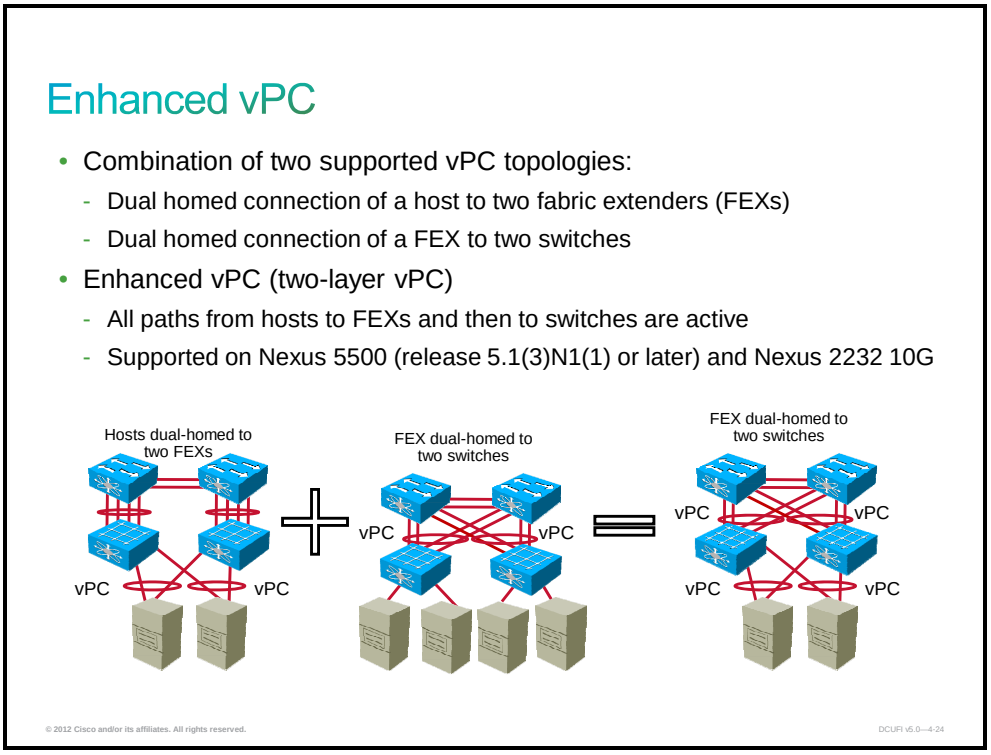
The storage VDC should provide only storage-related features. You can enable the FCoE feature set in only one VDC. When you configure FCoE VLANs in the storage VDC, you can use only the VLANs from the VLAN range that is allocated to the storage VDC. The storage VDC can use resources from a Cisco Nexus 7000 F-series Ethernet module. The storage VDC does not support rollback.

When you share interfaces between the storage VDC and another VDC, the shared interfaces do not support certain features, such as Switched Port Analyzer (SPAN), private VLANs, port channels, access mode, and the mac-packet-classify feature.

https://t.me/learningnets

FCoE with Enhanced vPC and Cisco Adapter FEX

This topic explains how to implement FCoE with enhanced vPC and Cisco Adapter FEX.



The Enhanced vPC feature, introduced in a previous lesson, combines two dual-homing scenarios is one solution:

- Dual-homed connection of a host to two fabric extenders (FEXs)
- Dual-homed connection of a FEX to two switches

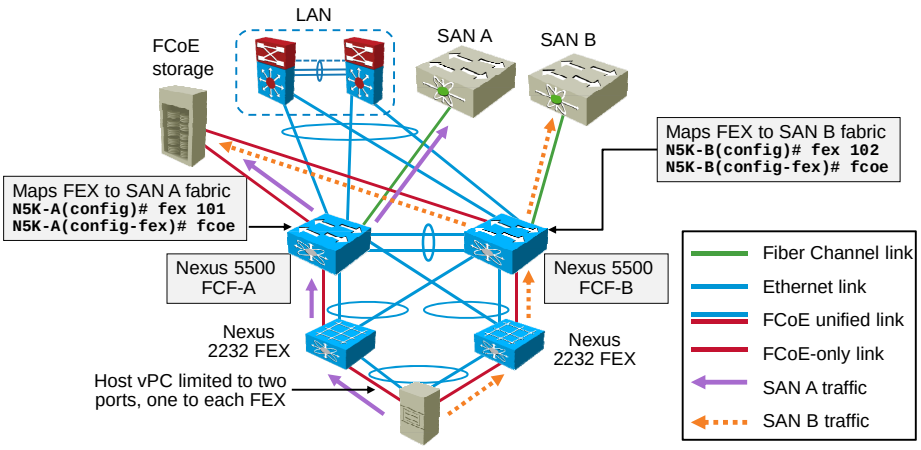
The Enhanced vPC solution can be used to provide FCoE connectivity. Enhanced vPC with FCoE is supported on the Cisco Nexus 5500 Platform running Cisco Nexus Operating System (NX-OS) Software Release 5.1(3)N1(1) or a later release and the Cisco Nexus 2232 Fabric Extender that is equipped with 10 Gigabit Ethernet interfaces.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

FCoE over Enhanced vPC

- Ethernet traffic dual-homed
- FCoE traffic single homed
- Host facing VFC is bound to physical interface and not port channel



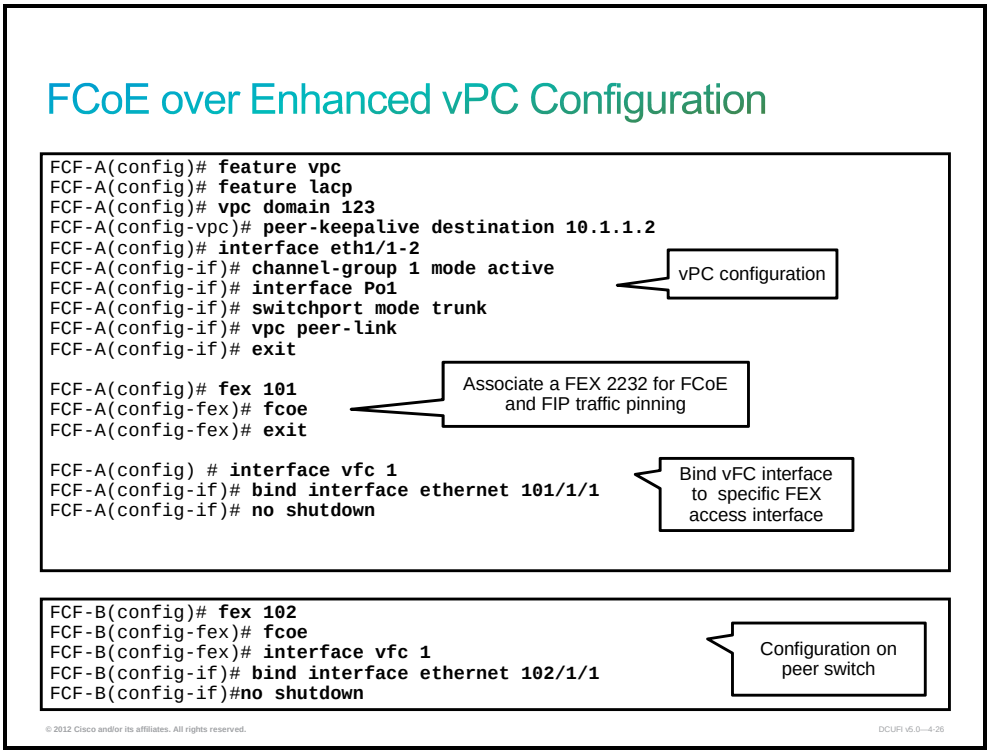
© 2012 Cisco and/or its affiliates. All rights reserved.

DCUFI v5.0—4-25

In the FCoE over Enhanced vPC solution, only Ethernet traffic from the host is dual-homed to the FEXs and then further dual-homed to the parent switches. The Ethernet traffic is consequently load balanced over the parallel paths via the FEXs and parent switches.

The FCoE traffic is single-homed. The vFC interfaces on the parent switches are bound to the physical access interfaces, and not the port channels. As a result, the FCoE traffic toward a given SAN flows to a single FEX and then a single parent switch. In this example, the FCoE traffic that is bound for SAN A flows over the left FEX and FCF-A. The FCoE traffic that is bound for SAN B flows over the right FEX and FCF-B.

https://t.me/learningnets



The configuration example illustrates the relevant settings on both Cisco Nexus 5500 Platform parent switches.

The first configuration part explains how to enable the required features, and configure the vPC parameters, such as the vPC domain name, peer keepalives, and the peer link. The respective configuration on the peer device (FCF-B) is not shown.

The second configuration section on FCF-A shows how to associate a FEX and bind the vFC interface. The equivalent configuration on FCF-B is shown here. You use these commands:

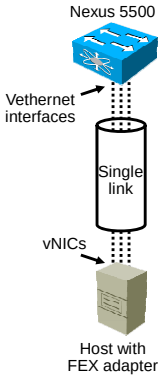
- The **fcoe** command is used to associate a Cisco Nexus 2232 Fabric Extender for pinning FCoE and FIP traffic. Optionally you could use the VSAN ID parameters to associate the FEX with a specific VSAN. The VSAN ID range is from 1 to 4094.
- The **bind interface** command is used to bind the vFC interface to a host access interface. When you bind an interface to a virtual Fibre Channel interface to enable FCoE traffic, you must use slot number 1. The port number can be from 1 to 32.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Cisco Adapter Fabric Extender FCoE

- Single physical link split into multiple virtual channels
- Channel
  - Identified by a unique channel number
  - Channel scope limited to the physical link
  - Connects a server vNIC with a switch Vethernet interface
  - Uses tagging with VNTag identifiers
- FCoE switch side:
  - Nexus 5500
  - Nexus 2232 connected to Nexus 5500
- FCoE server-side:
  - Cisco UCS P81E Virtual Interface Card for UCS C-Series
  - Third-party adapters that support the VNTag technology
  - For example Broadcom BCM57712 Convergence NIC



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

The Cisco Adapter FEX feature combines the advantages of the FEX link architecture with server I/O virtualization to create multiple virtual interfaces over a single Ethernet interface. Cisco Adapter FEX divides a single physical link into multiple channels. Each channel is identified by a unique channel number, and its scope is limited to the physical link.

The physical link connects a port on a server network adapter with an Ethernet port on the switch. This connection allows the channel to connect a virtual network interface card (vNIC) on the server with a virtual Ethernet (vEthernet) interface on the switch.

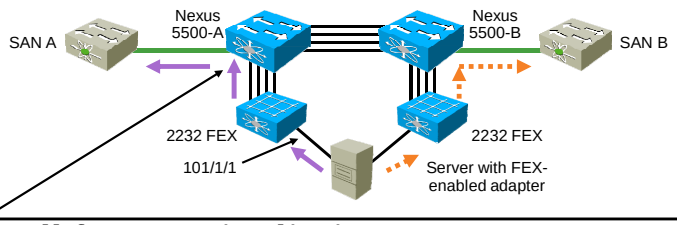
Packets on each channel are tagged with a virtual network tag (VNTag) that has a specific source virtual interface identifier (VIF). The VIF allows the receiver to identify the channel that the source used to transmit the packet.

The FCoE Cisco Adapter Fabric Extender solution is available with the Cisco Nexus 5500 Platform switches and with the Cisco Nexus 2232 Fabric Extenders that are connected to a Cisco Nexus 5500 Platform parent switch.

The FCoE Cisco Adapter FEX implementation is designed to work with server network adapters, such as the Cisco UCS P81E Virtual Interface Card for the Cisco UCS C-Series Rack-Mount Server (UCS P81E VIC) or third-party adapters that support the VNTag technology, such as the Broadcom BCM57712 Convergence Network Interface Card.

https://t.me/learningnets

Cisco Adapter Fabric Extender FCoE Configuration




```
N5500-A(config)#install feature-set virtualization
N5500-A(config)#feature-set virtualization

N5500-A(config)#fex 101
N5500-A(config-fex)#fcoe

N5500-A(config-fex)#interface ethernet 101/1/1
N5500-A(config-if)#switchport mode vntag

N5500-A(config-if)#interface vethernet 2
N5500-A(config-if)#bind interface eth 101/1/1 channel 1
N5500-A(config-if)#switchport mode trunk
N5500-A(config-if)#switchport trunk allowed vlan 33

N5500-A(config-if)#interface vfc 4
N5500-A(config-if)#bind interface vethernet 2
N5500-A(config-if)#no shutdown
```

Associate a FEX 2232 for FCoE and FIP traffic pinning

Access interface enabled for adapter FEX

Vethernet interface configuration

vFC interface bound to Vethernet interface

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

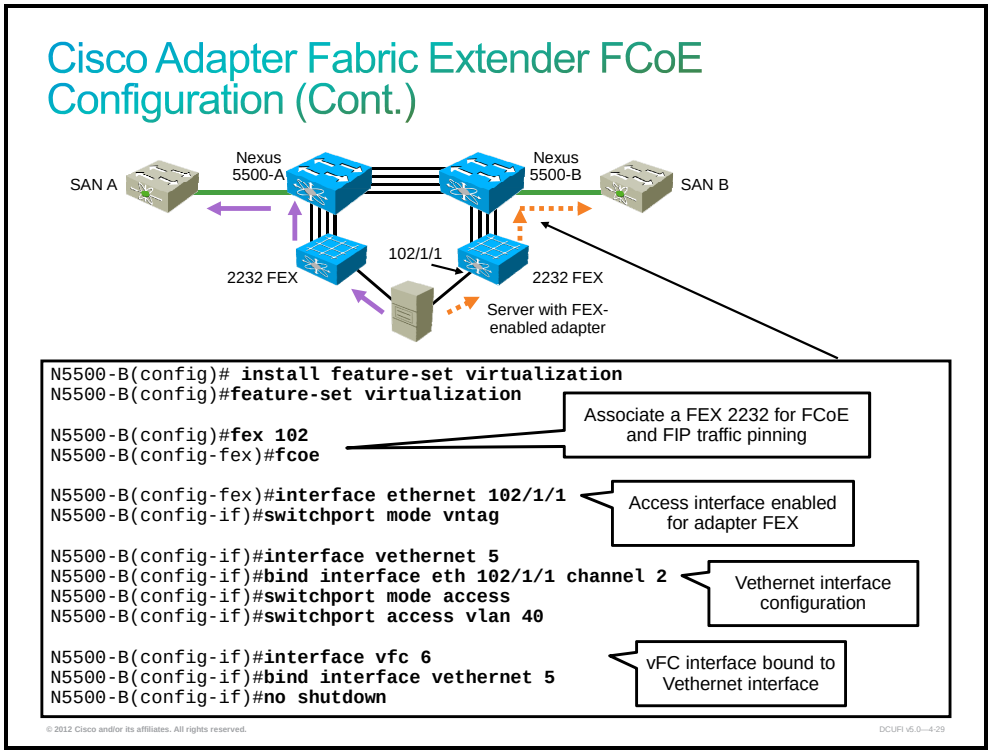
The FCoE Cisco Adapter FEX configuration is performed on the Cisco Nexus 5500 parent switch. The example illustrates the configuration on the switch Cisco Nexus 5500-A, that provides connectivity to SAN A.

First, you must enable the features that are needed for Cisco Adapter FEX operations. You have to associate the Cisco Nexus 2232 Fabric Extender for FCoE and FIP traffic pinning.

The vEthernet 2 interface is bound to the host access port 101/1/1 and Cisco Adapter Fabric Extender channel 1. The Vethernet interface is configured for trunking and carries VLAN 33, which is the FCoE VLAN in this example. Finally the vFC interface is bound to the Vethernet interface.

This configuration example does not use the auto-create feature of Cisco Adapter FEX. The auto-create feature would allow you to create the vEthernet interfaces automatically and populate them with the setting that is inherited from port profiles.

https://t.me/learningnets



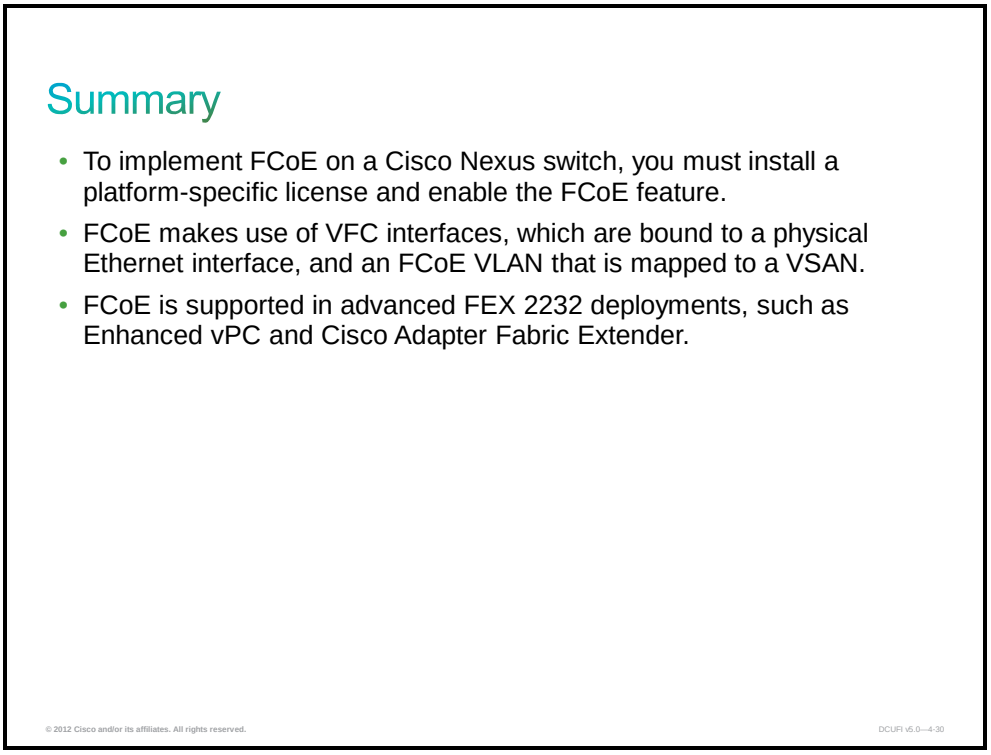
This example illustrates the configuration on the second parent switch that provides FCoE connectivity to SAN B.

The configuration uses the host access interface 102/1/1. The vEthernet interface 5 runs over the host access interface 102/1/1 and uses the channel ID 2. The vFC interface 6 is bound to the vEthernet interface 5.

https://t.me/learningnets

Summary

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



References

For additional information, refer to these resources:

- To learn more about configuring FCoE on the Cisco Nexus 5000 and 5500 Platform switches, refer to *Cisco Nexus 5000 Series NX-OS Fibre Channel over Ethernet Configuration Guide, Release 5.1(3)N1(1)* at: http://www.cisco.com/en/US/docs/switches/datacenter/nexus5000/sw/fcoe/513_n1_1/b_Cisco_n5k_fcoe_config_gd_re_513_n1_1.html
- To learn more about configuring FCoE on the Cisco Nexus 7000 switches, refer to *Cisco NX-OS FCoE Configuration Guide for Nexus 7000 and MDS 9500* at: http://www.cisco.com/en/US/docs/switches/datacenter/sw/nx-os/fcoe/configuration/guide/b_Cisco_NX-OS_FCoE_Configuration_Guide_chapter_011.html

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Lesson 5

Configuring SAN Switching Features

Overview

The Cisco Nexus 7000 Series Switches, Cisco Nexus 5000 and 5500 Platform switches, and the Cisco MDS switches support a plentiful set of Fibre Channel functions. In this lesson, you will learn how to configure the relevant SAN switching features.

Objectives

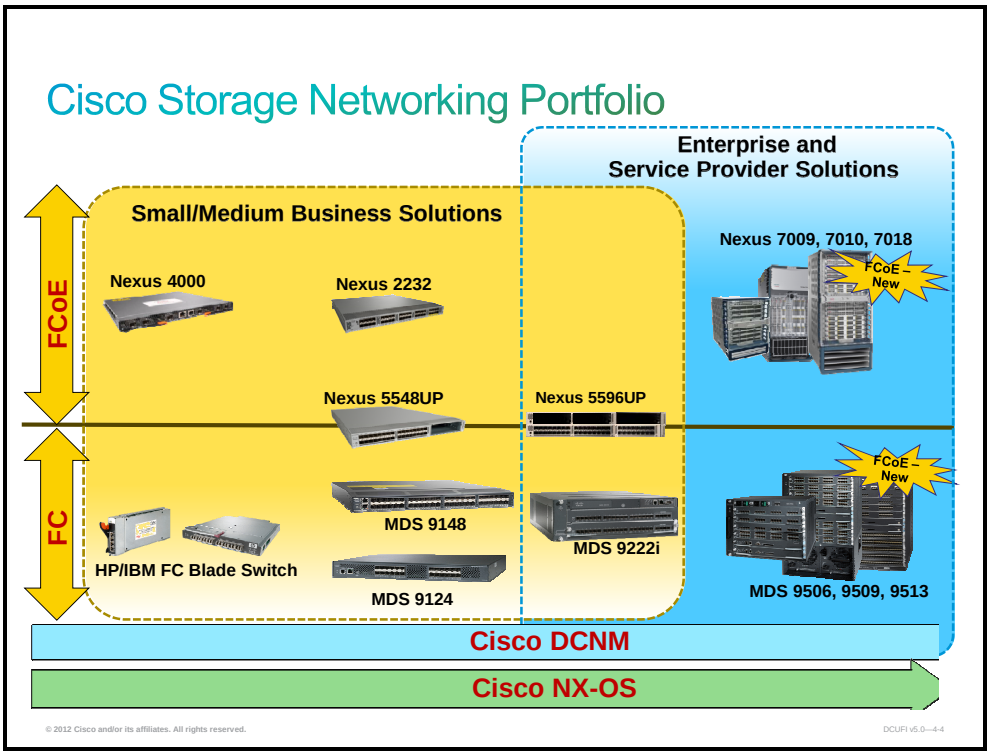
Upon completing this lesson, you will be able to configure the relevant SAN switching features on the Cisco Nexus and Cisco MDS Families of switches to support the design requirements. This ability includes being able to meet these objectives:

- Describe FCoE products in the Cisco MDS Family of switches
- Configure Fibre Channel interfaces
- Configure domain parameters on a Cisco Nexus switch
- Configure and manage VSANs
- Configure interfaces to support VSAN trunking
- Configure the Cisco Nexus and MDS Families of switches to support SAN port channels
- Manage the FLOGI and FCNS databases on the Cisco Nexus 5000 Series Switches
- Manage Cisco Fabric Services on a Cisco MDS switch

https://t.me/learningnets

FCoE on Cisco MDS

This topic describes the Fibre Channel over Ethernet (FCoE) solution in the Cisco MDS Family of switches



The Cisco storage network portfolio covers both small and medium businesses as well as enterprise and service provider requirements.

The already strong presence Cisco has with the Fibre Channel product portfolio is now enhanced to include FCoE as well.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Cisco MDS 9000 8-Port FCoE Module

- 8-Port Line Rate FCoE Module
- Multihop FCoE
- Support 10GE SFP+ SR/LR
- Bridge Unified Fabric to FC SANs
- Enabling storage services to Unified Fabric

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

The Cisco MDS 9000 10-Gbps 8-Port FCoE Module, part number DS-X9708-K9, enables multihop FCoE on Cisco MDS switches thus bridging traditional Fibre Channel SAN with unified fabric.

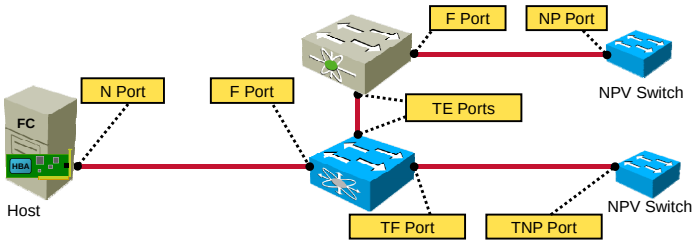
https://t.me/learningnets

Fibre Channel Interfaces

This topic describes how to configure Fibre Channel interfaces.

Fibre Channel Interfaces

Port	Description
E Port	Expansion port
F Port	Fabric port is connected to a peripheral device operating as a node port (N port)
NP Port	NP ports operate like N ports that function as proxies for multiple physical N ports
TE Port	Trunking expansion port creates an extended ISL (EISL) between two switches
TF Port	Trunking fabric port expand the functionality of F ports to support VSAN trunking
TNP Port	Trunking NP port may be connected to a trunked F port (TF port).
SD Port	SPAN destination port functions as a switched port analyzer (SPAN)



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

A Fibre Channel interface supports the following modes: E, F, NP, TE, TF, TNP, SD, and Auto.

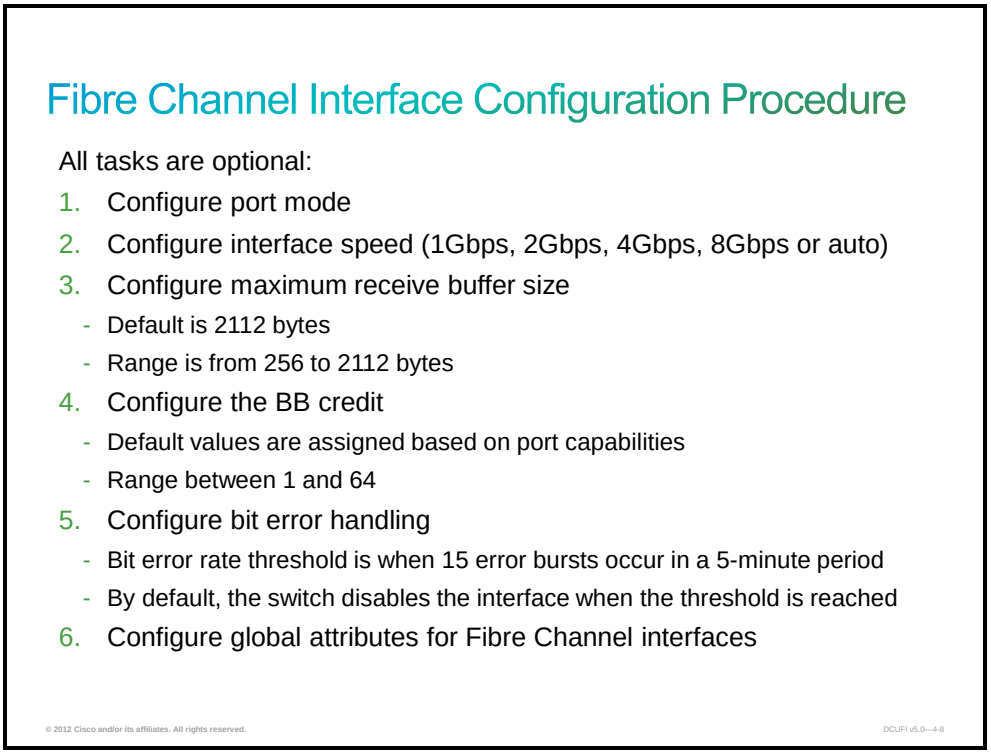
- **E Port**: This is an expansion port (E Port) mode where an interface functions as a fabric expansion port. This port is connected to another E Port to create an interswitch link (ISL) between two switches. E Ports carry frames between switches for configuration and fabric management. They serve as a conduit between switches for frames that are destined to remote node ports (N Ports) and node loop ports (NL Ports).
- **F Port**: This is a fabric port (F Port) mode where an interface functions as a fabric port. This port is connected to a peripheral device (host or disk) that is operating as an N Port. An F Port can be attached to one N Port only.
- **NP Port**: This is a port on a device that is in N-Port Virtualization (NPV) mode and is connected to the core switch via an F Port. Proxy N ports (NP Ports) function like N Ports except that in addition to providing N Port operations, they also function as proxies for multiple N Ports.
- **TE Port**: This is a trunking E Port (TE Port) mode where an interface functions as a trunking expansion port. This port is connected to another TE Port to create an Enhanced Inter-Switch Link (EISL) between two switches. TE Ports are specific to Cisco switches and expand the functionality of E Ports to support the following:
  - Virtual storage area network (VSAN) trunking
  - Transport quality of service (QoS) parameters
  - Fibre Channel Trace Route (fctrace) feature

  When an interface is in TE Port mode, all frames that are transmitted are in the EISL frame format, which contains VSAN information. Interconnected switches use the VSAN ID to multiplex traffic from one or more VSANs across the same physical link.

https://t.me/learningnets

- **TF Port:** This is a trunking F Port (TF Port) mode where an interface functions as a trunking E Port. This interface is connected to another trunked N Port (TN Port) or trunked NP Port (TNP Port) to create a link between a core switch and an NPV switch or a host bus adapter (HBA) to carry tagged frames. TF Ports are specific to Cisco switches and expand the functionality of F Ports to support VSAN trunking. In TF Port mode all frames are transmitted in the EISL frame formation, which contains VSAN information.
- **TNP Port:** This is a trunking NP Port (TNP Port) mode where an interface functions as a trunking E Port. This interface will be connected to a trunked F Port (TF Port) to create a link to a core N-Port ID Virtualization (NPIV) switch from an NPV switch to carry tagged frames.
- **SD Port:** This is a SPAN destination port (SD Port) mode where an interface functions as a Switched Port Analyzer (SPAN). An SD Port monitors network traffic that passes through a Fibre Channel interface. It uses a standard Fibre Channel analyzer (or similar switch probe) that is attached to the SD Port. SD Ports cannot receive frames and only transmit a copy of the source traffic. This feature is nonintrusive and does not affect switching of network traffic for any SPAN source port.
- **Auto Mode:** Interfaces that are configured in auto mode will automatically detect the mode to which they should be configured during interface initialization.

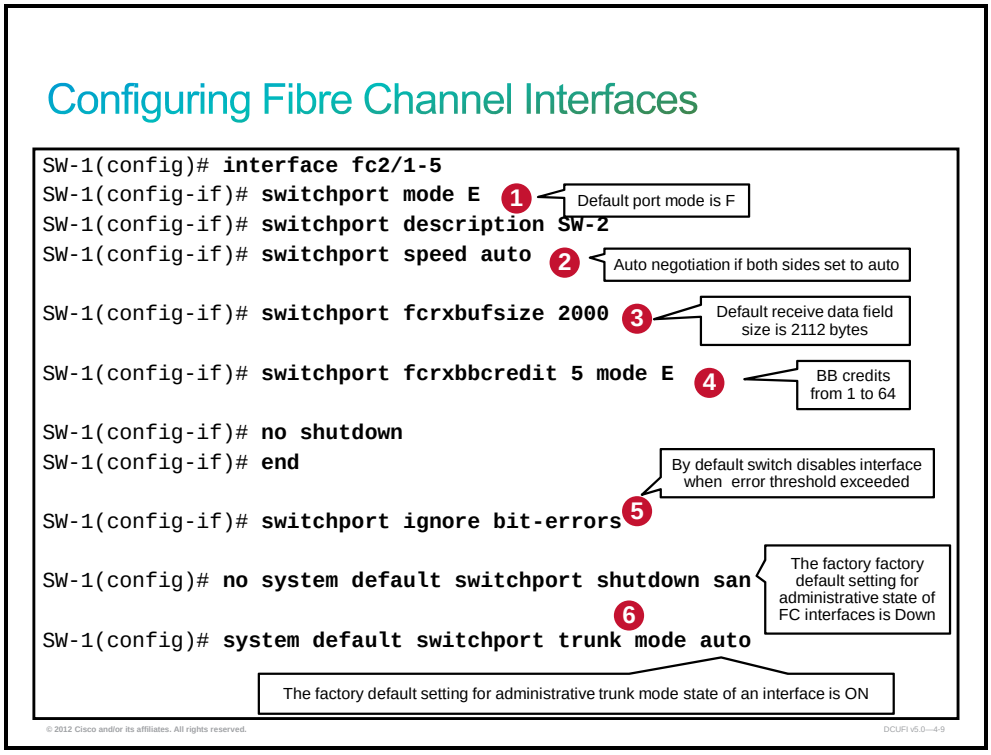
https://t.me/learningnets



When you configure physical Fibre Channel interfaces, you can tune their default attributes. You can perform the tuning using this procedure:

1. Configure the port mode.
2. Configure the interface speed. The available settings are 1, 2, 4, or 8 Gb/s or auto. Autosensing speed (setting auto) is enabled on all 4-Gb/s interfaces by default. This configuration enables the interfaces to operate at speeds of 1, 2, or 4 Gb/s on the 4-Gb/s ports. When autosensing is enabled for an interface that is operating in dedicated rate mode, 4 Gb/s of bandwidth is reserved, even if the port negotiates at an operating speed of 1 Gb/s or 2 Gb/s.
3. Configure the maximum receive buffer size. The default value is 2112 bytes. The range is from 256 to 2112 bytes. You can configure this parameter for native Fibre Channel interfaces but not for virtual Fibre Channel (vFC) interfaces. If the default data field size is 2112 bytes, the frame length will be 2148 bytes.
4. Configure the BB credit. The default BB credit values are assigned based on port capabilities. The range is between 1 and 64.
5. Configure bit error handling. The bit error rate threshold is used by the switch to detect an increased error rate before performance degradation seriously affects traffic. The bit errors can occur because of faulty cabling, faulty or bad Gigabit Interface Converter (GBIC) or small form-factor pluggable (SFP), speed configuration mismatch, and similar issues. The bit error rate threshold is reached when 15 error bursts occur in a 5-minute period. By default, the switch disables the interface when the threshold is reached.
6. Configure global attributes for Fibre Channel interfaces. These attributes are applied globally to all future switch port configurations, even if you do not individually specify them at that time.

© 2012 Cisco Systems, Inc.
https://t.me/learningnets

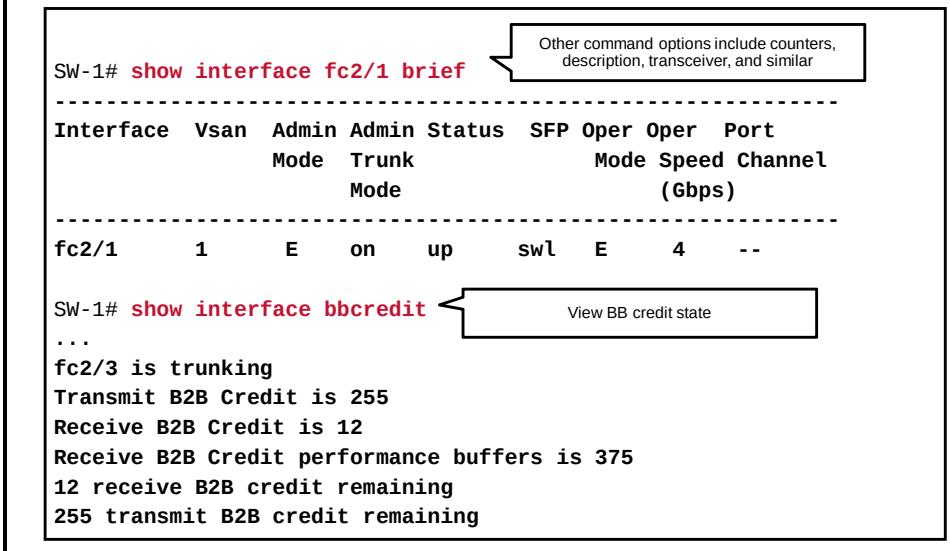


This sample configuration illustrates how to set the various attributes of a Fibre Channel interface. In this scenario you change the default port mode (fabric) to expansion, configure an interface description, set the speed to autosensing, modify the receive buffer size and BB credit value, and activate the interface. In addition, you configure the Fibre Channel interfaces to remain in an operational state even when the bit error threshold is exceeded.

Finally, you tune the systemwide parameters. You change the factory default setting for the administrative state of Fibre Channel interfaces from Down to Up using the **no system default switchport shutdown san** command. This command is applicable only to interfaces for which no user configuration exists for the administrative state. The **system default switchport trunk mode auto** command is used to configure the default setting for the administrative trunk mode state of an interface as Auto. The default setting is trunk mode on.

https://t.me/learningnets

Verifying Fibre Channel Interfaces



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

You can verify the status of physical Fibre Channel interfaces using the **show interface** command. The command allows appending various options, such as interface name, brief, counters, description, transceiver, and others to narrow down the output to the required data.

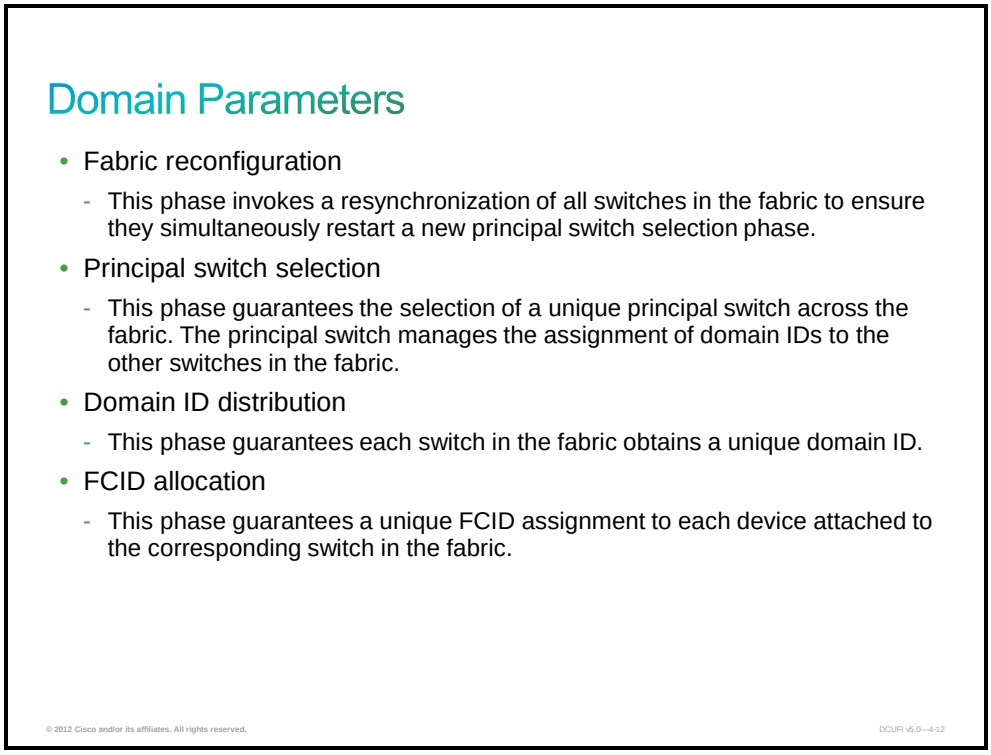
The **show interface bbcredit** command is used to view the configured and remaining transmit and receive B2B credits.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Domain Parameters

This topic describes how to configure domain parameters on a Cisco Nexus switch.



The Fibre Channel domain (fcdomain) feature performs principal switch selection, domain ID distribution, Fibre Channel ID (FCID) allocation, and fabric reconfiguration functions. Domains are configured on a per-VSAN basis. If a domain ID is not configured, then the local switch uses a random ID.

To successfully configure domain parameters and prevent fabric segmentation, it is necessary to understand the anticipated behavior of the fcdomain phases. The four phases that are associated with the fcdomain process are as follows:

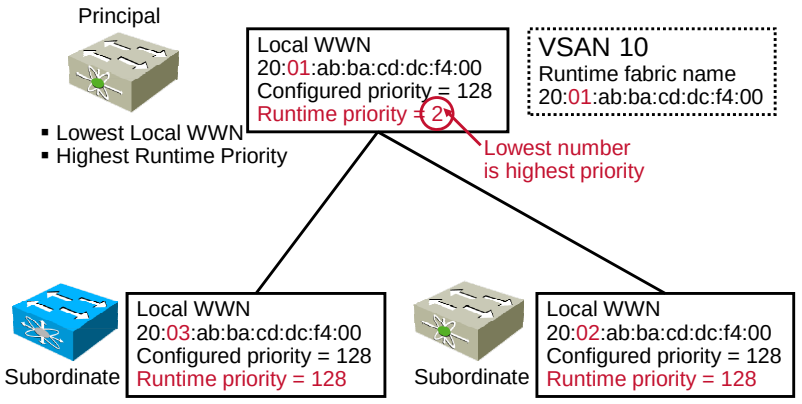
- **Principle switch selection:** This phase guarantees the selection of a unique principle switch across the fabric.
- **Domain ID distribution:** This phase guarantees that each switch in the fabric obtains a unique domain ID.
- **FCID allocation:** This phase guarantees a unique FCID assignment to 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.

The domain IDs in a single fabric (VSAN) must be unique across all switches.

https://t.me/learningnets

Principal Switch

- Selected using a procedure that guarantees a unique principal switch
- Responsible for ensuring domains IDs are unique in a fabric
- Normally located in the core SAN fabric.



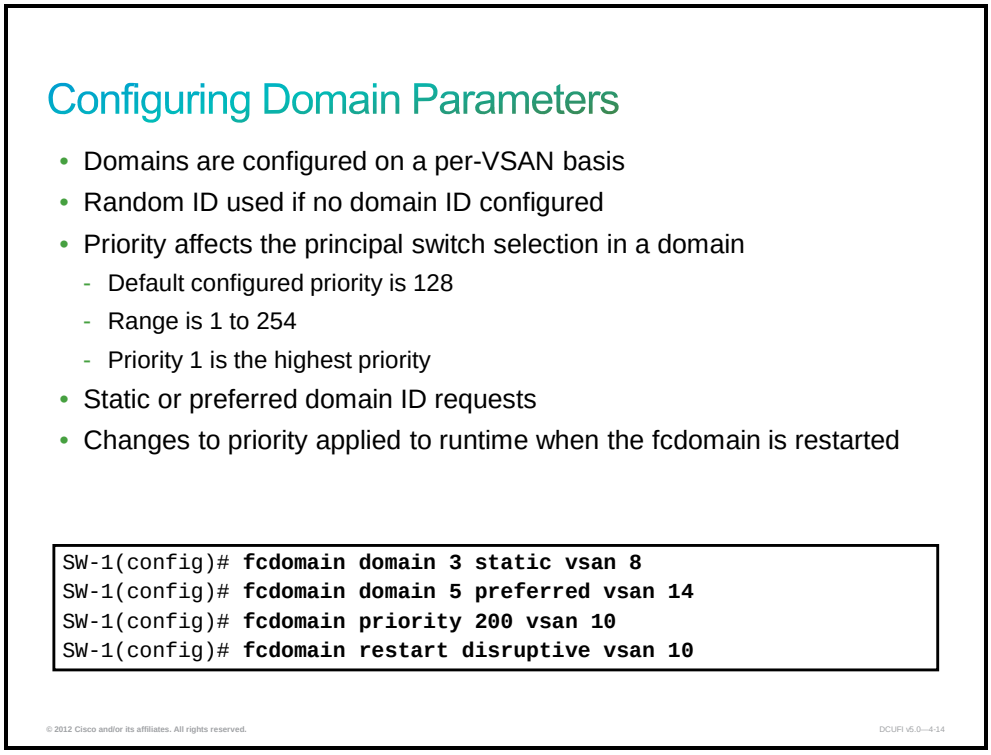
© 2012 Cisco and/or its affiliates. All rights reserved.

The principal switch selection phase guarantees a unique principal switch across the fabric. The principal switch is responsible for ensuring that domain IDs are unique in a single fabric.

The principal switch would normally be located in the core SAN fabric.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets



By default, the configured switch priority is 128. The valid range to set the priority is between 1 and 254. Priority 1 is the highest priority.

Any new switch cannot become the principal switch when it joins a stable fabric. During the principal switch selection phase, the switch with the highest priority becomes the principal switch. If two switches have the same configured priority, the switch with the lower world wide name (WWN) becomes the principal switch.

Modifying the priority of a switch requires the fcdomain to be restarted.

The figure shows how to configure the priority on the Cisco Nexus switches. If you configure a Nexus switch as an edge switch, it should not become the principal switch in a fabric.

Cisco Nexus switches can be configured to make either static or preferred domain ID requests.

- **Static requests:** These requests specify the only domain ID that the requester (subordinate switch) accepts from the principal switch. If the principal switch is unable to satisfy the request, it passes a different (available) domain ID back to the requester. The requester discards this domain ID, assigns itself the statically-configured domain ID, and isolates itself from the attached fabric.
- **Preferred requests:** These requests specify the domain ID that the requestor prefers, but the requestor accepts any domain ID that is provided by the principal switch.

You must issue the **fcdomain restart disruptive vsan** *vsan-id* command to apply the configuration domain changes to the runtime environment.

https://t.me/learningnets

Verifying Domain Parameters

```
SW-1# show fcdomain
VSAN 1
The local switch is a Subordinated Switch.

Local switch run time information:
        State: Stable
        Local switch WWN:    20:01:00:05:30:00:13:9f
        Running fabric name: 20:01:00:05:30:00:13:9e
        Running priority: 128
        Current domain ID: 0x4a(74)

Local switch configuration information:
        State: Enabled
        FCID persistence: Enabled
        Auto-reconfiguration: Disabled
        Contiguous-allocation: Disabled
        Configured fabric name: 20:01:00:05:30:00:28:df
        Configured priority: 128
        Configured domain ID: 0x00(0) (preferred)

Principal switch run time information:
        Running priority: 2
```

Local runtime WWN, priority, and domain ID

Priority of the principal switch is higher (numerically lower) than local value

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

You can use the **show fcdomain** command with no arguments to display the status of the fcdomain feature on all VSANs. The VSANs should be active or an error is generated. The output provides information on the local WWN, priority, and domain ID, as well as the principal switch priority.

 © 2012 Cisco Systems, Inc.

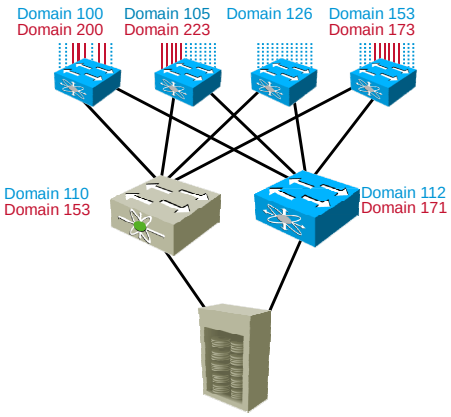
https://t.me/learningnets

VSAN Management

This topic describes how to configure and manage VSANs.

VSANs

- Each VSAN has its own principal switch and domain ID allocation policy.
- Principal switches for different VSANs can reside on different physical switches.
- Each switch has a separate domain ID for each active VSAN.
- Domain IDs can overlap between VSANs.
- Domain ID and FC-ID allocation policy, static or dynamic.
- All ports are originally in VSAN 1.



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

VSANs provide isolation among devices that are physically connected to the same fabric. With VSANs, you can create multiple logical SANs over a common physical infrastructure. This functionality can achieve higher security and greater stability in Fibre Channel fabrics.

The two primary functions that the VSAN feature provides are hardware-based isolation of traffic and independent fabric services for each VSAN.

Hardware-based isolation of tagged traffic belonging to different VSANs requires no special drivers or configuration at the end nodes, such as hosts and disks. The traffic from multiple VSAN traffic can be transported over EISL ports that are configured as trunk ports. Traffic is tagged at the ingress port and the tag is removed at the egress port.

Each Fibre Channel fabric service maintains a separate database for each newly created VSAN. These services include the zone server, name server, and management server. Each service runs independently on each VSAN and is independently managed and configured.

Each VSAN has its own principal switch and domain ID allocation policy, either static or dynamic. Principal switches for different VSANs do not need to reside on the same physical switch. Each switch has a separate domain ID for each active VSAN.

As shown in the illustration, each switch that has end ports in a particular VSAN has a domain ID assigned to that particular VSAN. Core switches that trunk these VSANs also have assigned domain IDSs in these VSANs.

https://t.me/learningnets

VSAN IDs

VSAN ID	Description
VSAN 1	• Automatically configured by the switch as the default VSAN • All ports are originally in VSAN 1 • Always present, cannot be deleted
VSAN 2 through 4093	• User-configurable VSANs • A maximum of 4000 VSANs per physical switch can be created in this number range
VSAN 4079	• TF-TN port link • Carries the Exchange Virtual Fabrics Protocol (EVFP)
VSAN 4094	• Reserved VSAN known as the isolated VSAN • Used to isolate ports whose port VSAN has been deleted • Not propagated across switches • Always present, cannot be deleted

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

There are certain rules that you must follow when creating VSANs. VSAN 1, for instance, is automatically configured by the switch as the default VSAN. All ports that are configured are originally put into VSAN 1 until they are configured into another VSAN number.

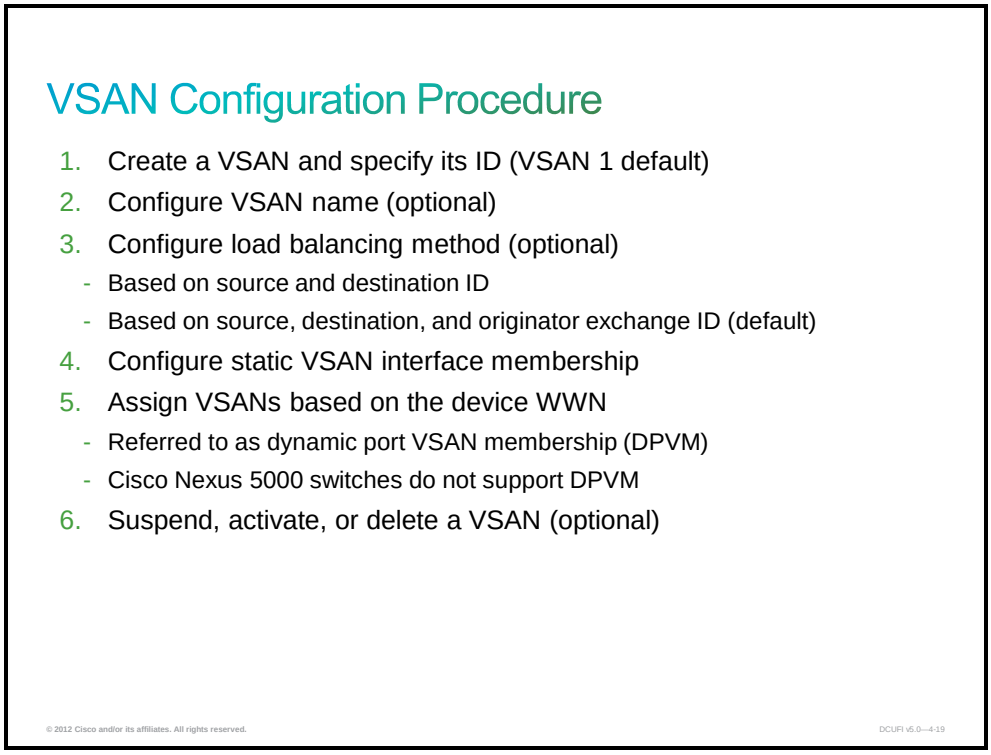
The VSAN numbers ranging from 2 through 4093 are the user-configurable VSANs.

VSAN 4094 is a reserved special VSAN that is called the isolated VSAN. It is used to temporarily isolate the ports whose VSAN has been deleted. VSAN 4094 is not propagated across switches; it is always present, and it cannot be deleted.

In the figure, VSAN 30 is not propagated across EISL, because it is not configured in the local switch, but it is configured on the remote switch. Instead of the host device on the local switch being able to connect to the remote switch, it has been placed in the isolated VSAN 4094 because the VSAN of the port, VSAN 30, has been deleted from the local switch configuration.

Note VSAN 0 and VSAN 4095 are reserved and cannot be used.

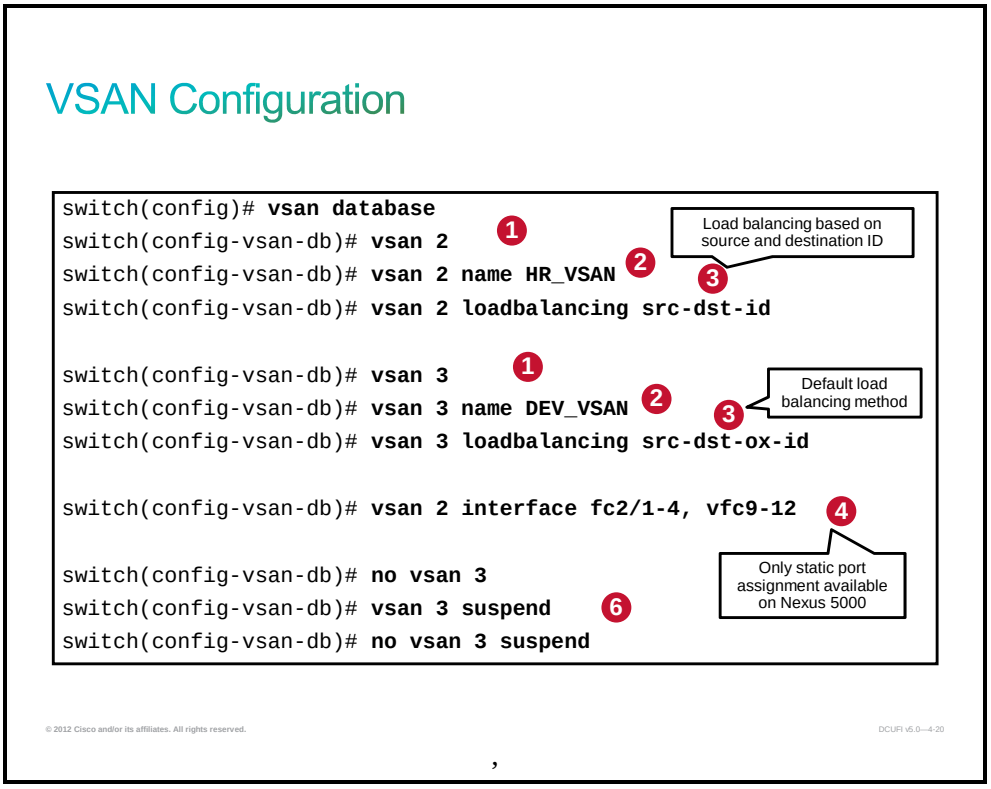
https://t.me/learningnets



Follow this procedure to create, modify, and manage VSANs on Cisco Nexus and MDS switches:

1. Create a VSAN and define its ID.
2. Optionally, configure the VSAN name. The name is a text string that identifies the VSAN for management purposes. The name can be from 1 to 32 characters long and it must be unique across all VSANs. By default, the VSAN name is a concatenation of VSAN and a four-digit string representing the VSAN ID. For example, the default name for VSAN 3 would be VSAN0003.
3. Optionally, configure the load-balancing method. Two methods are available. One method is based on the source ID (SID) and destination ID (DID). The other method uses the SID, DID, and the originator exchange ID (OXID) for load-balancing purposes. SID, DID, and OXID is the default for load-balancing path selection.
4. Assign static VSAN interface membership. Port VSAN membership on the switch is assigned on a port-by-port basis. By default, each port belongs to the default VSAN. Static VSAN interface assignment is the only method that is available on the Cisco Nexus 5000 Series Switches.
5. Assign VSANs based on the device WWN. This method is referred to as dynamic port VSAN membership (DPVM). Cisco Nexus 5000 switches do not support DPVM.
6. Optionally suspend, activate, or delete VSANs depending on the administrative requirements in your network. The administrative state of a VSAN can be set to active (default) or suspended. The active state of a VSAN indicates that the VSAN is configured and enabled. By enabling a VSAN, you activate the services for that VSAN. The suspended state indicates that the VSAN is configured but not enabled. Any ports that are associated with the suspended VSAN are disabled.

https://t.me/learningnets



To create and configure VSANs, enter configuration mode using the **config** command. Configure the VSAN database using the **vsan database** command. Define each individual VSAN ID using the **vsan** *x* command. Configure a name for the new VSAN with the **vsan** *x* **name** *abc* command. Configure the appropriate load-balancing mechanism using **the vsan** ***vsan-id*** **load balancing** command. Assign a specific Fibre Channel interface to the specified VSAN using the **vsan** *x* **interface fc***x/y*/**vfc***x* command.

Finally, you can delete, suspend, or activate a VSAN. When an active VSAN is deleted, all of its attributes are removed from the running configuration. Once deleted, all ports in that VSAN are made inactive and the ports are moved to the isolated VSAN (VSAN 4094). If the same VSAN is re-created, the ports do not automatically get assigned to that VSAN, but must be explicitly reconfigured.

© 2012 Cisco Systems, Inc.
https://t.me/learningnets

Verifying VSANs

```
switch# show vsan
vsan 1 information
         name:VSAN0001  state:active
         interoperability mode:default
         loadbalancing:src-id/dst-id/oxid
         operational state:up

vsan 2   information
         name: HR_VSAN  state:active
         interoperability mode:default
         loadbalancing:src-id/dst-id
         operational state:up
<output omitted>

vsan 4079:evfp_isolated_vsan

vsan 4094:isolated_vsan

switch# show vsan membership
vsan 1 interfaces:
vsan 2 interfaces:
    fc2/1             fc2/2             fc2/3             fc2/4
    vfc9              vfc10             vfc11             vfc12
vsan 4079(evfp_isolated_vsan) interfaces:
vsan 4094(isolated_vsan) interfaces:
```

Display VSAN attributes and state

View VSAN membership

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

To verify the current VSAN configuration and status, use the **show vsan** command. This command will display the following:

- VSANs created
- VSAN name
- Administrative state (active and suspended)
- Interoperability setting
- Load-balancing scheme
- Operational state (up or down)

Using the **show vsan** *x* command would display this information for a specific VSAN.

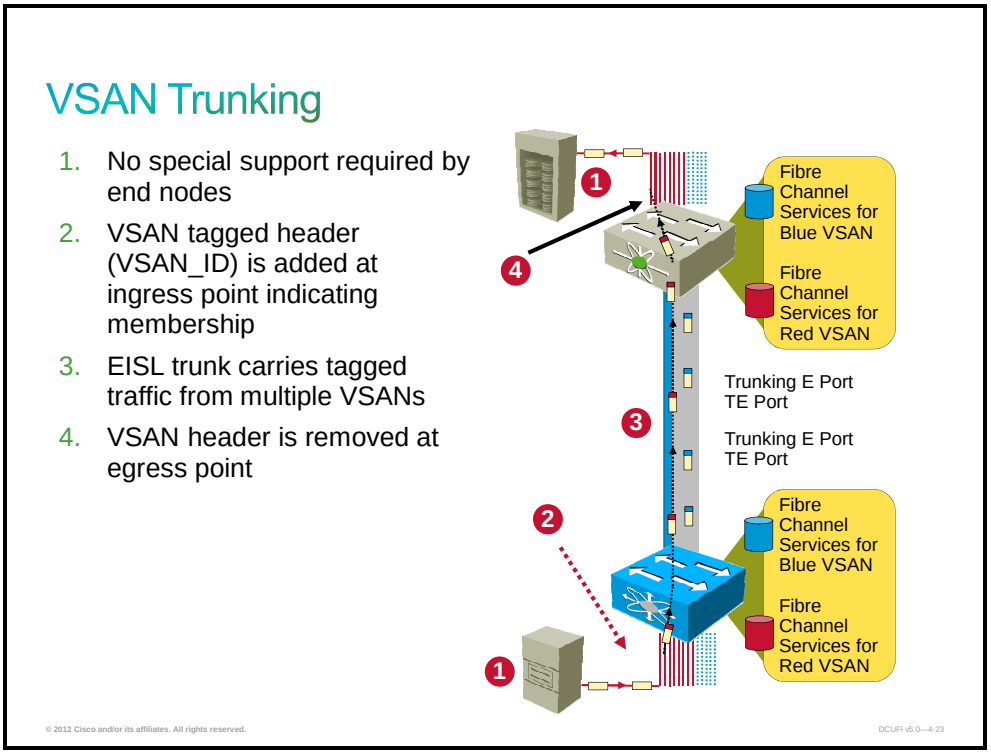
To verify port VSAN memberships, use the **show vsan membership** command. The report displays port VSAN assignments.

To report the membership for a specific VSAN, add the VSAN number to the command: **show vsan** *x* **membership**.

https://t.me/learningnets

VSAN Trunking

This topic describes how to configure interfaces to support VSAN trunking.



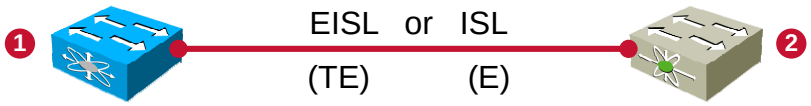
VSAN trunking enables interconnected ports to transmit and receive frames in more than one VSAN. Trunking is supported on E Ports and F Ports. Trunking mode is specific to Cisco switches that support Fibre Channel and FCoE.

The end nodes are not aware of the fact that there are VSAN trunks in the network. When a frame is transmitted over a trunk, the ingress switch encapsulates the frame with the VSAN tagged header, which identifies the VSAN to which the frame belongs. The EISL link carries the tagged traffic that originates from multiple VSANs. The VSAN header is removed at the egress switch and forwarded as a Fibre Channel or FCoE frame toward the destination.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Trunk Mode Negotiation Results



Administrative Trunk Mode (1)	Administrative Trunk Mode (2)	Operational Trunk Mode
ON	ON	TE Port
ON	OFF	E Port
ON	AUTO	TE Port
AUTO	AUTO	E Port

Default administrative trunk mode: On

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

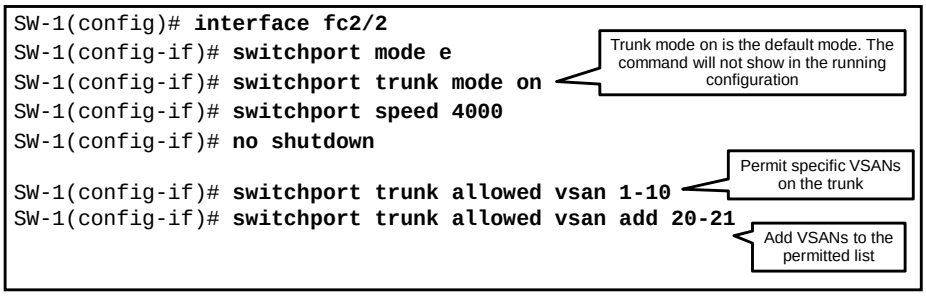
Three trunking mode options are available on Cisco Nexus and MDS switches:

- **On:** Explicitly enable trunking mode on.
- **Off:** Explicitly enable trunking mode off.
- **Auto:** A passive mode where the other end of the connection should indicate the desire to become a trunk. Otherwise, the trunk mode will be off.

As shown in the table, two connected switches negotiate the trunk mode when at least one of them has the administrative trunk mode set to On. The other may have the administrative trunk mode set to On or Auto.

https://t.me/learningnets

Configuring VSAN Trunks



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

The **switchport mode e** command configures the interface for E Port operations. By enabling the **switchport trunk mode on** command, the interface is being explicitly enabled for trunking. **Switchport trunk mode on** is the default option for ISLs.

The **switchport trunk** command has two options—the allowed VSAN list and the trunk mode configuration. You can limit the VSANs that are carried over the trunk with the **switchport trunk allowed vsan** command. You can add further VSANs to the trunk using the **switchport trunk allowed vsan add** command. A TE Port allows traffic for active allowed VSANs only. If the allowed list is not configured correctly, certain VSAN traffic will not be able to traverse the link. The default is to trunk for all VSANs.

https://t.me/learningnets

Verifying VSAN Trunks

```
switch# show interface fc2/2 brief
-------------------------------------------------------------
Interface  Vsan  Admin Admin Status  SFP  Oper Oper  Port
                 Mode  Trunk              Mode Speed Channel
                       Mode                    (Gbps)
-------------------------------------------------------------
fc2/2      1      E    on    trunking swl   TE   4    --

switch# show interface trunk vsan 1-1000
fc3/1 is not trunking
...
fc3/11 is trunking
    Belongs to san-port-channel 6
    Vsan 1 is up, FCID is 0xef0000
    Vsan 2 is up, FCID is 0xef0000
...
san-port-channel 6 is trunking
    Vsan 1 is up, FCID is 0xef0000
    Vsan 2 is up, FCID is 0xef0000
```

VSANs carried over the trunk

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

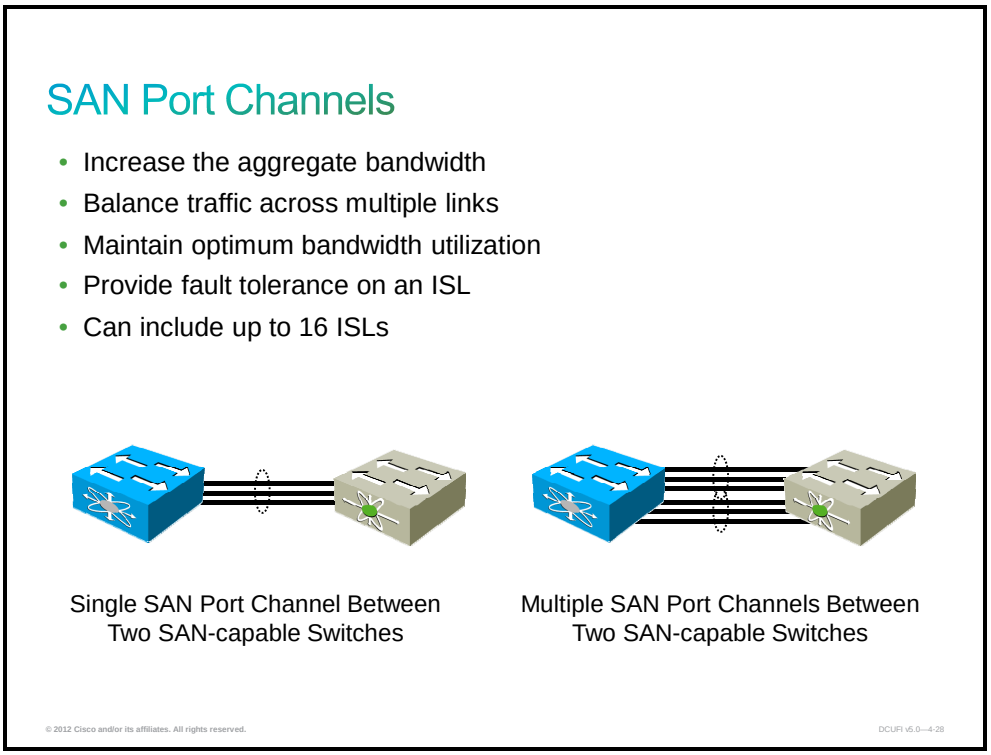
You can verify the VSAN trunk ports using various commands. The first example shows the **show interface** command, which displays the port mode and status.

The second example illustrates the use of the **show interface trunk vsan** command. It displays all interfaces that are in the trunking mode, including the allowed VSANs with their FCIDs.

https://t.me/learningnets

SAN Port Channels

This topic explains how to configure the Cisco Nexus and Cisco MDS Families of switches to support SAN port channels.



SAN port channels are an aggregation of multiple physical interfaces into one logical interface to provide higher aggregated bandwidth, load balancing, and link redundancy.

A SAN port channel has the following features:

- Provides a point-to-point connection over ISL (E Ports) to EISLs (TE Ports). Multiple ports can be combined in a SAN port channel.
- Increases the aggregate bandwidth on an ISL by distributing traffic among all functional links in the channel.
- Load-balances across multiple links and maintains optimum bandwidth utilization. Load balancing is based on the SID, DID, and exchange ID (XID).
- Provides high availability on an ISL. If one link fails, then traffic that was previously carried by this link is switched to the remaining links. If a link in a SAN port channel goes down, the upper layer protocol is not aware of it. Routing tables are not affected by the link failure. SAN port channels can contain up to 16 physical links.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Configuring SAN Port Channels

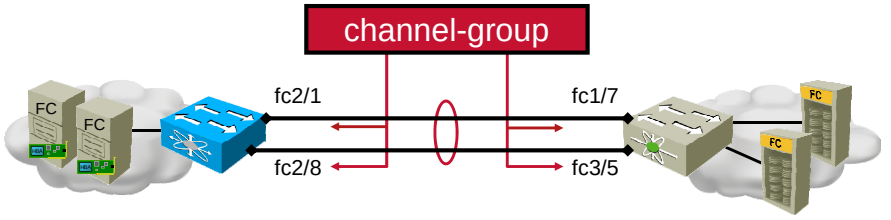
- Two configuration methods: manual and autocreation
- Manual configuration: SAN port channel interface, port channel, and member ports
- Autocreation based on autocreation protocol, shown below:

```
N5K-1(config)# interface fc2/1
N5K-1(config-if)# channel-group auto
N5K-1(config-if)# no shutdown

N5K-1(config-if)# interface fc2/8
N5K-1(config-if)# channel-group auto
N5K-1(config-if)# no shutdown
```

```
MDS-1(config)# interface fc1/7
MDS-1(config-if)# channel-group 1
MDS-1(config-if)# no shutdown

MDS-1(config-if)# interface fc3/5
MDS-1(config-if)# channel-group 1
MDS-1(config-if)# no shutdown
```



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

The port channel protocol expands the port channel functional model in Cisco SAN switches. The port channel protocol is enabled by default and contains a subprotocol, called the autocreation protocol, which automatically aggregates compatible ports into a SAN port channel.

If channel group autocreation is enabled, ISLs can be configured automatically into channel groups without manual intervention. The figure shows an example of channel group autocreation. Autocreation must be enabled in both the local and peer ports to negotiate a SAN port channel.

In the alternative manual configuration approach, you need to create the SAN port channel interface using the **interface san-port-channel** *x* command, configure a port channel, and configure the member ports. The valid range for a SAN Port Channel interface ID is 1 to 256. Both sides need to have symmetric settings.

A port is not allowed to be configured as part of a SAN port channel when the autocreation feature is enabled. The manual configuration and the autocreation are mutually exclusive.

https://t.me/learningnets

Verifying SAN Port Channels

```
switch# show san-port-channel summary
-----------------------------------------------------------------
Interface          Total Ports     Oper Ports   First Oper Port
-----------------------------------------------------------------
san-port-channel 1       2              1             fc2/1

switch# show san-port-channel database
san-port-channel 1
    Administrative channel mode is active
    Operational channel mode is active
    Last membership update succeeded
    First operational port is fc2/1
    2 port in total, 1 port up
    Ports:   fc2/1 [up] *

switch# show interface fc2/1
fc2/1 is trunking
    Hardware is Fibre Channel, FCOT is short wave laser
    Port WWN is 20:0a:00:0b:5f:3b:fe:80
    ...
    Port-channel auto creation is enabled
    Belongs to port-channel 1
<output omitted>
```

The first operational port (FOP) is the primary operational interface to carry control-plane traffic. No load balancing is involved for control traffic.

Status of autocreation protocol

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

The **show san-port-channel summary** command displays a summary of SAN port channels within the switch. A one-line summary of each SAN port channel provides the administrative state, the operational state, the number of attached and active interfaces (up), and the first operational port (FOP), which is the primary operational interface that is selected in the SAN port channel to carry control-plane traffic (no load balancing). The FOP is the first port that comes up in a SAN port channel and can change if the port goes down. The FOP is also identified by an asterisk (*).

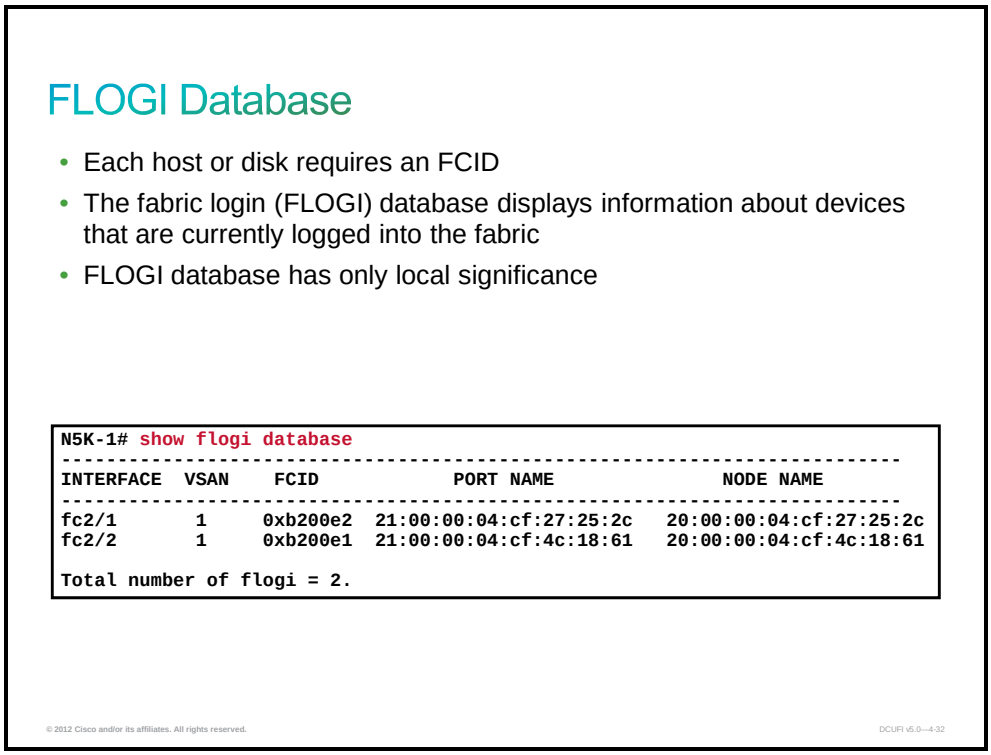
To display additional information about the administrative and operational channel modes, you can use the **show san-port-channel database** command. The **show interface** command provides the details on the autocreation function and additional parameters, such as port WWN (pWWN).

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

FLOGI and FCNS Databases

This topic describes how to manage the fabric login (FLOGI) and Fibre Channel Name Servers (FCNS) databases on the Cisco Nexus 5000 Series Switches.



In a Fibre Channel fabric, each host or disk requires an FCID. The FLOGI database displays information about devices that are currently logged into the fabric on the switch on which the command is issued.

The **show flogi database** command is used to verify if a storage device is displayed in the FLOGI database. If the required device is displayed in the FLOGI database, the FLOGI was successful. The FLOGI database is of local significance only.

https://t.me/learningnets

FCNS Database

- Stores name entries for all hosts in a VSAN
- Populated by the name server
  - Nx port is permitted to register attributes during a port login (PLOGI)
  - The attributes are deregistered when the Nx port logs out
- Distributed across all switches in a VSAN

```
N5K-1# show fcns database

VSAN 1:
--------------------------------------------------------------------------
FCID        TYPE  PWWN                    (VENDOR)        FC4-TYPE:FEATURE
--------------------------------------------------------------------------
0xb200e2    N     21:00:00:04:cf:27:25:2c (Seagate)       scsi-fcp
0xb200e1    N     21:00:00:04:cf:4c:18:61 (Cisco)         ipfc

Total number of entries = 2
```

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

The FCNS database displays information about devices that are currently registered in the fabric.

The name server stores name entries for all hosts in the FCNS database. An Nx port is permitted to register attributes during a port login (PLOGI) by the name server, and to obtain attributes of other hosts. These attributes are deregistered when the Nx port logs out either explicitly or implicitly. Use the **show fcns database** command to display this database.

The FCNS database is a distributed database across all switches in a single fabric (VSAN) containing entries for all devices in that fabric.

© 2012 Cisco Systems, Inc.

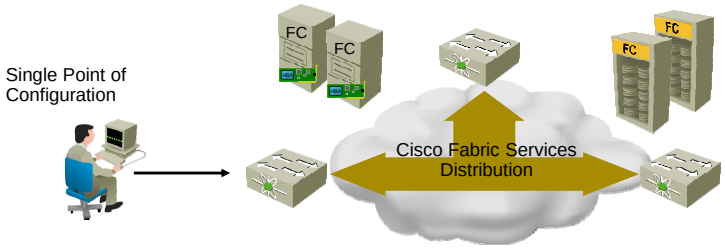
https://t.me/learningnets

Cisco Fabric Services on Cisco MDS

This topic explains how to manage Cisco Fabric Services on the Cisco MDS Switch.

Cisco Fabric Services on MDS

- Cisco Fabric Services is a facility for synchronizing and distributing application configurations to all switches in a fabric
- Both in-band and out-of-band protocol
- Single point of configuration ensures fabric-wide consistency
- Discovery of Cisco Fabric Services-capable switches and applications



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

Many features require configuration synchronization in all switches in the fabric. Maintaining configuration synchronization across a fabric is important to maintain fabric consistency. In the absence of a common replication infrastructure, this synchronization is achieved through manual configuration of each switch within the fabric. This manual process is tedious and error-prone.

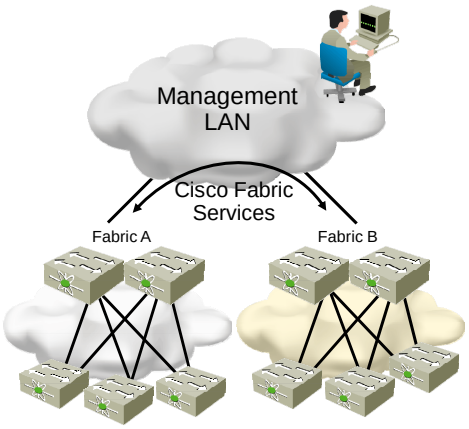
The Cisco Nexus Operating System (NX-OS) Software uses the Cisco Fabric Services infrastructure to enable efficient database distribution and to foster device flexibility. Cisco Fabric Services simplifies SAN provisioning by automatically distributing configuration information to all switches within a fabric. Several Cisco NX-OS applications use the Cisco Fabric Services infrastructure to maintain and distribute the contents of the database of a particular application.

Cisco Fabric Services provides a common infrastructure for automatic configuration synchronization in the fabric. It provides the transport function as well as a rich set of CiscoWorks Common Services to the applications. Cisco Fabric Services has the ability to discover Cisco Fabric Services-capable switches in the fabric and can discover the application capabilities of all Cisco Fabric Services-capable switches.

https://t.me/learningnets

Cisco Fabric Services Distribution Methods

- Switches attempt distribution over Fibre Channel first; upon failure, distributes over IP
- Does not send duplicate messages if distribution over both IP and Fibre Channel is enabled
- Cisco Fabric Services applications with the physical-all distribution scope can be used with IP distribution
- Two distribution modes:
  - Uncoordinated method distributes information that is not expected to conflict with information from a peer
  - Coordinated distribution used where an application can have only one distribution at a given time. A fabric-wide lock is applied to enforce this mode of distribution.



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

Cisco Fabric Services distribution can be performed out-of-band (OOB) over IP to synchronize Cisco Fabric Services-enabled applications for physically separate fabrics.

The switch attempts to distribute information over Fibre Channel first, and then over the IP network, if the first attempt over Fibre Channel fails. Cisco Fabric Services does not send duplicate messages if distribution over both IP and Fibre Channel is enabled.

Cisco Fabric Services distribution over IP supports the following features:

- Physical distribution entirely over an IP network
- Physical distribution over a hybrid Fibre Channel and IP network, with the distribution reaching all switches that are reachable over either Fibre Channel or IP

Note	Distribution can be over IP version 4 (IPv4) or IP version 6 (IPv6). Cisco Fabric Services cannot distribute over both IPv4 and IPv6 from the same switch.

The Cisco Fabric Services keepalive mechanism uses a configurable multicast address to detect network topology changes. The distribution for logical scope applications is not supported, because the VSAN implementation is limited to Fibre Channel.

Cisco Fabric Services data can be broadcast to all peers in the fabric or unicast to a single peer. Cisco Fabric Services supports the following distribution modes:

- **Uncoordinated distributions:** Used to distribute information that is not expected to conflict with information from a peer.
- **Coordinated distributions:** Used where an application can have only one distribution at a given time. A fabric-wide lock is applied to enforce this mode of distribution.

https://t.me/learningnets

Cisco Fabric Services Applications

Application	Distribution Mode	Scope	Cisco Fabric Services Enabled?
Call Home	Coordinated	Physical	No
Global Device Alias	Coordinated	Physical	Yes
DPVM	Coordinated	Physical	Yes
IVR	Both	Physical	No
iSNS	Uncoordinated	Physical	Yes
NTP	Coordinated	Physical	No
Port Security	Coordinated	Logical	No
RADIUS and TACACS+	Coordinated	Physical	No
Role (Role-Based Access Control)	Coordinated	Physical	No
Syslog	Coordinated	Physical	No
VSAN fctimer (fabric timeout values)	Coordinated	Physical	No
fcdomain allowed list	Coordinated	Logical	No
RSCN event timer	Coordinated	Logical	No
SCSI Flow Services	Coordinated	Selected	Yes
iSCSI Load Balancing	Coordinated	Physical	No
Fabric Startup Configuration Manager	Coordinated	Physical	Yes
FlexAttach Virtual pWWN	Coordinated	Physical	Yes

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

The table lists the Cisco NX-OS features and applications that use the Cisco Fabric Services infrastructure.

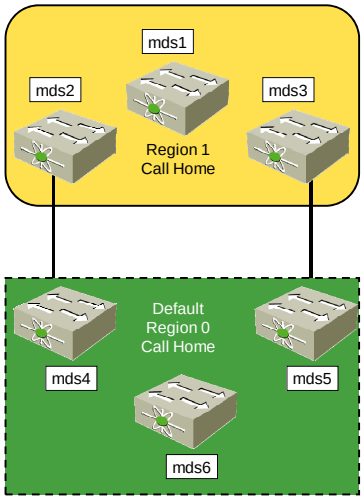
The scope of the application synchronization varies based on the application. Cisco Fabric Services deploys two distribution scopes:

- **Logical scope:** The configuration distribution is restricted to the VSAN. An example of such an application is port security, where the configuration database is applicable only within a given VSAN.
- **Physical topology:** The distribution is switch-based and is performed regardless of the existing VSAN configuration. Examples of such applications are Network Time Protocol (NTP) and DPVM, which are independent of VSANs. Note that DPVM membership is based on the WWNs of VSANs.

https://t.me/learningnets

Cisco Fabric Services Regions

- SANs may have several administrators requiring different application requirements
- Example: Apply different Call Home profiles on different sets of switches to alert the correct administrator
- Cisco Fabric Services Regions provide the ability to support distribution islands within a physical fabric
- Cisco Fabric Services regions apply only to applications that work in the physical scope



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

Large enterprise SANs may have several administrators, each with various levels of administrative responsibilities that dictate different application requirements. Demand from the field requested the ability to create multiple *islands of distribution* for a given application within the fabric. The Cisco Fabric Services distribution infrastructure allows applications to distribute information to other switches in a fabric. Different Cisco Fabric Services distribution scopes include the following:

- **Logical:** Distribution is limited to a VSAN.
- **Physical:** Distribution spans the physical fabric.

Cisco Fabric Services regions let you support distribution islands within a physical fabric. Cisco Fabric Services regions apply only to applications that work in the physical scope. For example, there may be a requirement for different Call Home profiles on different sets of switches to alert the appropriate administrator in the event of a system failure (Call Home is a physical scope application).

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

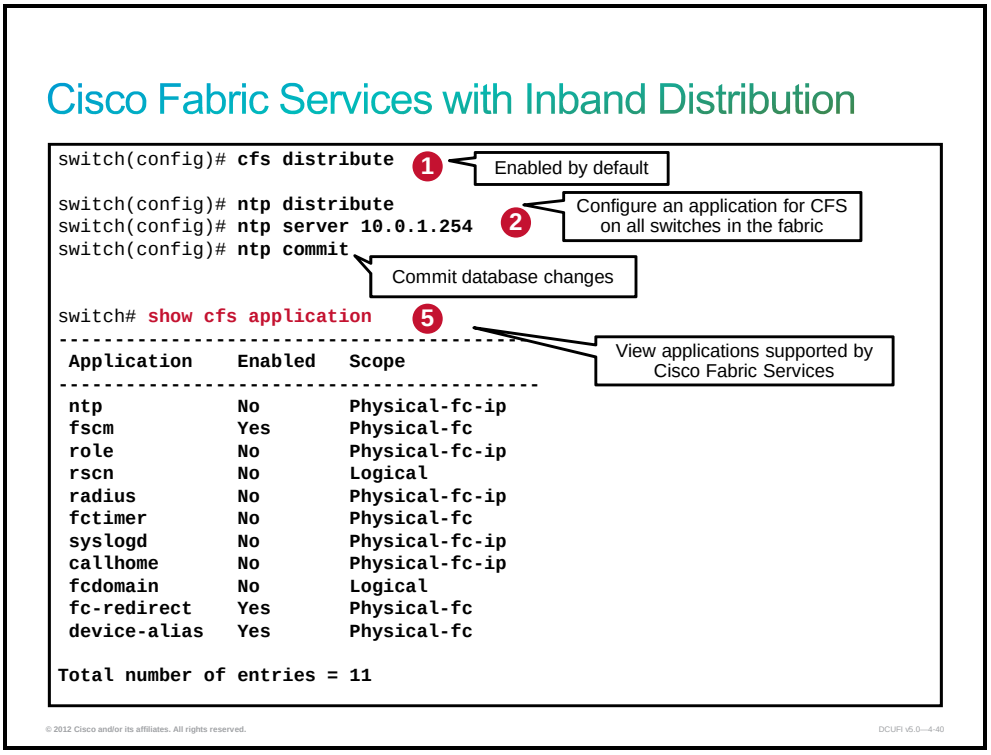
Cisco Fabric Services Implementation Procedure

1. Ensure Cisco Fabric Services distribution in global configuration mode
   - Enabled by default
   - Can be disabled and re-enabled
2. Configure a given application for Cisco Fabric Services
   - On all switches in the fabric
   - Commit the pending database changes
3. Configure distribution over IPv4 or IPv6 via one of two methods (optional)
   - IP multicast
   - Static peer addresses
4. Configure regions (optional)
5. View applications supported by Cisco Fabric Services (optional)
6. Display Cisco Fabric Services-capable switches in the fabric
   - Independent of individual application registrations
   - Local switch is indicated as [Local]

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

By default, Cisco Fabric Services distribution is enabled. Applications can distribute data and configuration information to all Cisco Fabric Services-capable switches in the fabric where the applications exist. This is the normal mode of operation. You can globally disable Cisco Fabric Services on a switch, including Cisco Fabric Services over IP, to isolate the applications using Cisco Fabric Services from fabric-wide distributions while maintaining physical connectivity. When Cisco Fabric Services is globally disabled on a switch, Cisco Fabric Services operations are restricted to the switch, and all Cisco Fabric Services commands continue to function as if the switch were physically isolated. Use the **cfs distribute enable** command to globally enable Cisco Fabric Services distribution for the switch.

https://t.me/learningnets

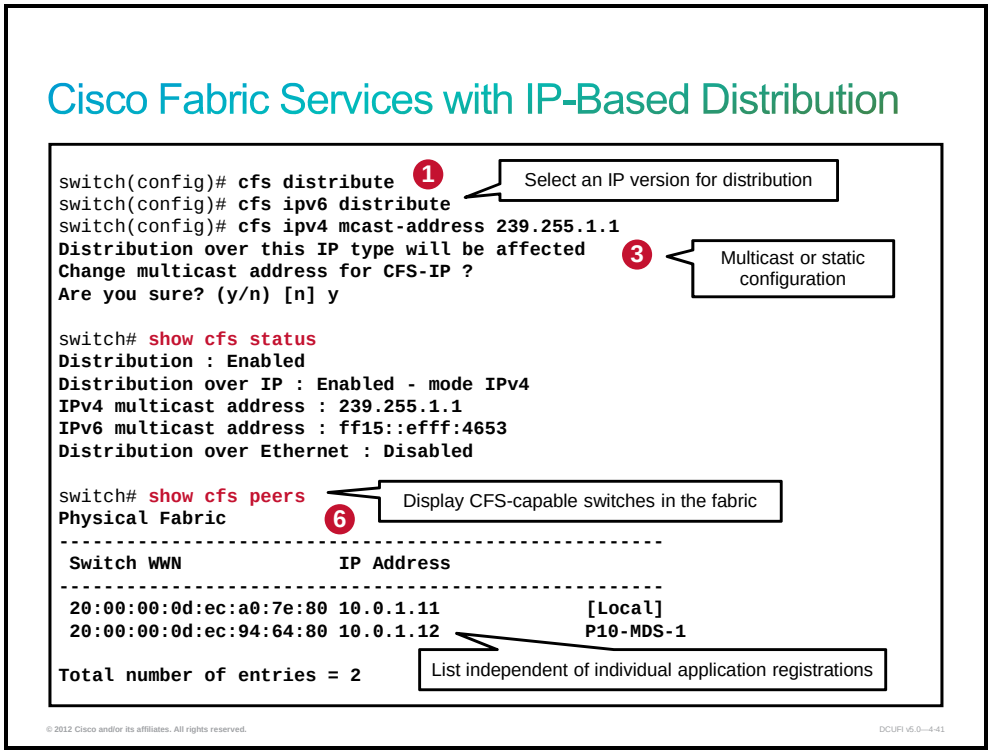


First you may use the **show cfs application** command to view the list of applications that support Cisco Fabric Services. Some applications are enabled by default. For others, the application does not appear in the list until the feature is first enabled. If not already enabled, enable distribution for the given application on all switches in the fabric. In this example, you enable Cisco Fabric Services distribution for NTP by using the **ntp distribute** command.

You configure the application. In this example, you configure an NTP server with the address 192.168.0.198 using the **ntp server 10.0.1.254** command. For applications using coordinated distributions, the switch that begins the configuration acquires a systemwide lock.

Then, you need to commit the pending database. In the example, you commit the preceding NTP configuration using the **ntp commit** command. A commit task saves the pending database for all application peers and releases the lock for all switches.

© 2012 Cisco Systems, Inc.
https://t.me/learningnets



This example shows how to configure the Cisco Fabric Services distribution to be performed out of band over IP to synchronize Cisco Fabric Services-enabled applications for physically separate fabrics.

To globally enable or disable Cisco Fabric Services over IPv4 or IPv6, issue the **cfs ipv4 distribute** or the **cfs ipv6 distribute** command as illustrated in the figure.

You can verify the Cisco Fabric Services over IP configuration using the **show cfs status** command.

The **show cfs peers** command displays all the switches in the physical fabric in terms of the switch WWN and the IP address. The local switch is indicated as [Local].

https://t.me/learningnets

Cisco Fabric Services with Regions

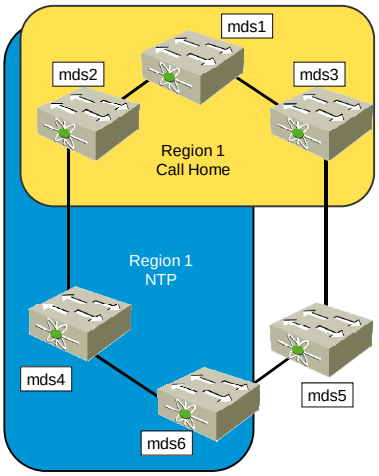
```
mds1/2(conf)# cfs region 1
mds1/2(conf-cfs-region)# callhome
mds1/2(conf-cfs-region)# ntp
```
4

```
mds3(conf)# cfs region 1
mds3(conf-cfs-region)# callhome
```
4

```
mds4/5(conf)# cfs region 1
mds4/5(conf-cfs-region)# ntp
```
4

```
mds1# show cfs regions brief
----------------------------------------
 Region         Application    Enabled
----------------------------------------
   1            ntp            yes
   1            callhome       no

mds1# show cfs regions
Region-ID  : 1
Application: ntp
Scope      : Physical-fc-ip
----------------------------------------------
 Switch WWN              IP Address
-------------------------------------------------------------------
 20:00:00:0d:ec:a0:7e:80 10.0.1.11                    [Local]
 20:00:00:0d:ec:94:64:80 10.0.1.12                     mds2
<output omitted>
```



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

By default, all switches reside in Region 0 for any given Cisco Fabric Services physical scope application.

In the example, you configure mds1 and mds2 in Call Home Region 1, mds3 in Call Home Region 1, and mds4 and mds6 in NTP Region 1.

You can use the **show cfs regions** command to verify Cisco Fabric Services region status. The command displays the complete Cisco Fabric Services region database for distribution-enabled applications.

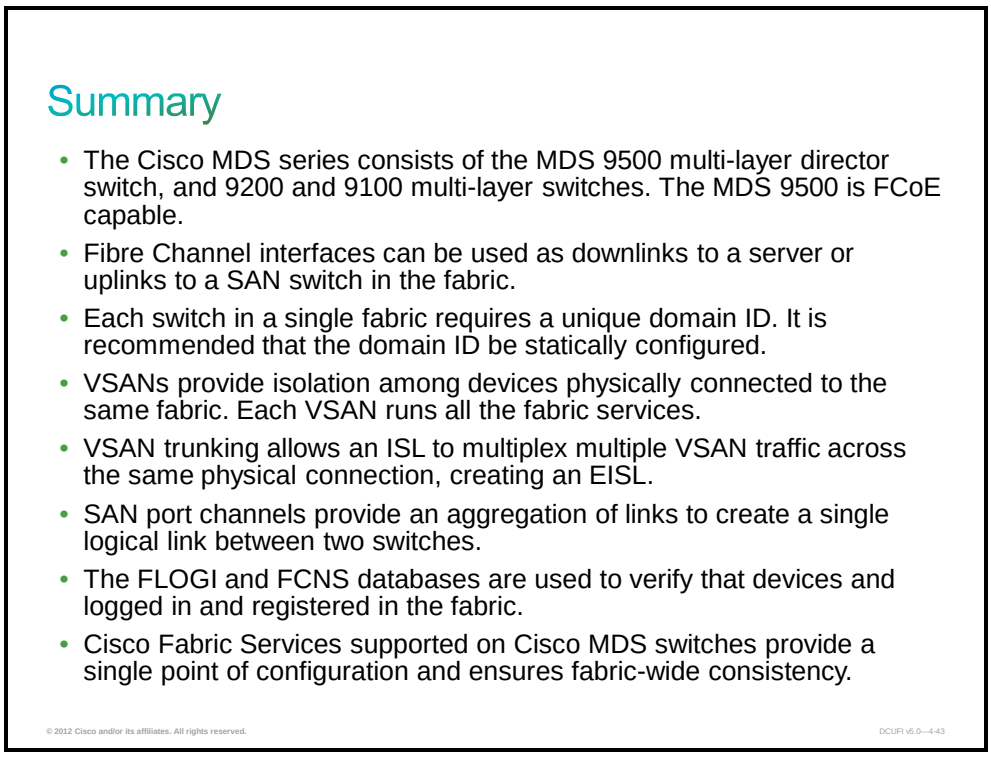
The **show cfs regions brief region** command restricts the display of the Cisco Fabric Services region database. Other variants of this command include the **show cfs regions region** command and the **show cfs regions name** command.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Summary

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



References

For additional information, refer to these resources:

- To learn more about configuring SAN switching on the Cisco Nexus 5000 and 5500 Platform switches, refer to *Cisco Nexus 5000 Series NX-OS SAN Switching Configuration Guide, Release 5.1(3)N1(1)* at: http://www.cisco.com/en/US/docs/switches/datacenter/nexus5000/sw/san_switching/513_n1_1/b_Cisco_n5k_nxos_sanswitching_config_guide_rel513_n1_1.html
- To learn more about configuring SAN switching on the Cisco Nexus 7000 switches, refer to *Cisco Nexus 7000 Series NX-OS SAN Switching Configuration Guide* at: http://www.cisco.com/en/US/docs/switches/datacenter/sw/nx-os/san_switching/configuration_guide/b_Cisco_Nexus_7000_NX-OS_SAN_Switching_Configuration_Guide.html

https://t.me/learningnets

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Lesson 6

Configuring NPV Mode

Overview

NPIV allows multiple applications on the N port to use different identifiers, and allows access control, zoning, and port security to be implemented at the application level. In contrast, NPV alleviates the domain ID limit by sharing the domain ID of the core switch among multiple edge switches.

In this lesson, you will learn how to implement N-Port ID Virtualization (NPIV), N-Port Virtualization (NPV), and FCoE NPV on the Cisco Nexus 5000 and 5500 Switches.

Objectives

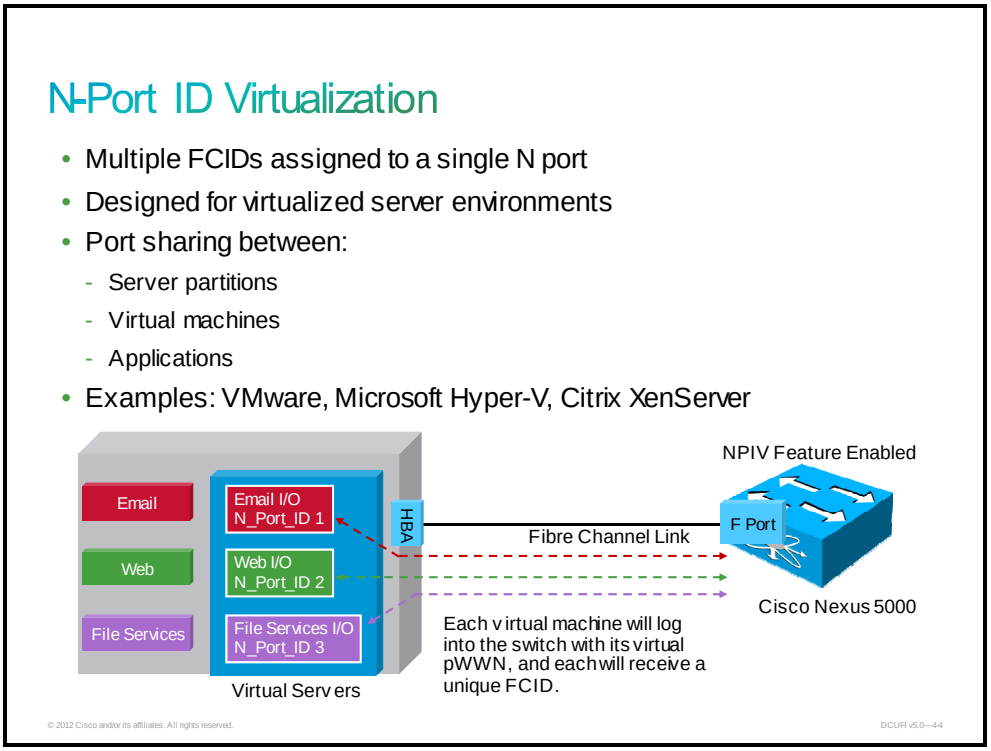
Upon completing this lesson, you will be able to configure the Cisco Nexus and Cisco MDS Families of switches to support NPIV, NPV, and FCoE NPV based on a given implementation plan. This ability includes being able to meet these objectives:

- Describe the use of the NPIV feature on the Cisco Nexus and MDS Families of switches
- Configure NPIV on the Cisco Nexus and MDS Families of switches
- Describe the NPV feature
- Configure NPV mode on the Cisco Nexus and MDS Families of switches
- Describe the FCoE NPV feature
- Configure the FCoE NPV feature on the Cisco Nexus Family of switches

https://t.me/learningnets

N-Port ID Virtualization

This topic identifies the purpose of the NPIV feature on the Cisco Nexus and MDS switches.



N-Port ID Virtualization (NPIV) allows a Fibre Channel host connection (N Port), to be assigned multiple node port (N Port) IDs or Fibre Channel IDs (FCIDs) over a single link. All FCIDs that are assigned can be managed on a Fibre Channel fabric as unique entities on the same physical host. Different applications can be used with NPIV, such as server partitions, virtual machines, or applications. Examples include VMware, Microsoft Hyper-V, and Citrix XenServer.

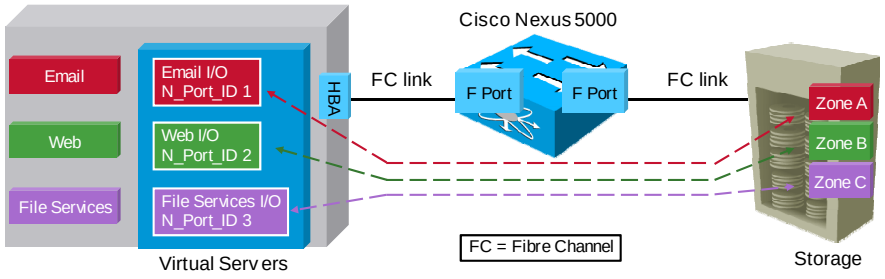
NPIV is supported on the Cisco Nexus 5000 and 5500 Platform switches to provide an attachment point for virtualized systems.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Implementing NPIV

- Separate fabric logins and pWWNs enable:
  - Port security
  - Zoning
  - Traffic management (for example, QoS)
- All virtual N Ports belong to the same VSAN
- Standards-based (T11)



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

DCUFI v5.0—4-5

In a virtual machine environment where many host operating systems or applications are running on a physical host, NPIV allows each virtual machine to be managed independently. The separation is possible as a result of different fabric logins (FLOGIs) and port WWNs (pWWNs) used by each virtual machine. The network can provide separation in zoning, aliasing, security, and quality of service (QoS) treatment.

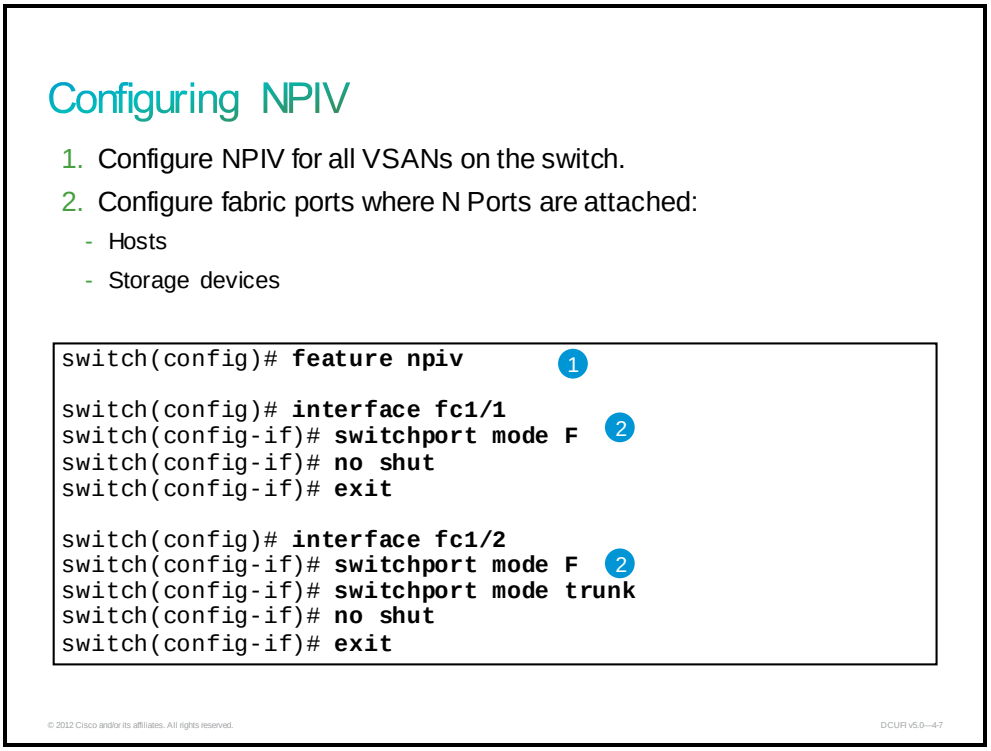
A host bus adapter (HBA) that supports the NPIV feature follows the standard login process. The initial connection and login to the fabric is performed through the standard F Port login (FLOGI) process. All subsequent logins for either virtual machines or logical part ions on a mainframe are transformed into fabric discovery (FDISC) login commands. The FDISC logins follow the same standard process. Each subsequent FDISC login will be a part of the same virtual storage-area network (VSAN) as the original FLOGI.

The NPIV has been defined by the Technical Committee T11 in the Fibre Channel: Link Services (FC-LS) specification.

https://t.me/learningnets

NPIV Configuration

This topic describes how to configure NPIV on the Cisco Nexus and MDS switches.



NPIV is not enabled by default. It must be enabled on core switches to which an NPV-enabled device is connected.

To complete the N Port virtualization configuration, the Fibre Channel fabric switch should be enabled for NPIV from the global configuration mode.

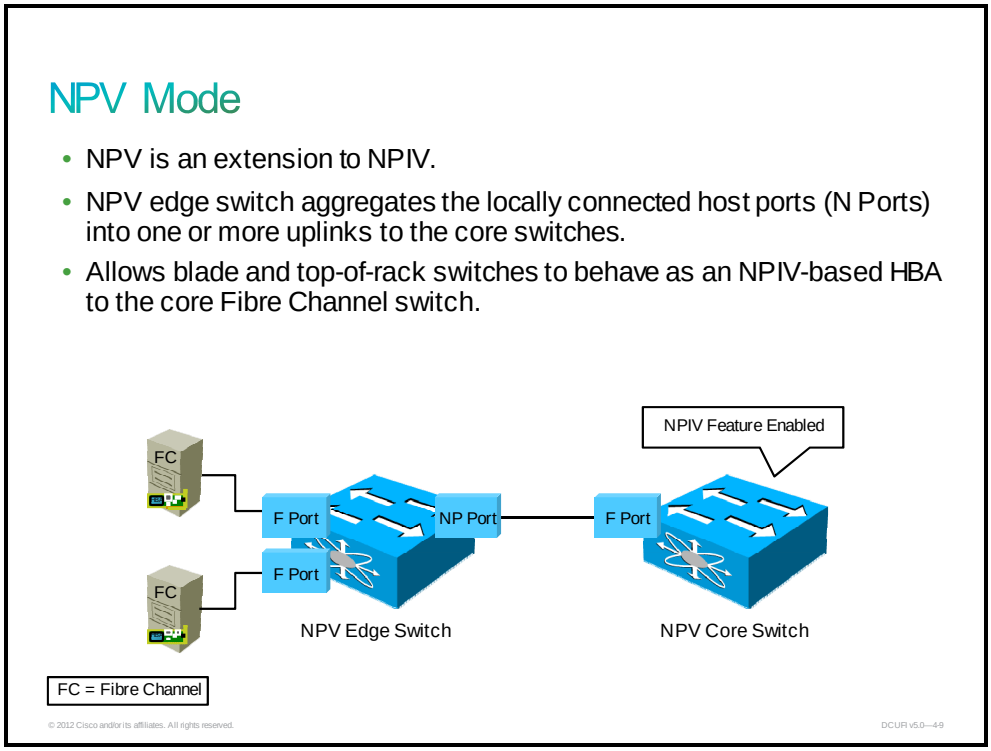
Next, you can configure the Fibre Channel interfaces where the servers are attached for the fabric mode using the **switchport mode F** command. The fabric mode is the default mode and therefore does not need to be explicitly configured. You would use the fabric mode for other systems, such as storage devices.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

NPV Mode

This topic describes the NPV feature.



Normally, each Fibre Channel switch has one domain ID and connects to the upper layer Fibre Channel switch through an expansion port (E Port). The host that is connected to the switch first sends the FLOGI to the switch. The switch intercepts the FLOGI and assigns the FCID to the host by sending an Accept frame. The Fibre Channel switch also provides other Fibre Channel services such as name registration, binding check, zoning check, and so on. The Fibre Channel switch runs the routing protocol Fabric Shortest Path First (FSPF). The switch computes the path to reach the remote domains using the Fibre Channel switch table. The Fibre Channel switch maintains the FCID to the port mapping table (station table) for a locally attached host. When the Fibre Channel switch receives a frame, it checks the domain ID of the destination ID (DID) first. If the domain ID is the same as its own domain ID, then the local station table is used to figure out the destination port. If the domain ID is different from its own domain ID, then the Fibre Channel switch searches the Fibre Channel switch table to identify the egress port.

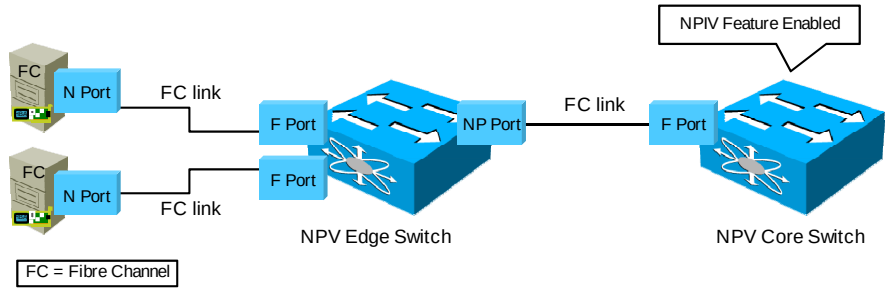
The N-Port Virtualization (NPV) feature is an extension to NPIV. The NPV allows the blade switch or top-of-rack (ToR) fabric device to behave as an NPIV-based HBA to the core Fibre Channel switch. The device aggregates the locally connected host ports (N Ports) into one or more uplinks (pseudo-interswitch links) to the core switch.

The core switch, to which the NPV edge switch is connected, needs to support NPIV.

https://t.me/learningnets

NPV Interfaces

- Classic NPV on Cisco Nexus 5000 and 5500 Platform switches:
  - Server interfaces are physical Fibre Channel ports in F mode.
  - External interfaces are physical Fibre Channel ports in NP Port mode.
- FCoE NPV on Cisco Nexus 5000 and 5500 Platform switches:
  - Supports FCoE-based interfaces
  - Available in Cisco NX-OS Release 5.0(3)N2(1) and later



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

The ports that connect to the upstream Fibre Channel switch are called border ports, or external ports. The interfaces that are connected to the upstream Fibre Channel switch are the proxy N port (NP Port) interfaces.

The ports that connect to the hosts are called server ports. The server ports must be configured as fabric ports (F Ports).

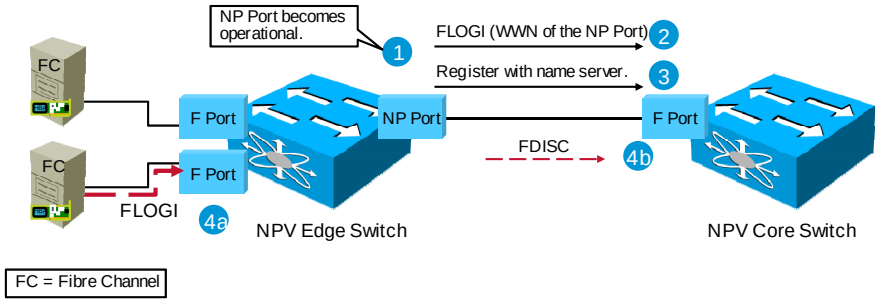
In classic NPV, both server and external interfaces are physical Fibre Channel ports. The FCoE NPV feature, available in Cisco Nexus Operating System (NX-OS) Software Release 5.0(3)N2(1) and later, offers the support for FCoE interfaces. The server and external ports in FCoE NPV are thus virtual Fibre Channel (vFC) interfaces. The FCoE NPV feature is discussed later in the lesson.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

NPV Initialization

1. NP Port becomes operational.
2. Edge switch logs itself into the core switch using a FLOGI request:
   - Using the pWWN of the NP Port
3. Edge switch registers itself with the name server on the core switch:
   - Using symbolic port name of the NP Port and IP address of the edge switch
4. Subsequent FLOGIs from servers are converted to fabric discovery (FDISC) messages.



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

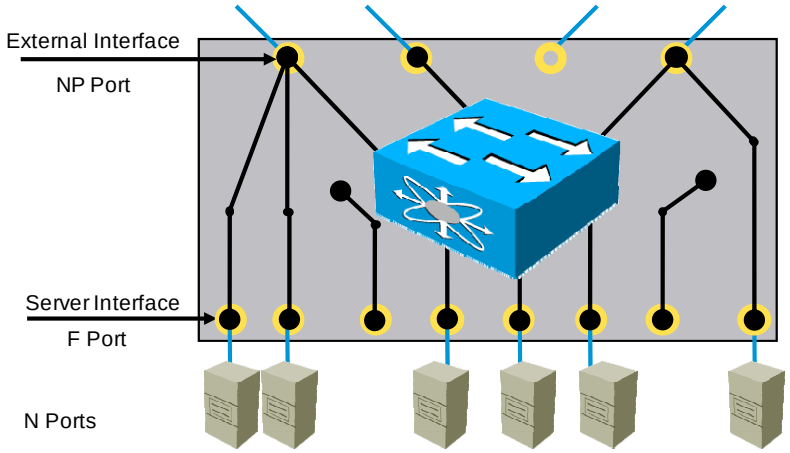
The login process for the NP Port uplink is the same as for an HBA that is NPIV-enabled; the only requirement of the core switch is that it supports the NPIV feature. As end devices log into the NPV-enabled edge switches, the FCID addresses that are assigned use the domain of the core switch. Because the connection is treated as an N Port and not an E Port to the core switch, the edge switch shares the domain ID of the core switch as FCIDs are being allocated. The edge NPV-enabled switch no longer requires a separate domain ID to receive connectivity to the fabric.

When an uplink port is enabled on an N-Port Virtualization device, it must first perform a FLOGI and process login (PRLI) to the core director and register with the Fibre Channel Name Server (FCNS). After this login process is completed for the physical uplink, the end devices can start their login processes. In the Cisco NPV process, an uplink port must first be logged into the core director for the end devices to start their login processes. As an end device initiates its login process into the edge NPV switch, it follows the standard login process with a FLOGI to the NPV edge switch. The edge switch will then convert the FLOGI to an FDISC login and allow the login to proceed down the NPV uplink. Several NPV uplinks can be connected to either a single or multiple core switches. The Cisco implementation allows connectivity from each edge switch to multiple core switches.

https://t.me/learningnets

NPV Traffic Distribution

- Different NP Ports can be connected to different core switches.
- Traffic automatically distributed from servers to NP Port uplinks.
- It can be manually distributed using traffic maps.



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

NPV supports automatic selection of NP Port uplinks. When a server interface is brought up, the NP Port uplink interface with the minimum load is selected from the available NP uplinks in the same VSAN as the server interface.

When a new NP Port uplink interface becomes operational, the existing load is not redistributed automatically to include the newly available uplink. Server interfaces that become operational after the NP Port uplink can select the new NP uplink.

In Cisco NX-OS Software Release 4.0(1a)N2(1) and later, NPV supports traffic maps. A traffic map allows you to specify the NP Port uplinks that a server interface can use to connect to the core switches. When an NPV traffic map is configured for a server interface, the server interface must select only from the NP Port uplinks in its traffic map. If none of the specified NP Port uplinks are operational, the server remains in a nonoperational state.

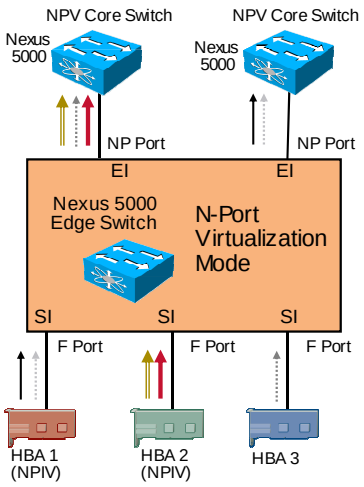
The NPV traffic map feature provides the following benefits:

- Facilitates traffic engineering by allowing configuration of a fixed set of NP Port uplinks for a specific server interface (or range of server interfaces)
- Ensures correct operation of the persistent FCID feature, because a server interface will always connect to the same NP Port uplink (or one of a specified set of NP Port uplinks) after an interface reinitialization or switch reboot

https://t.me/learningnets

NPV Disruptive Load Balancing

- Hosts are pinned to external interfaces (traffic maps or automatic distribution).
- Failed login request on one external interface is retried on another external interface.
- Disruptive load balancing:
  - Redistributes the server interfaces across all available NP Port uplinks when a new NP Port uplink becomes operational.
  - Reinitializes only server interfaces that are moved to a different uplink.
  - Disrupts the end devices that are attached to the moved server.
- Alternative: Manually reinitialize some server interfaces to distribute server traffic to new NP Port uplinks.



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

In Cisco NX-OS Software Release 4.0(0)N1(2a) and later, NPV supports disruptive load balancing. When disruptive load balancing is enabled, NPV redistributes the server interfaces across all available NP Port uplinks when a new NP Port uplink becomes operational. To move a server interface from one NP Port uplink to another NP Port uplink, NPV forces reinitialization of the server interface so that the server performs a new login to the core switch.

Only server interfaces that are moved to a different uplink are reinitialized. A system message is generated for each server interface that is moved. Redistributing a server interface causes traffic disruption to the attached end devices.

To avoid disruption of server traffic, you should enable this feature only after adding a new NP Port uplink, and then disable it again after the server interfaces have been redistributed.

If disruptive load balancing is not enabled, you can manually reinitialize some, or all, of the server interfaces to distribute server traffic to new NP Port uplink interfaces.

When you deploy NPV traffic management, follow these guidelines:

- Use NPV traffic management only when automatic traffic engineering does not meet your network requirements.
- Server interfaces configured to use a set of NP Port uplink interfaces cannot use any other available NP Port uplink interfaces, even if none of the configured interfaces are available.
- When disruptive load balancing is enabled, a server interface may be moved from one NP Port uplink to another NP Port uplink. Moving between NP Port uplink interfaces requires NPV to relogin to the core switch, causing traffic disruption.
- To link a set of servers to a specific core switch, associate the server interfaces with a set of NP Port uplink interfaces that all connect to that core switch.

https://t.me/learningnets

NPV Features

Feature	Regular Switching	NPV Edge
Fibre Channel services	• All Fibre Channel services provided (FLOGI, name server, zoning, domain server, FSPF, and management • FSPF, zoning, and name server databases are distributed among connected switches	Most Fibre Channel services are switched off.
Switching operation	Local switching enabled	Device acts as a proxy.
Domain ID	• Each switch consumes a domain ID • Theoretical limit of 239 domain IDs per VSAN	• Does not use a domain ID • No domain ID limitation
Scalability and manageability	• ISL between switches becomes a path within the FSFP Routing Table • Up to 16 ISLs may belong to a port channel	• Smaller number of switches to manage • Eliminates the need for server administrators to manage the SAN
QoS	Available	Not available

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

In NPV mode, a Cisco Nexus 5000 or 5500 Platform switch operates as a NP Port mode. It does not do any Fibre Channel switching itself. There are no local switching and zoning checks on the Cisco Nexus switches. The Cisco Nexus switch appears as a host to the upper layer switches and as a Fibre Channel switch to the server that is attached to it.

Because Cisco Nexus 5000 and 5500 Platform switches do not provide Fibre Channel services, they do not need a Fibre Channel domain ID. As a result, the Cisco Nexus 5000 and 5500 Platform switches increase the scalability and eliminate the switch-to-switch interoperability issues. They also make management easier because they do not get involved in the FSPF and Fibre Channel policy.

The regular switch mode has these characteristics:

- All Fibre Channel services are provided.
  - FLOGI, name server, zoning, domain server, FSPF, management.
  - FSPF, zoning, and name server databases are distributed among connected switches.
- Local switching is enabled.
- The interswitch link (ISL) between switches becomes a path within the FSPF routing table.
- Up to 16 ISLs may belong to a single port channel.
- Each switch consumes a domain ID.

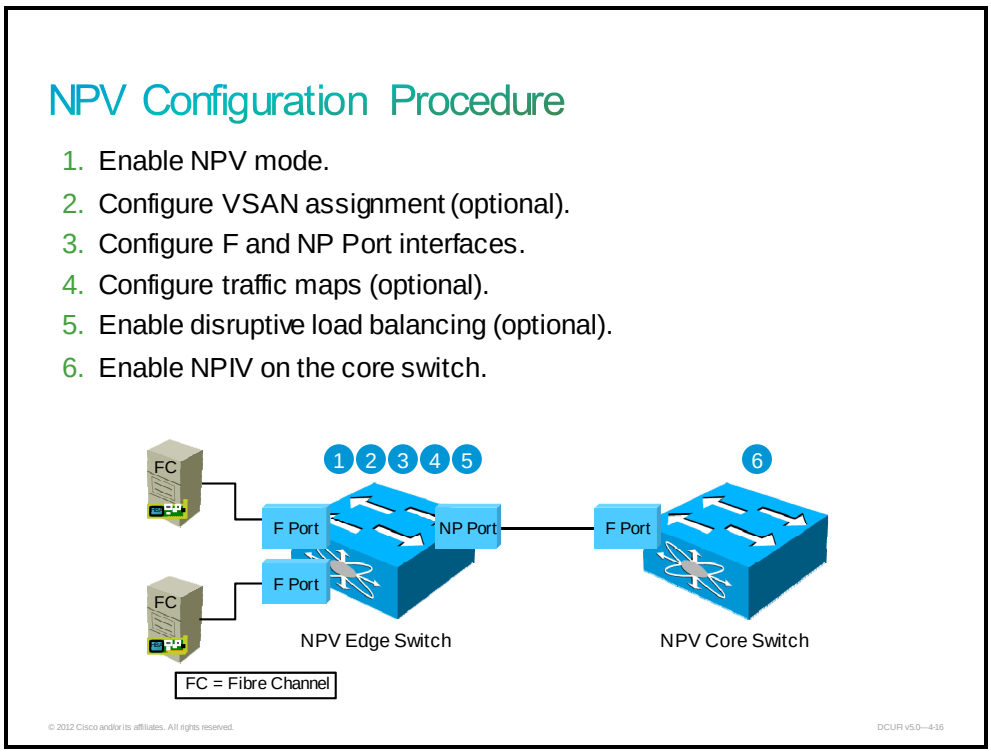
The NPV mode has these features:

- Most Fibre Channel services are switched off.
- The NPV-enabled switch now becomes a multiplexor.
- The NPV-enabled switch does not use a domain ID, which is the reason for the domain ID limitation.
- There are a smaller number of switches to manage.
- It eliminates the need for server administrators to manage the SAN.

https://t.me/learningnets

NPV Configuration

This topic describes how to configure NPV on the Cisco Nexus and MDS Switches.



Follow this procedure to configure the NPV feature on Cisco Nexus 5000 and 5500 Platform switches:

1. Enable NPV mode. Enabling NPV is provided by using the **npv enable** command. This command is issued globally. It will require the switch to erase the system configuration and reload the switch.

2. Create VSANs and assign interfaces to VSANs. If you do not explicitly configure any VSAN assignments, all ports are placed in VSAN 1.

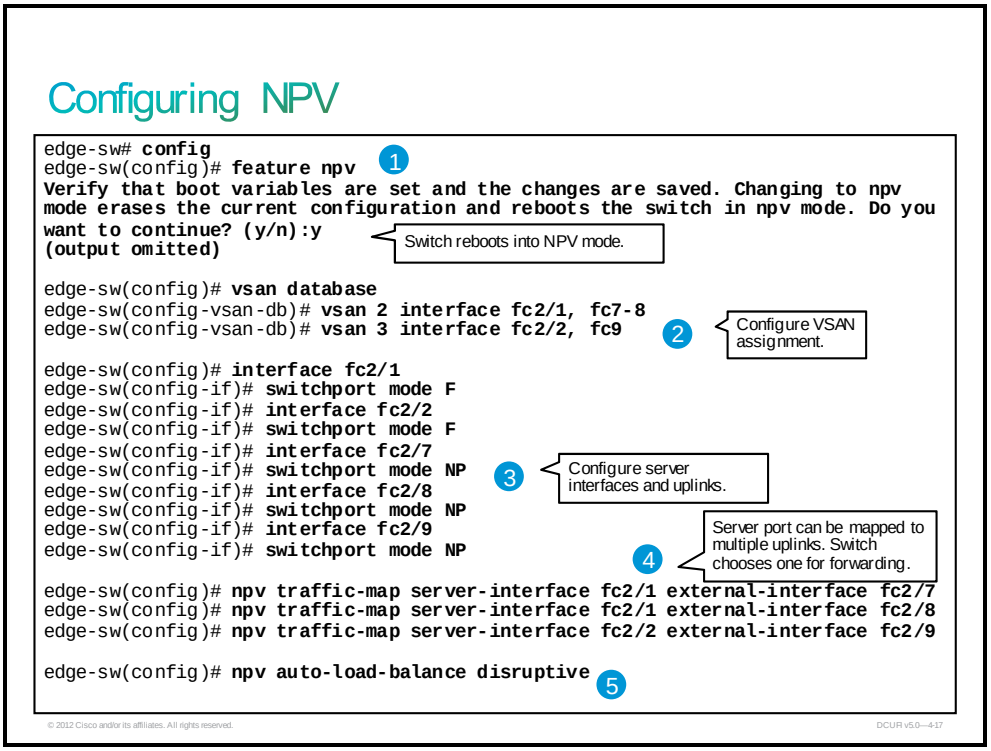
3. Configure F and NP Port interfaces. The F mode is the default mode. At least the external interfaces have to be configured for the NP Port mode.

4. Optionally, configure traffic maps. This step pins the server interfaces to the external interfaces. NPV supports automatic selection of NP Port uplinks. When a server interface is brought up, the NP Port uplink interface with the minimum load is selected from the available NP Port uplinks in the same VSAN as the server interface

5. Optionally, enable disruptive load balancing.

6. Enable NPIV on the core switch. This step was discussed earlier in the lesson.

https://t.me/learningnets



The configuration example starts with the enabling of the NPV mode. The configuration is erased and the switch reboots. You create two VSANs with IDs 2 and 3 and assign Fibre Channel interfaces to them.

Note If you are using unified ports on the Cisco Nexus 5500 Platform switch, the switch will need to reboot twice—once to load NPV mode, and for the second time to redefine unified ports as Fibre Channel interfaces.

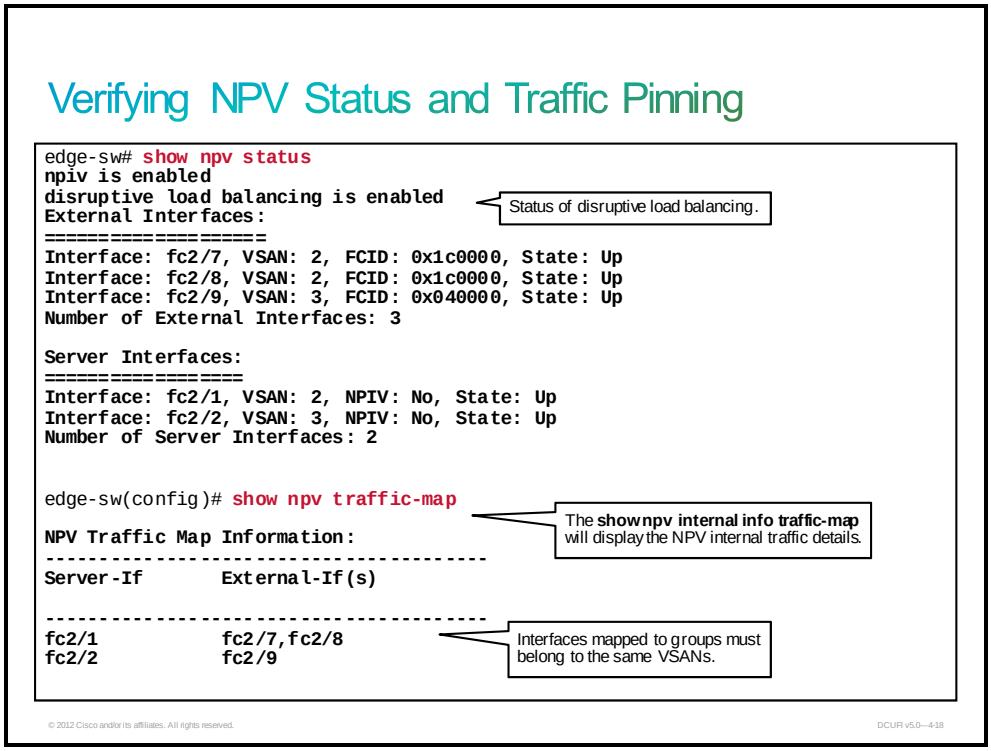
In step 4, you declare the Fibre Channel ports as either server or external interfaces by configuring their mode.

Step 5 shows the static traffic maps. The first map binds the two server interfaces from VSAN 2 to the uplink in VSAN 2. The second map binds interfaces in VSAN 3.

Finally, you enable disruptive auto load balancing to automatically load share traffic over newly added uplinks.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets



You can use the **show npv status** command to display the current NPV status, including the state of disruptive load balancing and all NPV interfaces.

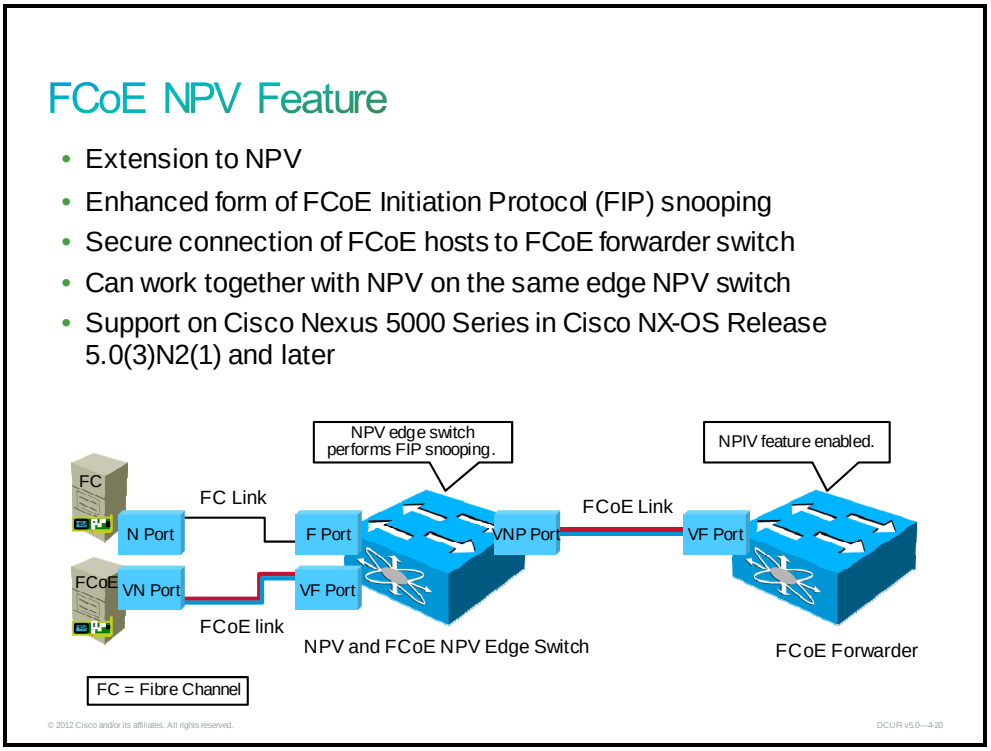
The **show npv traffic-map** command is used to display NPV traffic maps. You can examine the results that are shown in the figure to verify that the interfaces that are mapped to one another belong to the same VSAN.

The **show npv status** and **show npv traffic-map** commands are available only when the switch is in NPV mode.

https://t.me/learningnets

FCoE NPV Feature

This topic describes the FCoE NPV feature.



Beginning with Cisco NX-OS Software Release 5.0(3)N2(1), FCoE NPV is supported on the Cisco Nexus 5000 Series Switches. The FCoE NPV feature is an enhanced form of FCoE Initialization Protocol (FIP) snooping that provides a secure method to connect FCoE-capable hosts to an FCoE-capable FCoE forwarder switch. The FCoE NPV feature provides the following benefits:

- FCoE NPV does not have the management and troubleshooting issues that are inherent to managing hosts remotely at the FCoE forwarder.
- FCoE NPV implements FIP snooping as an extension to the NPV function while retaining the traffic-engineering, VSAN-management, administration, and trouble-shooting aspects of NPV.
- FCoE NPV and NPV together allow communication through Fibre Channel and FCoE ports at the same time. This provides a smooth transition when moving from Fibre Channel to FCoE topologies.

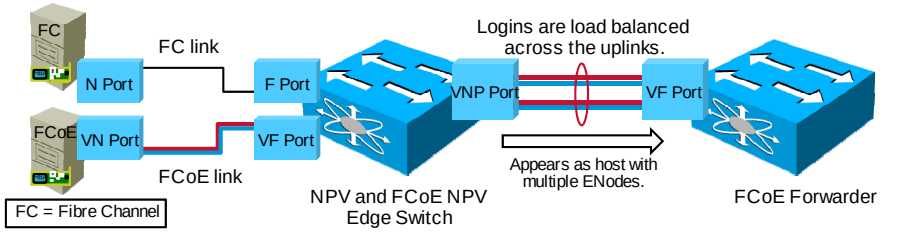
Note	When you enable FCoE NPV, the switch does *not* reload. This feature requires an extra license.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

FCoE NPV Edge Switch

- Performs proxy functions to load balance logins from the hosts evenly across the available FCoE forwarder uplink ports.
- VSAN to VLAN mapping:
  - Each server VSAN needs a dedicated VLAN.
  - VLAN carries FIP and FCoE traffic from the mapped VSAN.
  - VLAN-VSAN mapping must be consistent in the entire fabric.
  - Cisco Nexus 5000 Series Switches support 32 VSANs.
- On the VNP port, an FCoE NPV bridge emulates an FCoE-capable host with multiple ENodes, each with a unique end node MAC address.



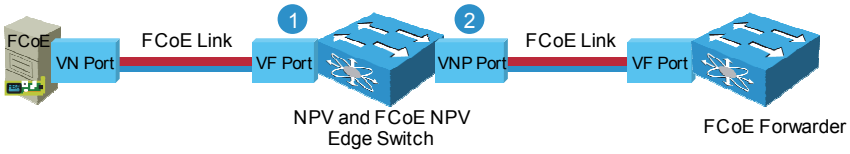
© 2012 Cisco and/or its affiliates. All rights reserved.

The figure shows the FCoE NPV bridge connecting hosts and FCoE forwarders. From a control plane perspective, FCoE NPV performs proxy functions toward the FCoE forwarder and the hosts in order to load balance logins from the hosts evenly across the available FCoE forwarder uplink ports. An FCoE NPV bridge is VSAN-aware and capable of assigning VSANs to the hosts.

https://t.me/learningnets

FCoE NPV Interfaces

1. VF ports:
   - Must be bound to a VLAN trunk Ethernet interface or a port channel.
   - FCoE VLAN must not be the native VLAN.
   - A port VSAN must be configured for the VF Port.
2. VNP ports:
   - Must be point-to-point links.
   - Individual Ethernet interfaces or members of an Ethernet port channel.
   - For each FCF, a vFC interface must be bound to the Ethernet interface
   - Binding to MAC address is not supported.
   - By default, the VNP port is enabled in trunk mode and carries multiple VSANs.
   - Have the STP automatically disabled in the FCoE VLAN.



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

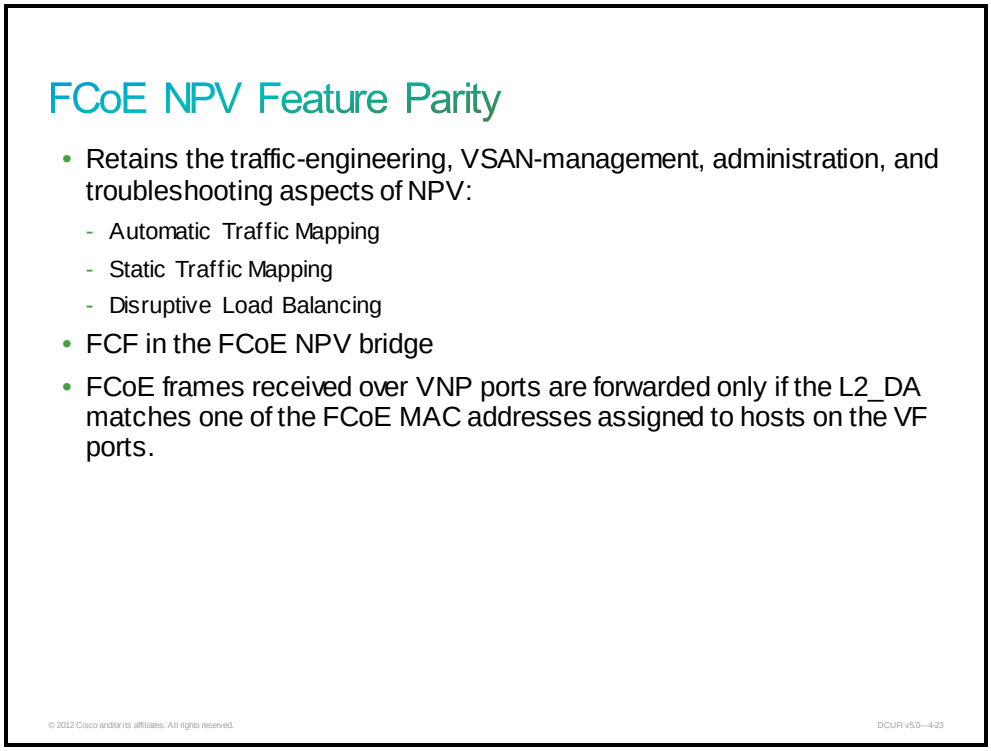
For each host directly connected over Ethernet interfaces on the FCoE NPV bridge, a virtual Fibre Channel (vFC) interface must be created and bound to the Ethernet interface. By default, the vFC interface is configured in the F mode (virtual fabric port [VF Port]).

The VF Port must be configured with the following parameters:

- A VF Port must be bound to a VLAN trunk Ethernet interface or a port-channel interface. The FCoE VLAN must not be configured as the native VLAN on the Ethernet interface.
- A port VSAN must be configured for the VF Port.
- The administrative state must be up.

Connectivity from an FCoE NPV bridge to the FCF is only supported over point-to-point links. These links can be individual Ethernet interfaces or members of an Ethernet port channel interface. For each FCoE forwarder-connected Ethernet interface, a vFC interface must be created and bound to the Ethernet interface. These vFC interfaces must be configured as virtual proxy N ports (VNPs). On the VNP port, an FCoE NPV bridge emulates an FCoE-capable host with multiple ENodes, each with a unique ENode MAC address. A VNP port interface binding to a MAC address is not supported. By default, the VNP port is enabled in trunk mode. Multiple VSANs can be configured on the VNP port. The FCoE VLANs that correspond to the VNP port VSANs must be configured on the bound Ethernet interface. The Spanning Tree Protocol (STP) is automatically disabled in the FCoE VLAN on the interfaces that the VNP ports are bound to.

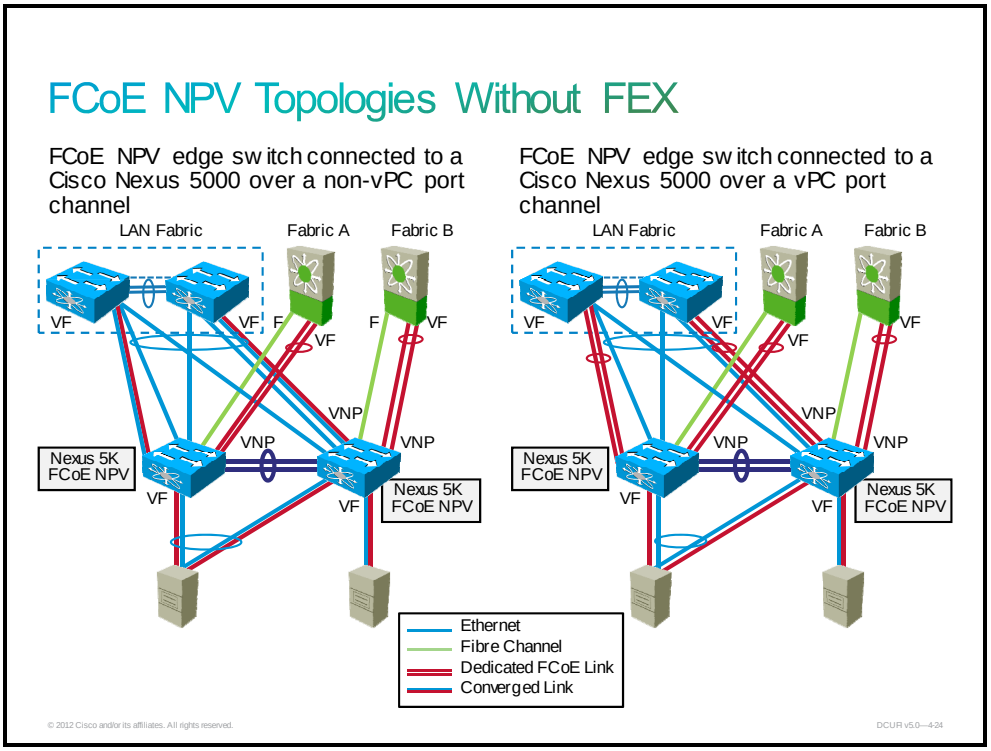
https://t.me/learningnets



The following NPV features apply for the FCoE NPV feature:

- Automatic Traffic Mapping
- Static Traffic Mapping
- Disruptive Load Balancing
- FCoE Forwarding in the FCoE NPV bridge
- FCoE frames received over VNP ports are forwarded only if the Layer 2 destination address (L2_DA) matches one of the FCoE MAC addresses that are assigned to hosts on the VF ports; otherwise, they are discarded.

https://t.me/learningnets



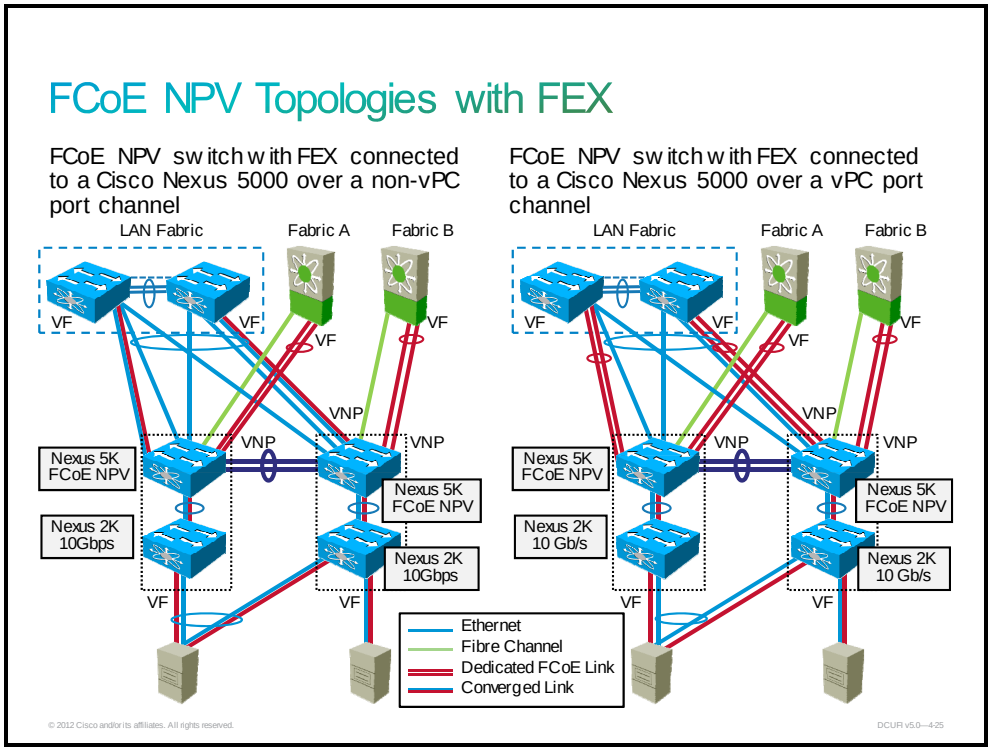
The figure illustrates two supported topologies that do not use the fabric extenders (FEXs). The left topology depicts a Cisco Nexus 5000 Series switch acting as an FCoE NPV device that is connected to a Cisco Nexus 5000 Series switch over a non-vPC port channel. The topology on the right illustrates a similar case with a vPC port channel.

When VNP ports are configured as vPC topologies between an FCoE NPV bridge and an FCF, the following limitations apply:

- vPC spanning multiple FCFs in the same SAN fabric is not supported.
- For LAN traffic, dedicated links must be used for FCoE VLANs between the FCoE NPV bridge and the FCF that is connected over a vPC.
- FCoE VLANs must not be configured on the interswitch vPC interfaces.
- VF Port binding to a vPC member port is not supported for an interswitch vPC.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets



These topologies illustrate two solutions that are based on fabric extenders with 10-Gb/s interfaces, and thus able to support FCoE. The scenarios are equivalent to the topologies that are shown in the previous page. The figure shows a non-vPC solution (on the left) and a vPC solution (on the right).

Unsupported topologies are not shown here, but include these topologies:

- Cisco Nexus 5010 Switch or Nexus 5020 Switch as an FCF connecting to the same FCoE NPV bridge over multiple VF Ports
- A 10-GB Fabric Extender connecting to the same FCoE NPV bridge over multiple VF Ports
- Cisco Nexus 5000 Series Switch as an FCoE NPV bridge connecting to a FIP snooping bridge or another FCoE NPV switch
- VF Port trunk to hosts in FCoE NPV mode

https://t.me/learningnets

FCoE NPV Interoperability and Port Limits

Cisco Nexus 5000 switch interoperates with these FCoE-capable switches:

- Cisco MDS 9000 Series Multilayer Switches enabled to perform FCoE forwarding functions (FCoE NPV and VE)
- Cisco Nexus 7000 Series Switches enabled to perform FCoE forwarding functions (FCoE NPV and VE)
- Cisco Nexus 4000 Series Switches enabled for FIP snooping without FCoE NPV functionality

Interface type	Nexus 5000	Nexus 5500	Nexus 2000 Series Fabric Extenders (10 Gb/s Interfaces)
VNP port bound to Ethernet interface	4 VNP ports	16 VNP ports	Not Supported
VNP port bound to Ethernet port channel interface	2 VNP ports	16 VNP ports	Not Supported
VNP port bound to virtual Ethernet (vEth) interface	Not Supported	Not Supported	Not Supported

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

The Cisco Nexus 5000 and 5500 Platform switches interoperate with the following FCoE-capable switches:

- Cisco MDS 9000 Series Multilayer Switches that are enabled to perform FCoE forwarding functions (FCoE NPV and VE)
- Cisco Nexus 7000 Series Switches that are enabled to perform FCoE forwarding functions (FCoE NPV and VE)
- Cisco Nexus 4000 Series Switches that are enabled for FIP snooping without FCoE NPV functionality

The table lists the FCoE configuration limits over Ethernet, Ethernet port channel, and virtual Ethernet interfaces. Furthermore, the FCoE NPV feature has the following upgrade and downgrade guidelines and limitations:

- You cannot perform a Cisco In-Service Software Downgrade (ISSD) to Cisco NX-OS Software Release 5.0(3)N1(1) or an earlier release if FCoE NPV is enabled and if VNP ports are configured.
- A warning is displayed if an ISSD is performed to Cisco NX-OS Release 5.0(3)N1(1) or an earlier release when FCoE NPV is enabled but VNP ports are not configured.
- Before performing a Cisco In-Service Software Upgrade (ISSU) on an FCoE NPV bridge, use the **disable-fka** command to disable the timeout value check (FKA check) on the core switch.

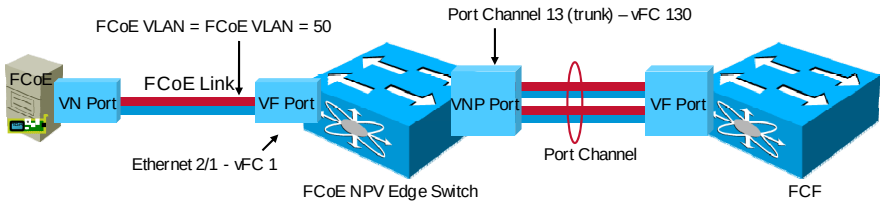
https://t.me/learningnets

FCoE NPV Configuration

This topic describes how to configure NPV on the Cisco Nexus switches.

FCoE NPV Configuration Procedure

1. Enable FCoE NPV mode using one of two methods:
   - Enable FCoE, and then enable NPV.
   - Enable FCoE NPV.
2. Enable default FCoE QoS policy (best practice).
3. Configure VSAN-to-VLAN mapping.
4. Configure server interfaces.
5. Configure uplinks.
6. Configure traffic maps (optional).
7. Enable disruptive load balancing (optional).



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

Follow this procedure to implement FCoE NPV:

1. Enable FCoE NPV mode using one of two methods. The methods are described in the next page.
2. Enable default FCoE QoS policy. This step ensures lossless forwarding for FCoE traffic and is considered a best practice.
3. Configure VSAN-to-VLAN mapping.
4. Configure server interfaces. These interfaces are vFC interfaces.
5. Configure uplinks. These interfaces are vFC interfaces.
6. Optionally, configure static traffic maps.
7. Optionally, enable disruptive load balancing for immediate rebalancing of traffic over newly added uplinks.

https://t.me/learningnets

Methods of Enabling FCoE NPV

Method	Enable FCoE and Then Enable NPV	Enable FCoE NPV
Commands	**feature fcoe** **feature npv**	**feature fcoe-npv**
Trigger of write erase and reload	Yes	No
Licensing	Storage Protocols Services Package (FC_FEATURES_PKG)	This method requires a separate license package (FCOE_NPV_PKG). This license is also included in the Storage Protocols Services Package license.
Description	When FCoE is enabled, the default mode of operation is Fibre Channel switching, and when you enable NPV, the mode changes to NPV mode.	FCoE NPV provides NPV feature just for FCoE traffic.

In both methods, after FCoE NPV mode has been configured on a switch, the FCoE feature cannot be enabled. A warning is displayed to reload the system first in order to enable FCoE.

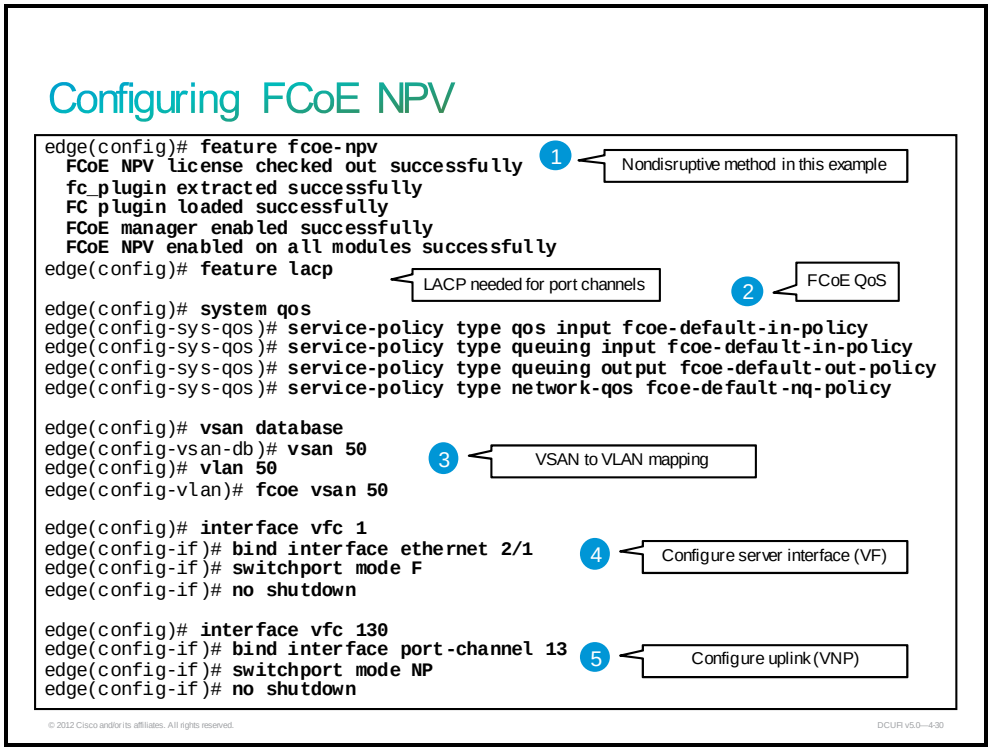
© 2012 Cisco and/or its affiliates. All rights reserved.

You can enable FCoE NPV by choosing one of the following methods:

- **Enable FCoE, and then enable NPV**: This method requires that you enable FCoE first using the **feature fcoe** command, and then you enable NPV by using the **feature npv** command. When FCoE is enabled, the default mode of operation is Fibre Channel switching, and when you enable NPV, the mode changes to NPV mode. Switching to NPV mode automatically performs a write erase and reloads the system. After the reload, the system comes up in NPV mode. To exit NPV mode and return to Fibre Channel switching mode, enter the **no feature npv** command. Exiting NPV mode also triggers a write erase and a switch reload. This method requires the Storage Protocols Services Package (FC_FEATURES_PKG) license.
- **Enable FCoE NPV**: When you enable FCoE NPV using the **feature fcoe-npv** command, the mode changes to NPV. When you use this method, a write erase and reload does not occur. This method requires a separate license package (the FCoE NPV Package [FCOE_NPV_PKG] license). This license is also included in the Storage Protocols Services Package license.

When FCoE NPV mode is configured on a switch, the FCoE feature cannot be enabled. A warning is displayed to reload the system first to enable FCoE.

https://t.me/learningnets



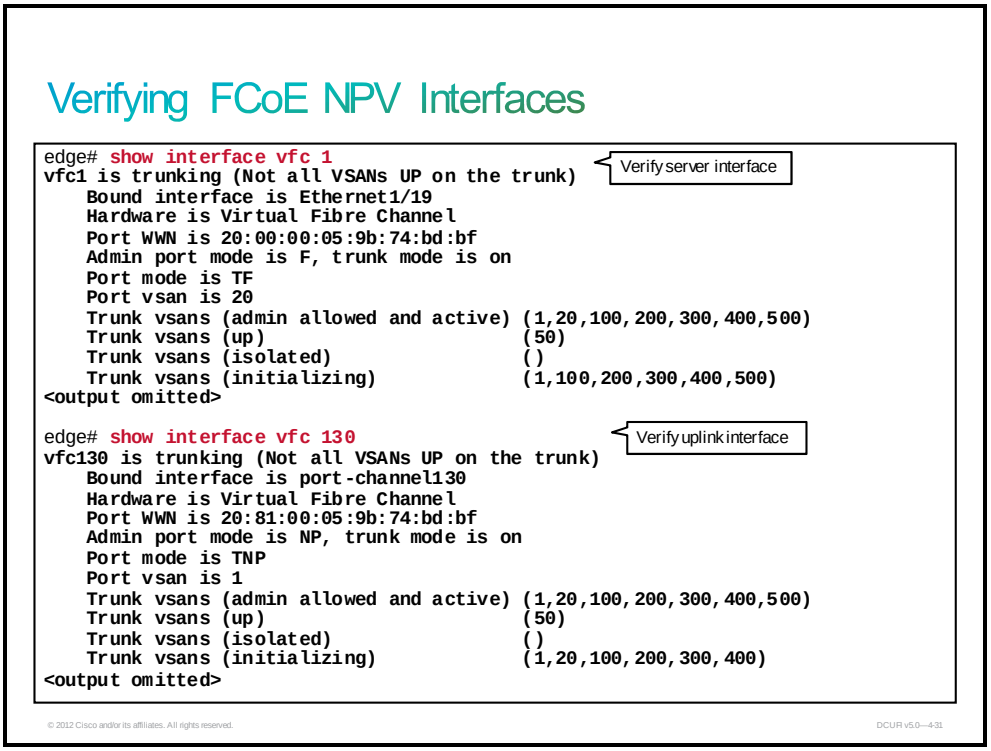
This example illustrates how to implement the first five configuration steps. The optional steps of static traffic maps and disruptive auto load balancing (steps 6 and 7) were shown in the previous topic.

The configuration starts by enabling the fcoe-npv feature, which allows the switch to retain the running configuration and stay operational. You enable the QoS policy with default FCoE parameters for QoS input, input and output queuing, and network forwarding.

In this scenario, you configure the VSAN 50 and map it to VLAN 50. The interface vFC 1 is the server interface and is bound to the Ethernet 2/1. The port Ethernet 2/1 is configured as a trunk port, and carries VLAN 50 as the FCoE VLAN (not shown in the example).

The interface vFC 130 is the external interface and is bound to the port channel 13. The port channel 13 is configured as a trunk port, and carries, among others, VLAN 50 as the FCoE VLAN.

https://t.me/learningnets



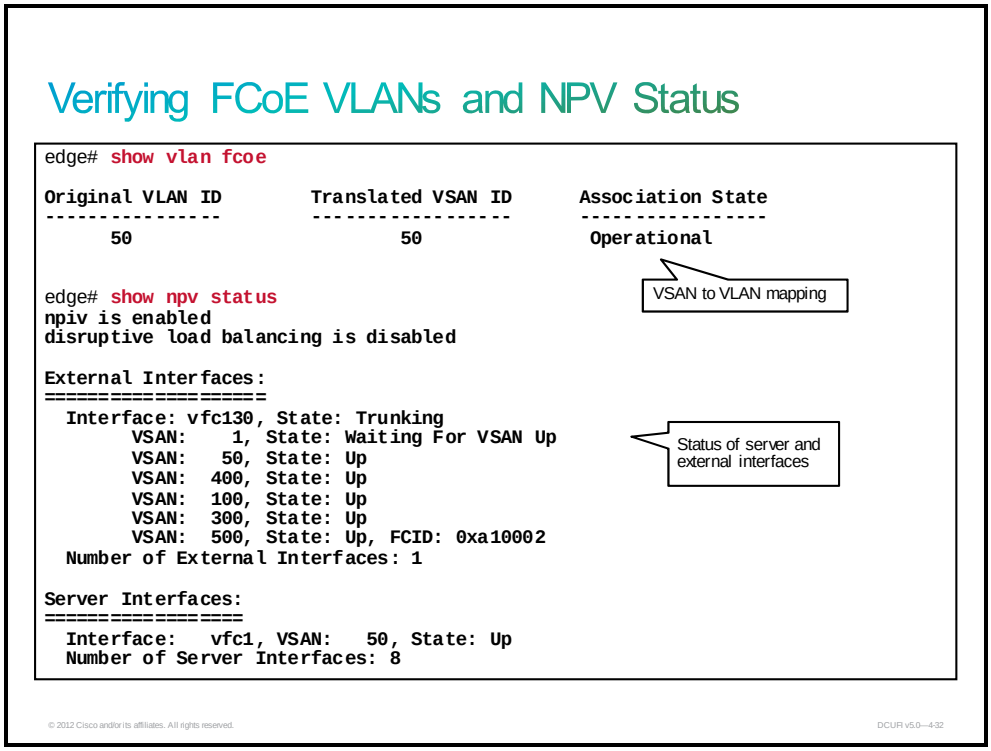
You may begin the FCoE NPV by verifying the server and external interfaces using the **show interface** command.

The interface vFC 1 is in TF mode (trunking fabric), and carries among others, VLAN 50. VLAN 50 is mapped to VSAN 50 in this scenario.

The interface vFC 130 is in TNP mode (trunking NP Port), and carries among others, VLAN 50.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets



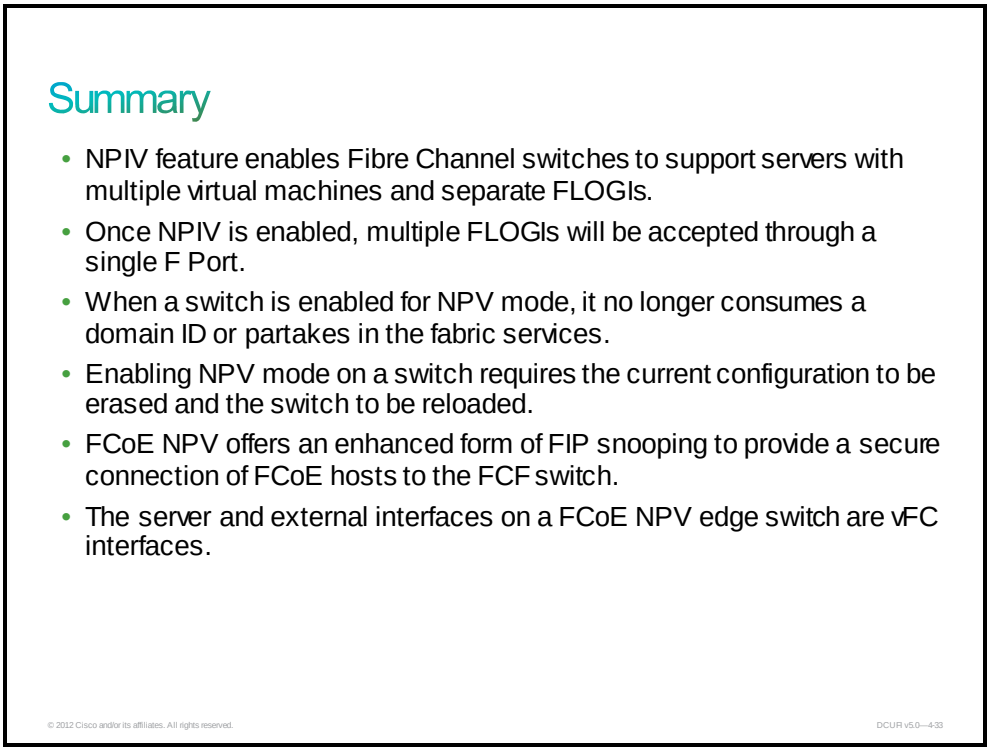
In addition, you can view all VLANs that have been enabled for FCoE and the VSAN IDs to which they are mapped. The **show vlan fcoe** command offers the relevant information.

The **show npv status** command is used to display a summary of the enabled features, the external ports, and the server interfaces.

https://t.me/learningnets

Summary

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



References

For additional information, refer to these resources:

- To learn more about configuring SAN switching on the Cisco Nexus 5000 and 5500 switches, refer to *Cisco Nexus 5000 Series NX-OS SAN Switching Configuration Guide, Release 5.1(3)N1(1)* at: http://www.cisco.com/en/US/docs/switches/datacenter/nexus5000/sw/san_switching/513_n1_1/b_Cisco_n5k_nxos_sanswitching_config_guide_rel513_n1_1.html

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Lesson 7

Using SAN Management Tools

Overview

In this lesson, you will be introduced to both management tools, Cisco Prime Data Center Network Manager (DCNM) for SAN and Cisco Device Manager, that are responsible for managing the SAN part of the data center network. Using those management tools makes it possible to configure and monitor the SAN using the GUI.

Objectives

Upon completing this lesson, you will be able to identify the SAN management tools that can be used to manage the Cisco Nexus 5000 and 7000 Series Switches and Cisco MDS 9000 Series Multilayer Switches. This ability includes being able to meet these objectives:

- Explain the purpose of Cisco Prime DCNM for SAN
- Explain the purpose of Cisco Device Manager

https://t.me/learningnets

Cisco Prime DCNM for SAN

This topic explains the purpose of Cisco Prime Data Center Network Manager (DCNM) for SAN.

Overview of Cisco Prime DCNM for SAN

- Part of the new converged management solution
- Recommended solution for managing data center networks
- Maintains all the features of Cisco Fabric Manager plus additional new features
- Needs to be deployed through appropriate licence
- Upgrade from Cisco Fabric Manager 5.x or Cisco Data Center Network Manager 5.x is transparent (SASU contract)

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

The Cisco Prime Data Center Network Management (DCNM) for SAN tool offers administrators a new converged approach to manage and monitor devices in the fabric through a GUI interface.

Cisco Prime DCNM for SAN replaces the former recommended tool for monitoring SAN, Cisco Fabric Manager. It maintains all the features of Cisco Fabric Manager and includes new features

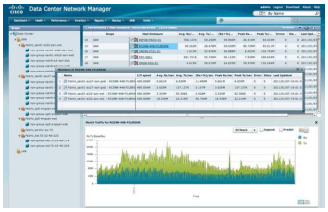
Cisco Prime DCNM is available with multiple licensing options. It can be licensed for SAN and LAN environments separately or together. Licenses are hosted on the management server. They are no longer hosted on the switch.

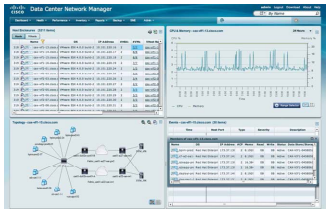
The existing Cisco Fabric Manager Release 5.x and Cisco DCMN Release 5.x remain applicable and can be transparently upgraded to the new converged Cisco DCNM. The only prerequisite for this upgrade would be an active Cisco Software Application Support plus Upgrades (SASU) contract.

https://t.me/learningnets

Cisco Prime DCNM for SAN Main Features

- Real-time fabric and network health monitoring
- VM-aware automated discovery and VM-path analysis
- VM-aware performance monitoring
- Detailed fabric topology views of datacenter infrastructure
- Comprehensive Fibre Channel over Ethernet (FCoE) management, including provisioning, discovery, and operation monitoring
- Custom reporting





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

Cisco Prime DCNM for SAN is a comprehensive Fibre Channel over Ethernet (FCoE) management, provisioning, discovery, and operation monitoring tool. It offers a real-time fabric and network health summary with a detailed view of individual network components. SAN administrators have the ability to proactively monitor and react to specific events based on their severity.

Cisco Prime DCNM for SAN also incorporates a discovery and analysis method, which provides a virtual path view through the physical fabric out to the storage array and to the data store. This capability is useful for viewing performance for every switch hop to a particular VMware ESX server and virtual machine (VM).

The VM-aware topology view shows all dependencies, from the VM out to the physical host, to the switch, and out to storage, with one-click access to attributes. Bottlenecks in VM and VMware ESX performance are easy to identify because VMware views provide end-to-end path information across multiple fabric clouds.

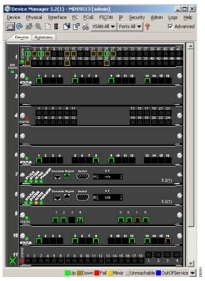
Configuration changes are easy to maintain because they can be prevalidated, which reduces the possibility of misconfiguration from the administrator side.

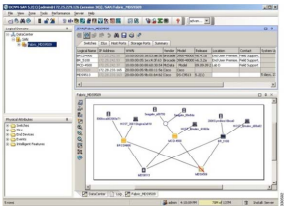
For reporting, Cisco Prime DCMN for SAN also includes custom reporting produced from predefined templates (such as Inactive Path, Zone Discrepancy, Switch Health, and so on) across all fabrics or for an individual fabric.

https://t.me/learningnets

Cisco Prime DCNM for SAN Basic Components

- DCNM-SAN Server
- DCNM-SAN Client
- DCNM-SAN Web Client
- Device manager
- Performance Manager
- Cisco Traffic Analyzer
- Network Monitoring
- Performance Monitoring





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

The Cisco Prime DCNM for SAN consists of the following basic components:

- **DCNM-SAN Server: Cisco Prime** DCNM-SAN Server is a platform for advanced MDS monitoring, troubleshooting, and configuration capabilities. It can be installed on any standalone computer that is intended to provide centralized Cisco MDS management services and performance monitoring. A computer that is configured as a Cisco DCNM-SAN Server can monitor multiple Fibre Channel SAN fabrics, and up to 16 clients (by default) can connect to a single Cisco DCNM-SAN Server concurrently.
- **DCNM-SAN Client:** Cisco Prime DCNM-SAN Client is a Java and Simple Network Management Protocol (SNMP)-based network fabric and device management tool, which displays real-time views of your network fabric. It offers complete configuration and status monitoring capabilities along with advanced troubleshooting tools.
- **DCNM-SAN Web Client:** With Cisco Prime DCNM-SAN Web Client you can monitor Cisco MDS, Cisco Nexus 7000 Series and Cisco Nexus 5000 Series switch events, performance, and inventory from a remote location using a web browser.
- **Cisco Device Manager:** The Cisco Device Manager provides a graphical view of the Cisco MDS 9000 Family chassis or Cisco Nexus 5000 Series switch chassis. There are two available views:
  - Device View, which is used to perform switch-level configuration and graphical representation of the switch configuration
  - Summary View, which monitors all ports on the switch as well as Fibre Channel and IP neighbor devices

https://t.me/learningnets

- **Cisco Performance Manager:** Cisco Performance Manager gathers information from across the fabric, and based on this captured data, presents detailed traffic analysis in the form of various graphs and charts. Performance Manager consists of three main operational stages:
  - Definition (setting up flows in the switches)
  - Collection (collect information on particular fabric)
  - Presentation (Results are presented on a web page through DCNM-SAN Web Server.)
- **Cisco Traffic Analyzer:** Cisco Traffic Analyzer is a tool that performs real-time analysis that is based on data captured from one or more Port Analyzer Adapters (PAAs).
- **Network Monitoring:** Network monitoring uses the SNMP protocol to perform SAN discovery and topology mapping. Results are presented in a customizable map with multiple viewing options (device view, summary view, and operation view).
- **Performance monitoring:** The performance monitoring tool at monitors ports (at configurable intervals) on the Cisco MDS 9000 Family switches in real time. Monitoring results, which are useful for further examination and troubleshooting, are shown in the form of graphs or charts.

https://t.me/learningnets

Cisco Prime DCNM for SAN Scope

Cisco DCNM for SAN supports a wide range of Cisco hardware platforms:

- Cisco MDS 9500 Series Switches
- Cisco MDS 9200 Series Switches
- Cisco MDS 9100 Series Switches
- Cisco Nexus 7000 Series Switches
- Cisco Nexus 5000 Series Switches
- Cisco Nexus 3000 Series Switches
- Cisco UCS 6100/6200 Series Fabric Interconnects

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

Cisco Prime DCNM for SAN supports a wide range of Cisco hardware platforms that are deployed in the data center including these products:

- Cisco MDS 9500 Series Multilayer Switches
- Cisco MDS 9200 Series Multilayer Switches
- Cisco MDS 9100 Series Multilayer Fabric Switches
- Cisco Nexus 7000 Series Switches
- Cisco Nexus 5000 Series Switches
- Cisco Nexus 3000 Series Switches
- Cisco Unified Computing System (UCS) 6100 and 6200 Series Fabric Interconnects

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

Cisco Prime DCNM for SAN Editions

Essentials Edition	Advanced Edition
Summary and host dashboards	Federation and VSAN scoping
Automated fabric discovery	VM-aware monitoring
Template reports	Multiple Fabrics performance monitoring
Real-time performance snapshot	Performance forecasting
Provisioning: Fibre Channel and FCoE provisioning wizards	Historical performance trending
Configuration management: Switch and fabric configuration using provisioning wizards	Event forwarding
Web services APIs	
Fabric topology views	

- Advanced edition contains all the features available in the Essentials Edition plus features listed in Advanced Edition

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

Cisco Prime DCNM for SAN is available in two editions:

- Essentials edition (free of charge)
- Advanced edition (license that is based on the number and type of switches)

Cisco Prime DCNM for SAN Essentials Edition is available for free, and it is included in every Cisco MDS 9000 Family hardware purchase. It is publicly available to download from Cisco.com. Essentials edition includes features like automated fabric discovery, real-time performance snapshots, switch and fabric configuration using the provisioning wizard, and so on.

The Cisco Prime DCNM for SAN Advanced Edition contains all the features that are available in the Essentials Edition and adds a few more capabilities like performance monitoring and trending, event forwarding, VM-aware path analysis, automated fabric discovery and so on. The Cisco Prime DCNM license for SAN Advanced Edition is deployed on the Cisco DCNM Server.

https://t.me/learningnets

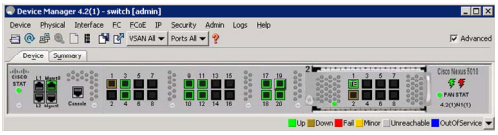
Cisco Device Manager

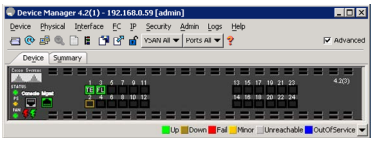
This topic explains the purpose of the Cisco Device Manager and how to install the software.

Cisco Device Manager

For managing a single switch or director

- Full switch configuration capabilities
  - VSANs
  - PortChannels
  - Interfaces
  - RMON alerts, and so on
- Real-time statistics monitoring
- Device and summary views
- License Manager
- View license information
- Install and remove license key files





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

The next tool the SAN administrator would use to manage the fabric is the Cisco Device Manager. Cisco Device Manager is used to manage and configure individual switches in the fabric. This topic discusses Cisco Device Manager.

Cisco Device Manager offers an alternative to the CLI for switch configuration. Cisco Device Manager allows you to configure values on a single switch through the Device View. In the Device View, you can open configuration and status dialog boxes by clicking various parts of the device, such as switch ports and modules.

Some of the functions that you can configure with Cisco Device Manager are virtual storage area networks (VSANs), SAN port channels, Remote Monitoring (RMON) alerts, general switch event filters, SNMP users and roles, and Fibre Channel services.

The Device tab graphically depicts the switch enclosure. The status of the switch modules, ports, fans, and power supplies can be determined at a glance. The Cisco Device Manager also provides an intuitive way to drill down for configuration and to examine specific statistics. It can also be used to view the different log information details.

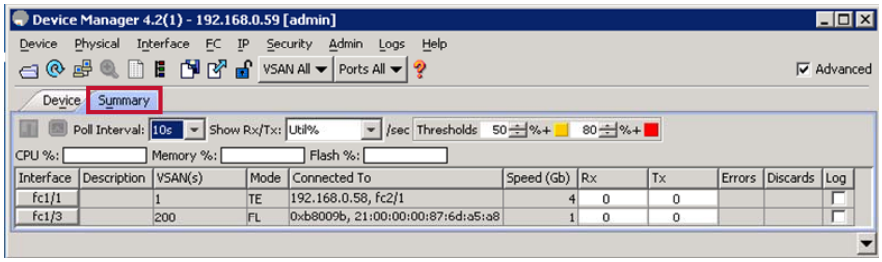
You can use the Cisco License Manager to view license information. This information includes the name of the feature package, type of license that is installed, and the expiration date. It can also be used to install or remove license key files.

https://t.me/learningnets

Cisco Device Manager (Cont.)

Displays a summary of interfaces on a single switch or director

- Configurable poll interval and thresholds
- CPU, memory, and flash usage
- Interface description, VSAN, port mode, and connectivity
- Rx/Tx utilization, errors, and so on



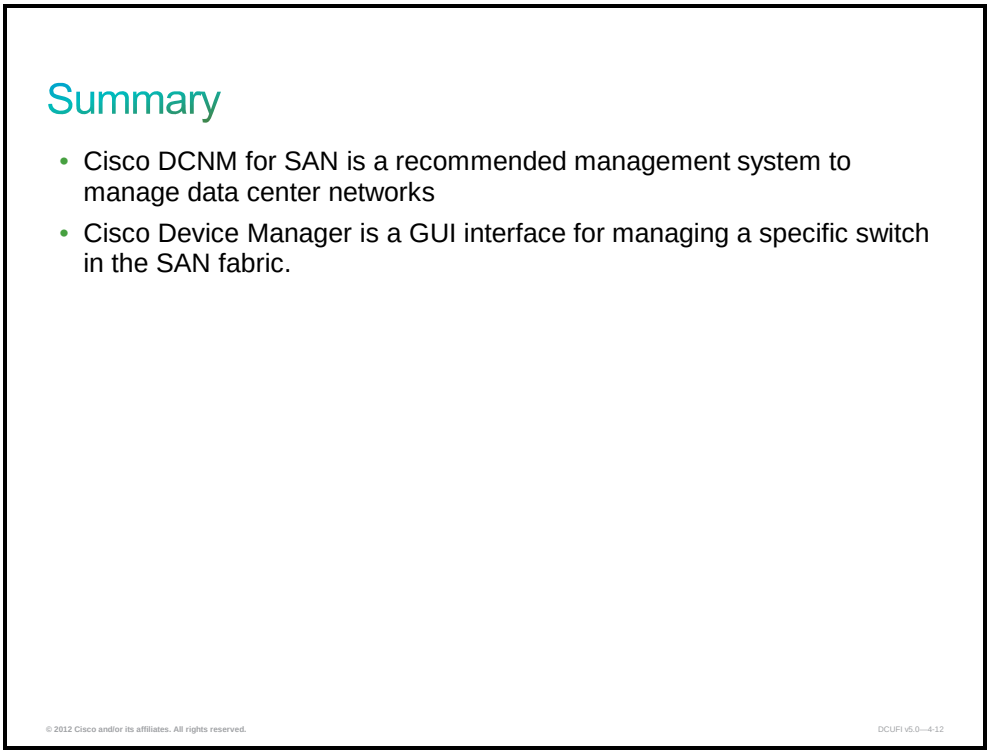
© 2012 Cisco and/or its affiliates. All rights reserved.

- To access the Summary View, click the Summary tab on the Device View main window. The Summary tab displays a summary of interfaces on a single switch.
- Attributes for the ports such as description, VSANs, mode, and the device that the port is connected to are provided.
- You can refresh the screen using the refresh icon.

https://t.me/learningnets

Summary

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

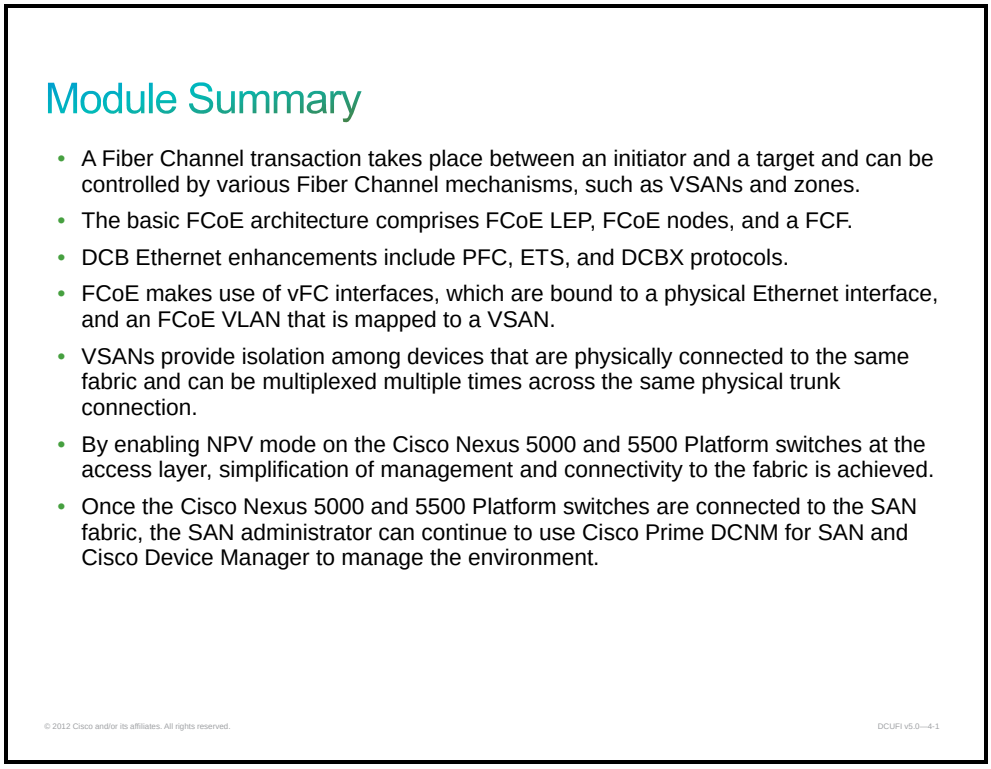


© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Module Summary

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



Traditionally, the SAN environment and the TCP/IP data network environment have been separate entities. The data network has used Ethernet as its Layer 2 protocol, while the SAN has used Fibre Channel. In the Fibre Channel SAN environment, communication takes place between an initiator and a target, where Small Computer Systems Interface (SCSI) commands, status, and block data are exchanged.

Some of the concerns in the data center are the lack of space, the amount of equipment that must be managed, the cabling infrastructure, cooling, and the power consumption of devices that are required to support these different networks.

Fibre Channel over Ethernet (FCoE) is a protocol that is used to consolidate the I/O infrastructure at the access layer. In the access layer, both Fibre Channel and TCP/IP can be carried over the same Ethernet physical wiring. However, TCP/IP and Fibre Channel impose different operational requirements on that wiring. TCP/IP traditionally relies on loss of packets to manage temporary periods of congestion. Fibre Channel is a lossless infrastructure with guaranteed delivery of frames.

To support different types of traffic being carried over a common infrastructure, the devices that are involved must be able to support the Data Center Bridging (DCB) Ethernet enhancements. These enhancements provide priority flow control (PFC), Enhanced Transmission Selection (ETS), and Data Center Bridging Exchange (DCBX) protocol for discovery and exchange of information between peers.

Due to the limitation in the number of domain IDs that will be supported in a SAN fabric, it is often advantageous to enable N-Port Virtualization (NPV) mode on the edge access switch. This mode permits the switch to partake in the fabric as a pass-through device with no domain ID usage. This mode is simple, easy to configure, and provides a seamless way of adding additional switches and devices to the fabric.

https://t.me/learningnets

When the Cisco Nexus 5000 and 5500 Platform switches are connected to a SAN fabric, the SAN administrator can continue to use the same tools to manage all the SAN switches. In a Cisco SAN, these tools are Cisco Prime Data Center Network Manager for SAN and Cisco Device Manager.

This module has discussed and explained the configuration steps that are required for implementing the Fibre Channel Protocol (FCP) and FCoE protocol, the DCB Ethernet enhancements, the NPIV, NPV, and FCoE NPV features.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Module Self-Check

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

Q1) What does the SCSI architecture define the relationship between? (Source: Describing Fibre Channel Protocol)

A) hosts and targets
B) hosts and disks
C) initiators and disks
D) initiators and targets

Q2) What is a Fibre Channel exchange? (Source: Describing Fibre Channel Protocol)

A) a number of words
B) a number of sequences
C) a TCP session
D) a number of frames

Q3) Which process occurs between the N Port of an initiator and the N Port of a target to exchange information about upper layer protocol support? (Source: Describing Fibre Channel Protocol)

A) PLOGI
B) NPV
C) PRLI
D) FLOGI
E) Fibre Channel exchange

Q4) Which Fibre Channel address is used by FSPF to route Fibre Channel frames through a fabric? (Source: Describing Fibre Channel Protocol)

A) FCID
B) nWWN
C) pWWN
D) MAC

Q5) How many domains are available to the fabric? (Source: Understanding Fibre Channel Protocol)

A) 256
B) 255
C) 244
D) 239

Q6) IVR resembles IP routing protocols because it enables connectivity between all endpoints and applications in physically separate locations. (Source: Describing Fibre Channel Protocol)

A) true
B) false

https://t.me/learningnets

Q7) Which two features are features of zones? (Choose two.) (Source: Describing Fibre Channel Protocol)

A) They are defined by physical port or pWWN.
B) Zoning configuration can change frequently.
C) An HBA can belong only to a single zone.
D) They provide routing and namespace separation.
E) They limit unicast traffic, but not multicast or broadcast.

Q8) Which feature should be enabled to support multiple fabric logins from a host? (Source: Understanding Fibre Channel Protocol)

A) NPV
B) NP Port
C) NPIV
D) Zoning

Q9) RSCNs provide a timely indication of which two of these events. (Choose two.) (Source: Understanding Fibre Channel Protocol)

A) servers joining or leaving the fabric
B) name server registration change
C) VSAN configuration change
D) new zone enforcement
E) change in interswitch connections

Q10) Which two components support FCoE? (Choose two.) (Source: Understanding FCoE Protocol)

A) hosts with HBAs
B) all Cisco Nexus 5000 models
C) all Cisco Nexus 2000 Fabric Extenders
D) Cisco Nexus 7000 in all VDCs
E) hosts with CNAs

Q11) Which component is responsible for the encapsulation and de-encapsulation that is necessary to transport FCoE? (Source: Understanding FCoE Protocol)

A) FCoE Logical Endpoint
B) FCoE node
C) Fibre Channel forwarder
D) FCoE forwarder

Q12) How does an FCoE node acquire the MAC address that it will use for FCoE traffic? (Source: Understanding FCoE Protocol)

A) server-provided MAC address
B) local MAC address
C) locally defined MAC address
D) fabric-provided MAC address

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Q13) FIP establishes virtual Fibre Channel links between which two combinations of types of ports? (Choose two.) (Source: Understanding FCoE Protocol)

A) N Ports and F Ports
B) VN Ports and VF Ports
C) VE Ports and VF Ports
D) E Ports and E Ports
E) VE Ports and VE Ports
F) VN Ports and VE Ports

Q14) Which method does a blade switch that is installed between the FCoE host and the FCF switch use to forward FCoE frames in a remote-attached FCoE topology? (Source: Understanding FCoE Protocol)

A) By FIP snooping
B) FIP spoofing
C) MAC address lookup using the burned-in MAC address of the host
D) MAC address lookup, where the FPMA includes the FC-MAP that uniquely identifies the CNA
E) MAC address lookup, where the FPMA includes the FCID that uniquely identifies the CNA

Q15) Which benefit results from having different FIP and FCoE EtherType IDs? (Source: Understanding FCoE Protocol)

A) The FCF switch can process FIP frames while transparently forwarding FCoE frames based on the MAC address.
B) The FCF switch can process FCoE frames while transparently forwarding FIP frames based on the MAC address.
C) Blade switches in remote-attached topologies can process FIP frames while transparently forwarding FCoE frames based on the MAC address.
D) Blade switches in remote-attached topologies can process FCoE frames while transparently forwarding FIP frames based on the MAC address.

Q16) Which specification of Data Center Bridging Ethernet enhancements do Cisco Nexus switches implement? (Source: Identifying DCB Enhancements)

A) Cisco proprietary
B) IEEE 802.1
C) www.T11.org
D) ANSI

Q17) Priority groups are associated with which Data Center Bridging Ethernet enhancement feature? (Source: Identifying Data Center Bridging Enhancements)

A) Priority Flow Control
B) Enhanced Transmission Selection
C) End-to-End Congestion Management
D) Layer 2 Multipathing

Q18) Which Data Center Bridging Ethernet enhancement feature supports pause on a virtual channel? (Source: Identifying Data Center Bridging Enhancements)

A) Enhanced Transmission Selection
B) Link Level Flow Control
C) Layer 2 Multipathing
D) Priority Flow Control

https://t.me/learningnets

Q19) Which protocol performs discovery of Data Center Bridging capabilities in a peer? (Source: Identifying Data Center Bridging Enhancements)

A) Data Center Bridging Exchange Protocol
B) Data Center Bridging Exchange Control Protocol
C) Cisco Discovery Protocol
D) Link-Layer Discovery Protocol

Q20) Which of the following defines the encapsulation of FLOGI, FDISC, LOGO, and ELP frames? (Source: Configuring FCoE)

A) FIP discovery
B) FCoE virtual link maintenance
C) FCoE virtual link instantiation
D) FCoE discovery

Q21) Which mode should be enabled on the Ethernet interface to support the transport of both Ethernet and FCoE frames? (Source: Configuring FCoE)

A) switchport mode trunk
B) switchport mode access
C) switchport mode fcoe
D) fcoe mode on

Q22) Which message would be sent to an individual virtual link to reset the virtual interface? (Source: Configuring FCoE)

A) FCoE Clear Virtual Link
B) FCoE Virtual Link Clear
C) FIP Virtual Link Clear
D) FIP Clear Virtual Link

Q23) Which license is required on a Cisco Nexus 5000 or 5500 Platform switch to support FCoE? (Source: Configuring FCoE)

A) Enterprise license
B) Storage Protocol Services license
C) Advanced Services license
D) Storage Protocol Advanced Services license

Q24) Which command would be used to associate an interface with a VSAN? (Source: Configuring FCoE)

A) **switch(config)# interface ethernet 1/3**
switch(config-if)# vsan 2
B) **switch(config)# vsan database**
switch(config-vsan-db)# vsan 2 interface ethernet 1/3
C) **switch(config)# vsan database**
switch(config-vsan-db)# vsan 2 interface vfc 1/3
D) **switch(config)# vsan database**
switch(config-vsan-db)# vsan 2 interface vfc 10

Q25) How does a Cisco Nexus switch process bit errors on Fibre Channel interfaces? (Source: Configuring SAN Switching Features)

A) It disables the interface when a bit error occurs.
B) It does not disable the interface until a threshold is reached.
C) You can configure an error threshold that affects the bit error processing.
D) By default, it increases error counters and keeps the interface up.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Q26) Which service ensures that domain IDs are unique throughout a single fabric (or VSAN)? (Source: Configuring SAN Switching Features)

E) Route Bridge
F) FSPF
G) Principle Switch
H) STP

Q27) What is the default switch priority for fcdomains? (Source: Configuring SAN Switching Features)

A) 64
B) 128
C) 192
D) 254

Q28) What is the default load-balancing method for traffic within a VSAN? (Source: Configuring SAN Switching Features)

A) S_ID/D_ID
B) S_ID/D_ID/OX_ID
C) S_MAC/D_MAC
D) S_MAC/D_MAC/Port_ID

Q29) What tagging method is used for VSAN trunks? (Source: Configuring SAN Switching Features)

A) 802.1Q
B) Inter-Switch Link (ISL)
C) Extended ISL (EISL)
D) 802.1Q or ISL

Q30) Which three attributes change as a Fibre Channel frame is forwarded in an IVR scenario? (Choose three.) (Source: Configuring SAN Switching Features)

A) source FCID
B) destination FCID
C) transit VSAN ID
D) current VSAN ID
E) source FCID when IVR NAT is enabled
F) destination FCID when IVR NAT is enabled

Q31) How many physical links can be configured in a single SAN port channel? (Source: Configuring SAN Switching Features)

A) 4
B) 8
C) 12
D) 16

Q32) Which of the following commands displays information about devices that are logged into a switch? (Source: Configuring SAN Switching Features)

A) **show flogi database**
B) **show fcns database**
C) **show mac-table**
D) **show fspf database**

https://t.me/learningnets

Q33) Which method is a valid Cisco Fabric Services distribution method? (Source: Configuring SAN Switching Features)

A) In-band with fallback to out-of-band
B) IPv4 and IPv6 at the same time
C) IPv4 with fallback to Fiber Channel
D) IPv6 with fallback to Fiber Channel

Q34) Which two port types support trunking on an NPV-enabled switch? (Choose two.) (Source: Configuring NPV Mode)

A) TNP
B) TF
C) NP
D) TE
E) F

Q35) Which mode relays the FLOGI and FDISC requests to an upstream Fibre Channel switch? (Source: Configuring NPV Mode)

A) NPV
B) NPIV
C) Switch
D) NPV and NPIV

Q36) What must be enabled on the NPV core switch if a Cisco Nexus 5000 Switch is in NPV mode and acts as an NPV edge switch? (Source: Configuring NPV Mode)

A) NPV
B) NPIV
C) NPV and NPIV
D) trunking

Q37) Which port type connects a Cisco Nexus 5000 switch in NPV mode to an upstream Fibre Channel switch? (Source: Configuring NPV Mode)

A) N Port
B) F Port
C) E Port
D) NP Port

Q38) How does an FCoE NPV edge switch appear to the upstream switch? (Source: Configuring NPV Mode)

A) as an HBA with multiple FC-IDs
B) as an end node with multiple MAC addresses
C) as an end node with multiple FCIDs but one MAC address
D) as a CNA with multiple end nodes, each with a unique end node MAC address
E) as a CNA with multiple end nodes but one MAC address

Q39) Which setting is required when FCoE NPV is implemented? (Source: Configuring NPV Mode)

A) VSAN-to-VLAN mapping
B) FCoE VLAN definition in the storage VDC
C) traffic maps
D) LACP feature

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Q40) Which component of the Cisco Prime DCNM for SAN, at configurable intervals, monitors the ports of the Cisco MDS 9000 Series switches in real time? (Source: Using SAN Management Tools)

A) DCNM-SAN Server
B) Device Manager
C) Performance Manager
D) Performance monitoring

Q41) Which capability is included only in the Advanced Edition of Cisco Prime DCNM for SAN? (Source: Using SAN Management Tools)

A) fabric topology views
B) automated fabric discovery
C) summary and host dashboards
D) VM-aware monitoring

Q42) Which tool would be used by operators to monitor and obtain reports for Cisco SAN switch events, performance, and inventory? (Source: Using SAN Management Tools)

A) Cisco Device Manager
B) Cisco Performance Manager
C) DCNM-SAN Server
D) Cisco Fabric Manager

Q43) Which SAN management tool is used to configure and manage a single switch in the SAN fabric? (Source: Using SAN Management Tools)

A) DCNM-SAN Server
B) Cisco Device Manager
C) Cisco Performance Manager
D) Network monitoring

https://t.me/learningnets

Module Self-Check Answer Key

Q1) D
Q2) B
Q3) C
Q4) A
Q5) D
Q6) B
Q7) B, E
Q8) C
Q9) B, D
Q10) B, E
Q11) A
Q12) D
Q13) B,E
Q14) E
Q15) C
Q16) B
Q17) B
Q18) D
Q19) A
Q20) C
Q21) A
Q22) D
Q23) B
Q24) D
Q25) B
Q26) C
Q27) B
Q28) B
Q29) C
Q30) D, E, F
Q31) D
Q32) A
Q33) A
Q34) A, B
Q35) A
Q36) B
Q37) D
Q38) D
Q39) A
Q40) D
Q41) D
Q42) C
Q43) B

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Module 5

Cisco Nexus Switch Management

Overview

In this module, you will learn how to identify which management tools are available for the Cisco Nexus Family of switches and how to configure the relevant tool to support the given design. You will learn how to configure and use the connectivity management processor (CMP) for management tasks, create users and roles for administrative access, and identify the system management features that are available.

Module Objectives

Upon completing this module, you will be able to identify and configure the management tools that are available on the Cisco Nexus Family of switches. This ability includes being able to meet these objectives:

- Evaluate and configure the CMP on the Cisco Nexus 7000 Series Switches and identify the management options that are available
- Identify which user management features are available on the Cisco Nexus Family of switches and explain how to configure those options to support the management infrastructure requirement
- Identify which system management features are available on the Cisco Nexus Family of switches and explain how to apply licensing and configure the switches to support the network requirements

https://t.me/learningnets

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

Lesson 1

Using the CMP

Overview

In this lesson, you will learn how to evaluate and configure the connectivity management processor (CMP) on the Cisco Nexus 7000 Series Switches and identify the management options that are available. By using the CMP, you will be able to manage the switch independently from the main control processor, so that console access remains available remotely even during switch reload.

Objectives

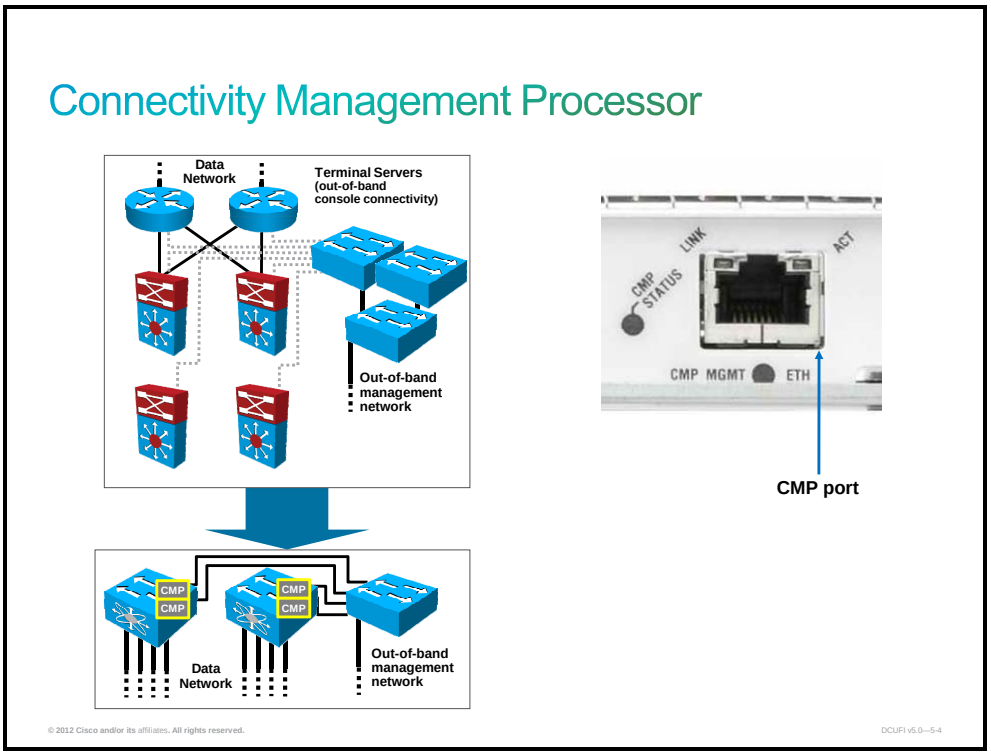
Upon completing this lesson, you will be able to evaluate and configure the CMP on the Cisco Nexus 7000 Series Switches and identify the management options that are available. This ability includes being able to meet these objectives:

- Describe the use of the CMP for management purposes
- Identify the configuration options that are available for the CMP
- Verify the CMP configuration
- Upgrade the CMP image
- Use the CMP to monitor and manage the control plane of the Cisco Nexus 7000 Series Switches

https://t.me/learningnets

Connectivity Management Processor

The CMP is a separate processor on the supervisor modules of the Cisco Nexus 7000 Series Switches. It provides an out-of-band (OOB) management interface. This topic describes the use of CMP for management purposes.



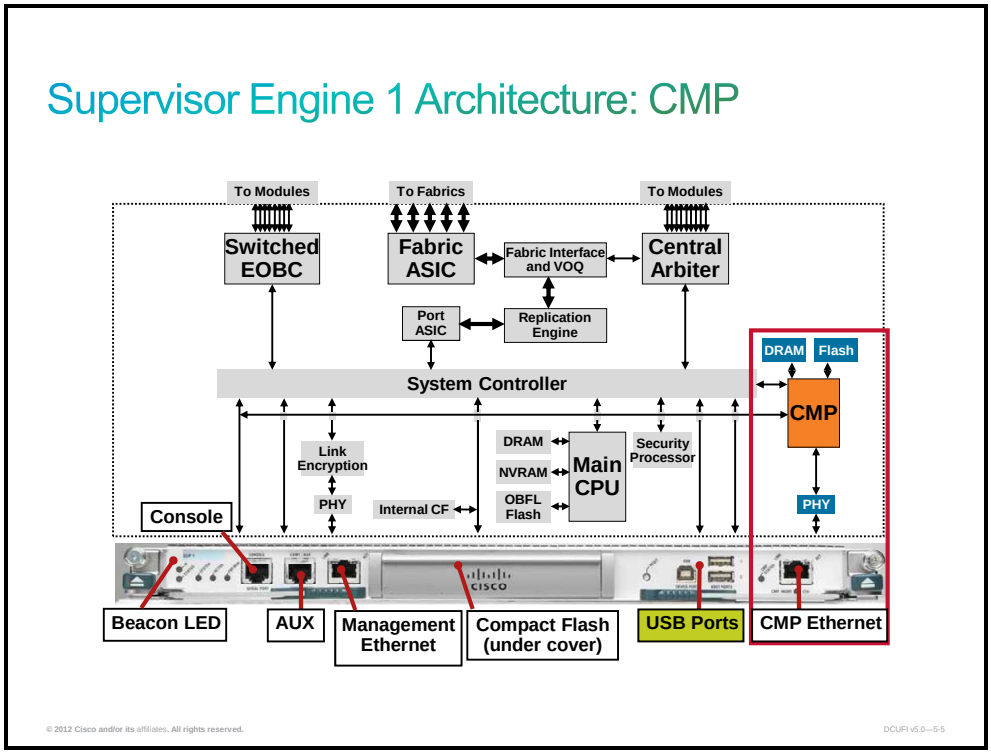
The CMP provides out-of-band (OOB) management and monitoring capability independent from the primary operating system. The CMP enables lights-out remote monitoring and management of the supervisor module, all other modules, and the Cisco Nexus 7000 Series system without the need for separate terminal servers. The CMP can reset all system components, including power supplies; it can also reset the host supervisor module to which it is attached, allowing a complete system restart.

These are key features of the CMP:

- **Dedicated operating environment:** Provides independent remote system management monitoring capabilities
- **Monitoring of supervisor status and initiation of resets:** Removes the need for separate terminal server devices for OOB management
- **System reset while retaining OOB Ethernet connectivity:** Provides visibility during the entire boot process
- **Capability to initiate a complete system power shutdown and restart:** Eliminates the need for local operator intervention
- **Login authentication:** Provides secure access to the OOB management environment
- **Access to supervisor logs:** Gives access to critical log information, which enables rapid detection and prevention of potential system problems
- **Control capability:** Ability to take complete console control of the supervisor
- **Dedicated front-panel LEDs:** Clearly identifies CMP status separately from the supervisor

© 2012 Cisco Systems, Inc.

https://t.me/learningnets



The figure shows the CMP Supervisor 1 Module architecture.

The CMP delivers the remote control through these ways:

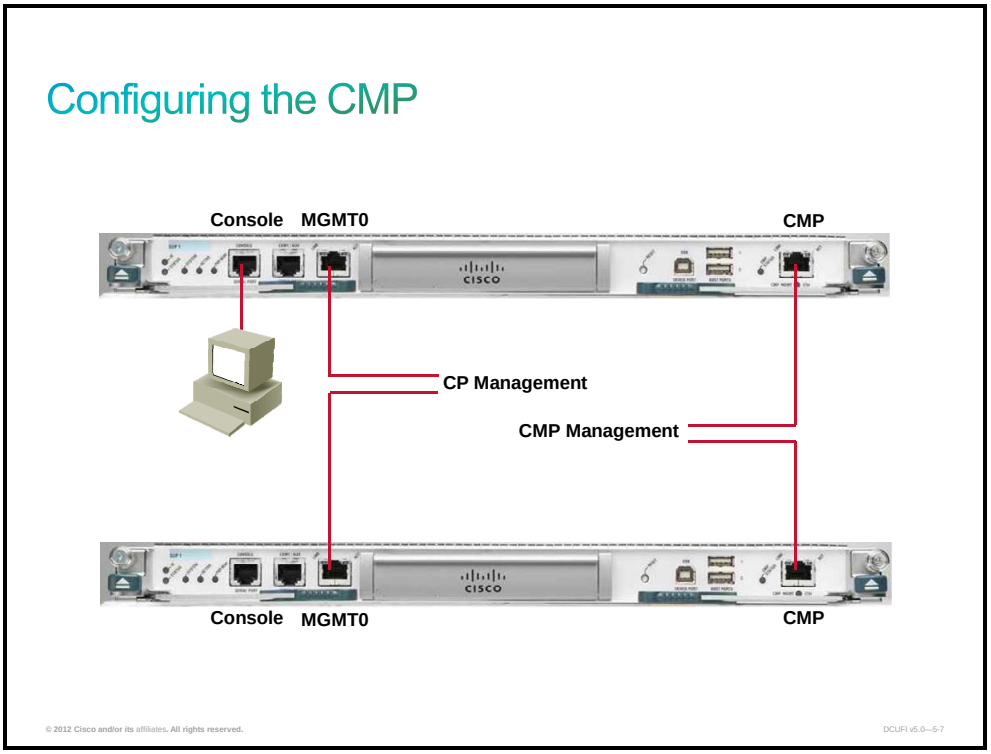
- A dedicated processor
- Its own memory
- Its own bootflash memory
- Separate Ethernet management port

The CMP runs an independent image that is stored in flash memory. Each CMP remains operational even if its supervisor module is in standby mode or the device is down because of issues such as over-temperature alarms. Each CMP gets power from an auxiliary power bus in the device. The device remains operational so long as you have at least one power cable that is attached to it.

https://t.me/learningnets

Configuring the CMP

There are a number of configuration options that are available for the CMP such as logging CMP messages, configuring the Cisco Discovery Protocol, and configuring CMPs on a dual-supervisor system. This topic identifies the configuration options that are available for the CMP.



The CMP runs in active mode on both supervisor modules, even when only one supervisor module is active, so you must configure each CMP individually. You can configure the unique IP address for each CMP from the active control processor by using Cisco Nexus Operating System (NX-OS) Software commands through either the CLI or scripts. To perform all other CMP configuration functions, connect directly to the CMP that you are configuring to perform those functions.

The following are best practices you should follow when performing an initial system setup on a dual-supervisor system:

- Connect four Ethernet cables to the system.
- Each supervisor requires two cables—one for the CMP port and the other for the management port.
- The CMP ports should be connected to the OOB management network.
- The Cisco Nexus 7000 Series system requires three IP addresses.
- Assign a unique IP address for each CMP.
- One IP address is shared by two supervisors.
- The active supervisor owns the IP address.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Accessing the CMP from the CP

- Access the CMP on the active supervisor module.

```
N7010# attach cmp
N7010-C1-cmp5 login: admin
Password: <password>
N7010-cmp#
```

- Enter configuration mode on the CMP.

```
N7010-cmp# configure terminal
N7010-cmp(config)#
```

- Exit the CMP console and returns to the Cisco NX-OS CLI on the CP.

```
N7010-cmp# ~,
N7010#
```

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

You can access the CMP from the active supervisor module. Before you begin, ensure that you are in the default virtual device context (VDC).

When the control processor and CMP are both operational, you can log into the CMP through the control processor using your username and password that is configured in the Cisco NX-OS Software or the admin username and password. If the control processor is configured with RADIUS or TACACS+, then your authentication is also managed by RADIUS or TACACS+. If the control processor is operational, the CMP accepts logins from users with network-admin privileges. The CMPs use the same authentication mechanism to configure the control processor (that is, RADIUS, TACACS+, or local). The control processor automatically synchronizes the admin password with the active and standby CMP so that you can use the admin username and password when a control processor is not operational.

https://t.me/learningnets

Configuring the CMP

- Configure the default gateway for the cmp-mgmt interface

```
N7010-cmp# configure terminal
N7010-cmp(config)# ip default-gateway 192.0.2.10
```

- Enter interface configuration mode and configure the IP address with appropriate mask for the cmp-mgmt interface.

```
N7010-cmp(config)# interface cmp-mgmt
N7010-cmp(config-if)# ip address 192.0.2.1/24
```

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

The figure shows the configuration steps to follow to configure CMP management interface from CMP.

Configuring the CMP (Cont.)

- Enter interface configuration mode for the cmp-mgmt interface.

```
N7010(config)# interface cmp-mgmt module 5
N7010(config-if-cmp)#
```

- Configure the IP address for this cmp-mgmt interface.

```
N7010(config-if-cmp)# ip address 192.0.2.1/24
```

- Configure the default gateway for this cmp-mgmt interface.

```
N7010(config-if-cmp)# ip default-gateway 192.0.2.10
```

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

The figure shows the configuration steps to follow to configure the CMP management interface from the control processor.

https://t.me/learningnets

Access to the CMP

- Enable SSH server (default setting).

```
N7010-cmp(config)# ssh server enable
```

- Enable Telnet server.

```
N7010-cmp(config)# telnet server enable
```

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

You can access the CMP via any of the following options:

- Control processor
- Secure Shell (SSH)
- Telnet

To access the CMP using SSH or Telnet, you must enable those sessions on the CMP. (By default, the SSH server session is enabled.)

https://t.me/learningnets

Configuring CDP for the CMP

- Enter interface configuration mode for the cmp-mgmt interface.

```
N7010(config)# interface cmp-mgmt module 5
N7010(config-if-cmp)#
```

- Enable CDP on this interface. This is enabled by default.

```
N7010(config)# interface cmp-mgmt module 5
N7010(config-if-cmp)# cdp enable
```

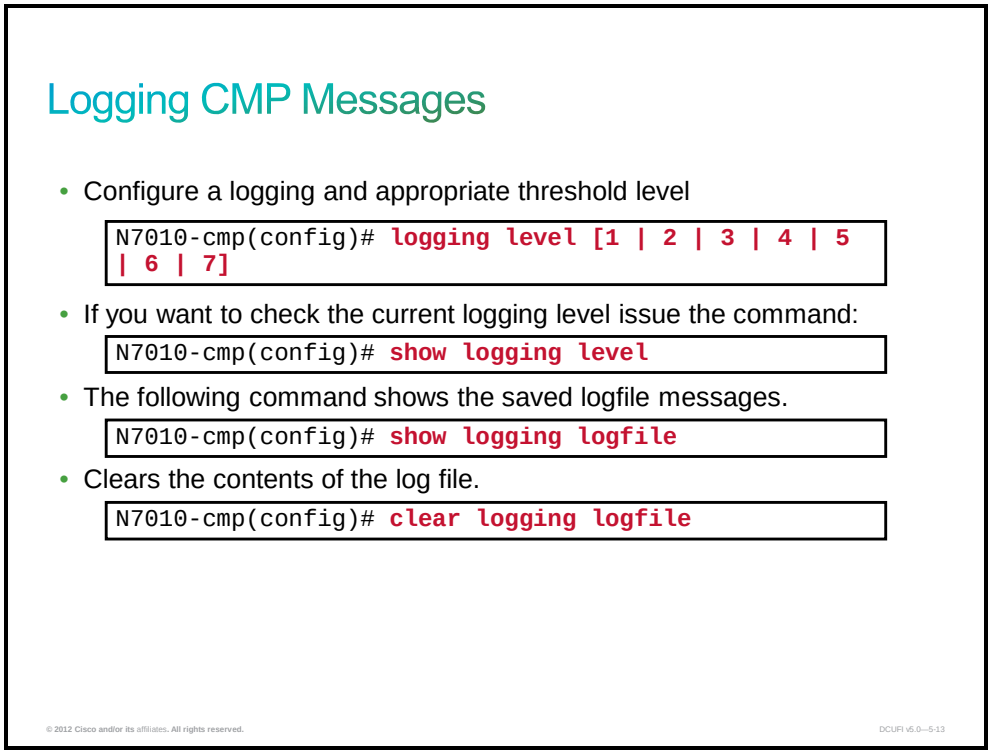
© 2012 Cisco and/or its affiliates. All rights reserved.

The Cisco Discovery Protocol is a media- and protocol-independent protocol that runs on all equipment that is manufactured by Cisco including routers, bridges, access and communication servers, and switches.

Cisco Discovery Protocol is enabled by default. You can disable Cisco Discovery Protocol and then re-enable it at a later time.

Cisco Discovery Protocol is enabled by default on an interface. You can disable Cisco Discovery Protocol on an interface.

https://t.me/learningnets



You can save up to 256 CMP messages in a log file, and you can specify a severity threshold for the messages that you want saved. When the file has 256 messages, the CMP automatically removes the oldest message whenever it saves a new message. When you specify a severity level, the CMP saves messages for that level and all levels below it in the log file.

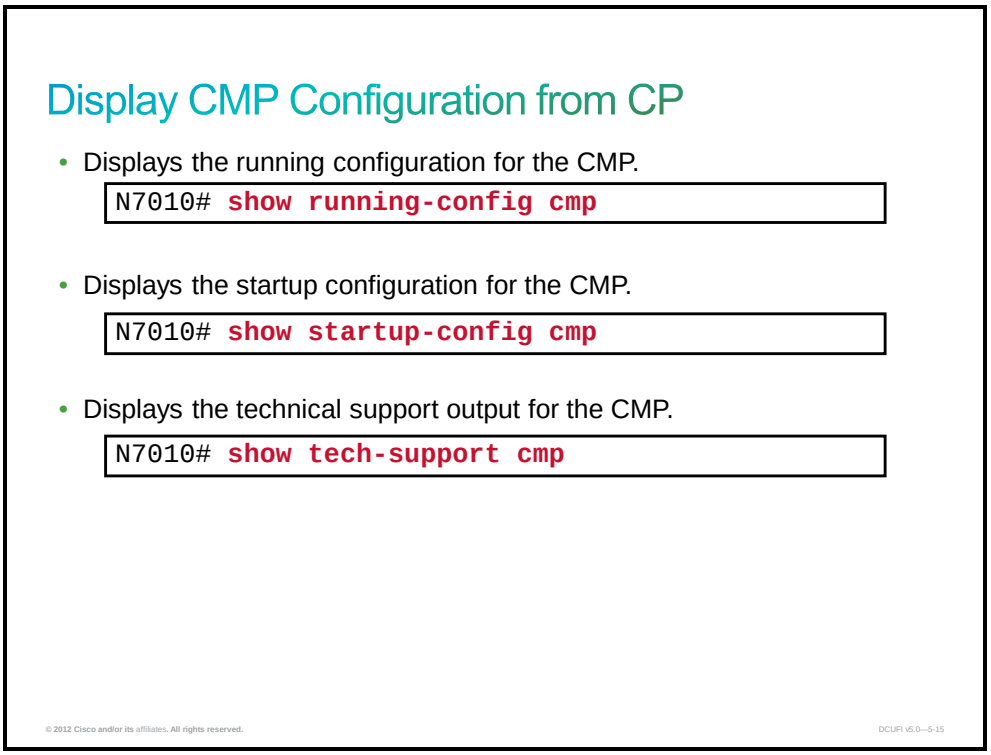
You can display all of the messages that are saved in the CMP log file or a specific number of the most recently saved messages by using the **show logging** command.

You can clear the contents of the log file by using the **clear logging logfile** command.

https://t.me/learningnets

Verifying the CMP

Once the CMP has been configured, it is important to verify the configuration for management purposes. This topic explains how to verify the CMP configuration.



The figure shows commands that are used to verify the CMP running and startup configuration. Also, you can display technical support output. All these commands are issued on the control processor.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Display CMP Information from the CMP

- Displays information about the CMP hardware.

```
N7010-cmp# show hardware
```

- Displays information about the cmp-mgmt interface.

```
N7010-cmp# show interface
```

- Displays information about the CMP processes.

```
N7010-cmp# show processes
```

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

On the CMP, you can display information about the CMP hardware, processes, and cmp-mgmt interface.

Display CMP Information from the CMP (Cont.)

- Displays the running configuration for the CMP

```
N7010-cmp# show running-config
```

- Displays the users logged into the system

```
N7010-cmp# show users
```

- Displays the software image versions for the supervisor CP and the CMP

```
N7010-cmp# show version
```

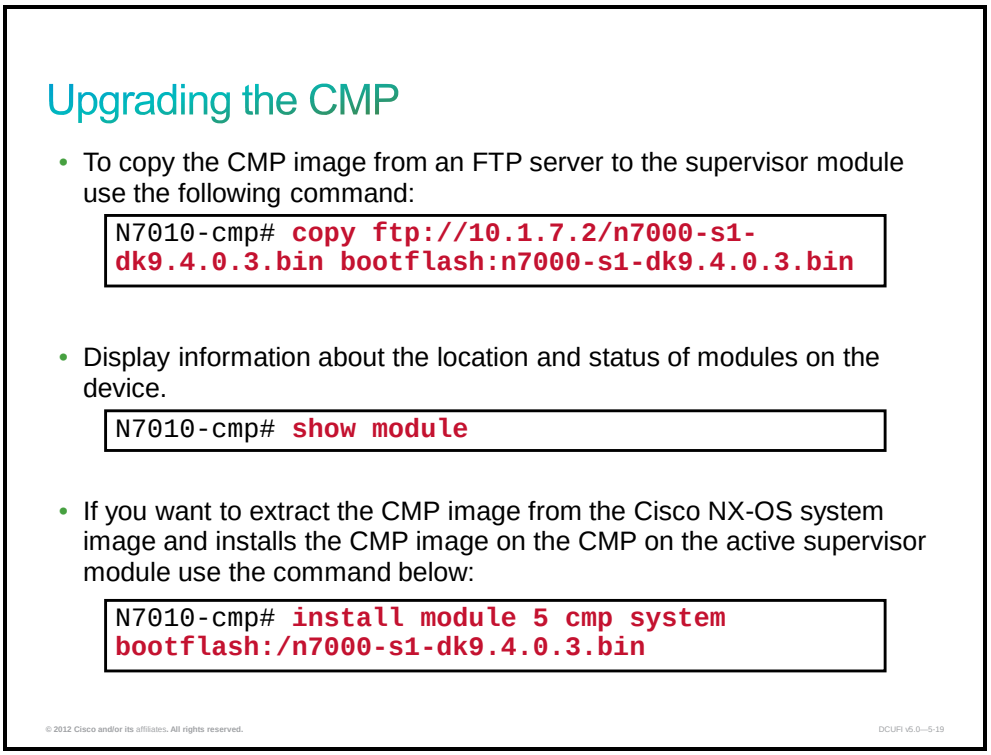
© 2012 Cisco and/or its affiliates. All rights reserved.

The figure shows the commands that are used to verify the CMP configuration, users that are logged into the system, and software image versions for the supervisor control processor and the CMP.

https://t.me/learningnets

Upgrading the CMP

The CMP is a separate management processor with its own RAM, bootflash, and front panel management Ethernet port. This topic describes how to upgrade the CMP image.



You can upgrade the CMP image, which is part of the Cisco NX-OS Software system image and contains a subset of commands to support the CMP features.

Note	The CMP image is independent of the control processor image, so the version of the CMP image might not match the version of the control processor image. To make sure that the CMP is running the latest compatible image, use the **install all** command from the Cisco NX-OS Software CLI on the control processor.

Use the following procedure if you want to update the CMP image only:

Step 1 Copy the CMP image from an FTP server to the supervisor module.

Step 2 Verify the location and status of supervisors.

Step 3 Install the CMP image on the CMP that is on the active supervisor.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Upgrading the CMP (Cont.)

- To extract the CMP image from the Cisco NX-OS system image and installs the CMP image on the CMP on the standby supervisor module, if present, use command:

```
N7010-cmp# install module 6 cmp system
bootflash:/n7000-s1-dk9.4.0.3.bin
```

- In order to complete upgrade, reload the CMP for the active supervisor module by issuing the following command:

```
N7010-cmp# reload cmp module 5
```

- Reload the CMP for the standby supervisor module, if present, to complete the upgrade procedure.

```
N7010-cmp# reload cmp module 6
```

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

Step 4 Install the CMP image on the CMP that is on the standby supervisor.

Step 5 Reload the CMP for the active supervisor.

Step 6 Reload the CMP for the standby supervisor.

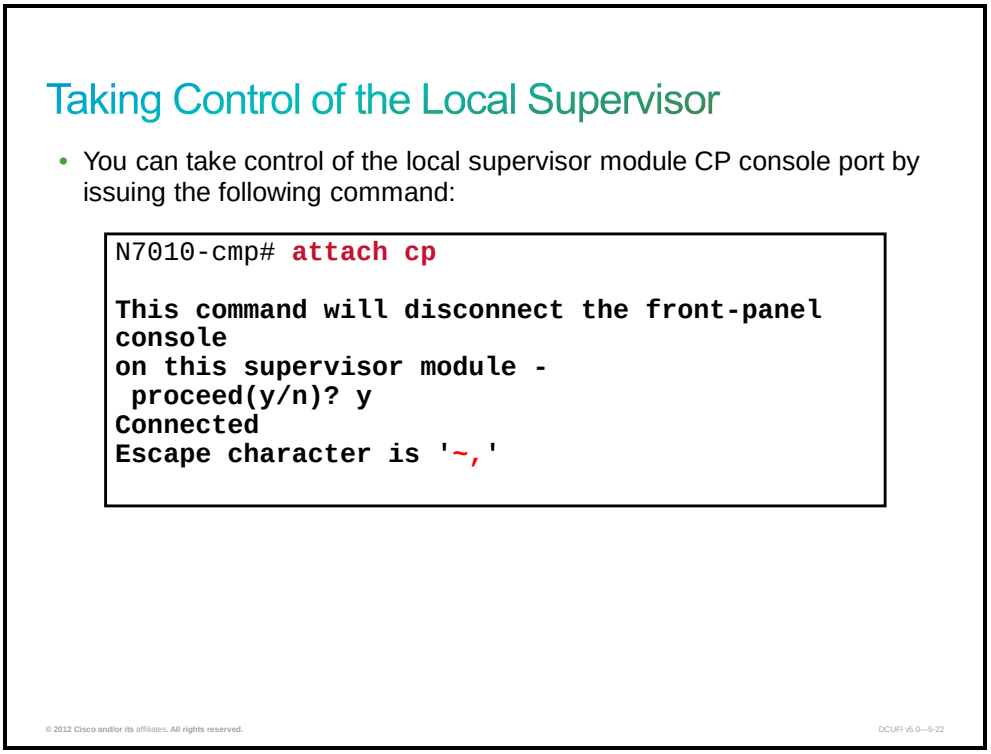
Step 7 Display the BIOS and software image versions of the CMP using the **show version** command.

.

https://t.me/learningnets

Using the CMP

As a separate management processor, the CMP can be used to monitor and manage the control plane of the Cisco Nexus 7000 Series Switches when the control plane is unavailable, or as an alternative management interface. This topic describes how to use the CMP to monitor and manage the control plane of the Cisco Nexus 7000 Series Switches.



Use the **attach cp** command on the CMP to take control of the local supervisor module control processor console port.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Saving Console Output on the CMP

- Enables the logging of console output on the CMP in a file of the size specified by this integer in this command or in a default sized file (50 KB) if a file size is not specified.

```
N7010(config)# capture cp console 100
```

- Changes the KB size of the console output log file. Specify an integer between 10 and 100. The default is 50.

```
N7010(config)# capture cp size 100
```

- Displays all of the logs in the log file.

```
N7010# show capture all
```

- Displays the most recently logged output. You include an integer to specify the number of lines to display.

```
N7010# show capture last
```

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

Beginning with Cisco NX-OS Software Release 5.0, you can log console output on the CMP to help you troubleshoot problems that you might encounter when reloading the CP onto your Cisco Nexus 7000 Series switch.

Rebooting Procedures

- If you reboot a supervisor module from the Cisco NX-OS CLI on the CP, the CMP also reboots. Use the **reload soft command** to reboot only the supervisor module CP and not the CMP.
- You can reboots the supervisor module by issuing command:

```
N7010-cmp# reload cp
```

- In order to reboots the whole Cisco NX-OS device use the command:

```
N7010-cmp# reload system
```

- At the end you can reboot CMP itself by issuing command:

```
N7010-cmp# reload cmp
```

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

You can reboot the control processor from the CMP. The CMP can reset all system components, including power supplies; it can also reset the host supervisor module to which it is attached, allowing a complete system restart.

To reboot the entire Cisco NX-OS device from the CMP, use the **reload system** command.

You can also reboot the CMP from the CMP.

https://t.me/learningnets

Summary

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

Summary

- The CMP is a separate processor on the Cisco Nexus 7000 Series supervisor module that is in addition to the main CP.
- The CMP provides a second network interface to the device and can be used when the CP is not reachable.
- The CMP can be accessed to verify its configuration settings.
- The CMP can be upgraded independently from Nexus device
- The CMP can be accessed to perform system operations such as taking over the CP console or restarting the CP.

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

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Lesson 2

Configuring User Management

Overview

In this lesson, you will learn how to identify which user management features are available on the Cisco Nexus Series Switches and how to configure those options to support the management infrastructure requirement. Those management features enable you to configure administrative access to your devices that is safe, encrypted, and in accordance with security policy

Objectives

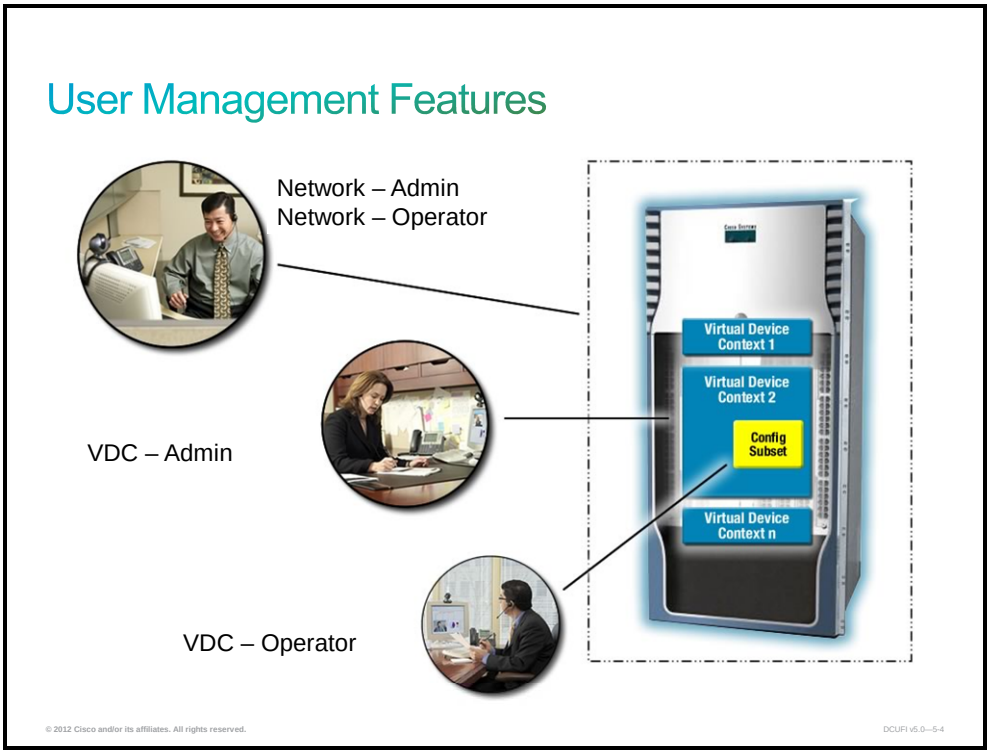
Upon completing this lesson, you will be able to identify which user management features are available on the Cisco Nexus Family of switches and explain how to configure those options to support the management infrastructure requirement. You will be able to meet these objectives:

- Explain which user management features are available on the Cisco Nexus Family of switches
- Configure AAA on the Cisco Nexus Family of switches
- Configure the Cisco Nexus Family of switches to support SSH management connections
- Configure user accounts and roles for local authentication on the Cisco Nexus Family of switches

https://t.me/learningnets

User Management Features

User management is the process of using secure connections for management connectivity and restricting administrative access either locally or through a central authentication, authorization, and accounting (AAA) server. This topic provides an overview of the user management features on the Cisco Nexus Family of switches.



On the Cisco Nexus 7000 Series Switches, user roles contain rules that define the operations that are allowed for the user who is assigned the role.

The Cisco Nexus Operating System (NX-OS) Software provides four default user roles:

- **Network-admin:** Complete read-and-write access to the entire Cisco NX-OS Software device (only available in the default virtual device context [VDC])
- **Network-operator:** Complete read access to the entire Cisco NX-OS Software device (only available in the default VDC)
- **vdc-admin:** Read-and-write access that is limited to a VDC
- **vdc-operator:** Read access that is limited to a VDC

Note	You cannot change the default user roles.

You can create custom roles within a VDC.

The VDCs on the same physical device do not share user roles. Each VDC maintains an independent user role database.

The **network-admin** and **network-operator** are the default roles on the Cisco Nexus 5000 and 5500 Platform switches.

© 2012 Cisco Systems, Inc.

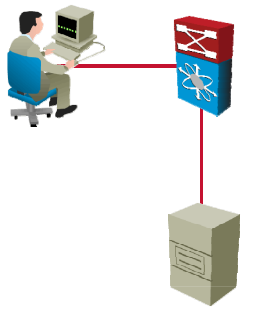
https://t.me/learningnets

Authentication, Authorization, and Accounting

AAA is the process of identifying users, authorizing their access rights, and accounting (logging) the tasks they performed while logged into the switch through a central AAA server. This topic explains how to configure AAA on the Cisco Nexus Family of switches.

About AAA

- AAA Security Services
  - Authentication
  - Authorization
  - Accounting
- AAA Service Configuration Options
  - User Telnet or SSH login authentication
  - Console login authentication
  - Cisco TrustSec authentication
  - 802.1X authentication
  - EAPoUDP authentication for Network Admission Control (NAC)
  - User management session accounting
  - 802.1X accounting



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

The AAA feature allows you to verify the identity of, grant access to, and track the actions of users managing a Cisco NX-OS device. Cisco NX-OS devices support RADIUS or TACACS+ protocols.

Based on the user ID and password combination that you provide, Cisco NX-OS devices perform local authentication or authorization, using the local database, or remote authentication or authorization, using one or more AAA servers. A preshared secret key provides security for communication between the Cisco NX-OS device and AAA servers. You can configure a common secret key for all AAA servers or for only a specific AAA server.

AAA security provides these services:

Authentication

Authentication identifies users, including login and password dialog, challenge and response, messaging support, and depending on the security protocol that you choose, encryption. Authentication is the process of verifying the identity of the person or device that is accessing the Cisco NX-OS device. The verification is based on the user ID and password combination that is provided by the entity trying to access the Cisco NX-OS device. Cisco NX-OS devices allow you to perform local authentication (using the local lookup database) or remote authentication (using one or more RADIUS or TACACS+ servers).

https://t.me/learningnets

Authorization

Authorization provides access control. AAA authorization is the process of assembling a set of attributes that describe what the user is authorized to perform. Authorization in the Cisco NX-OS Software is provided by attributes that are downloaded from AAA servers. Remote security servers, such as RADIUS and TACACS+, authorize users for specific rights by associating attribute-value pairs, which define those rights with the appropriate user.

Accounting

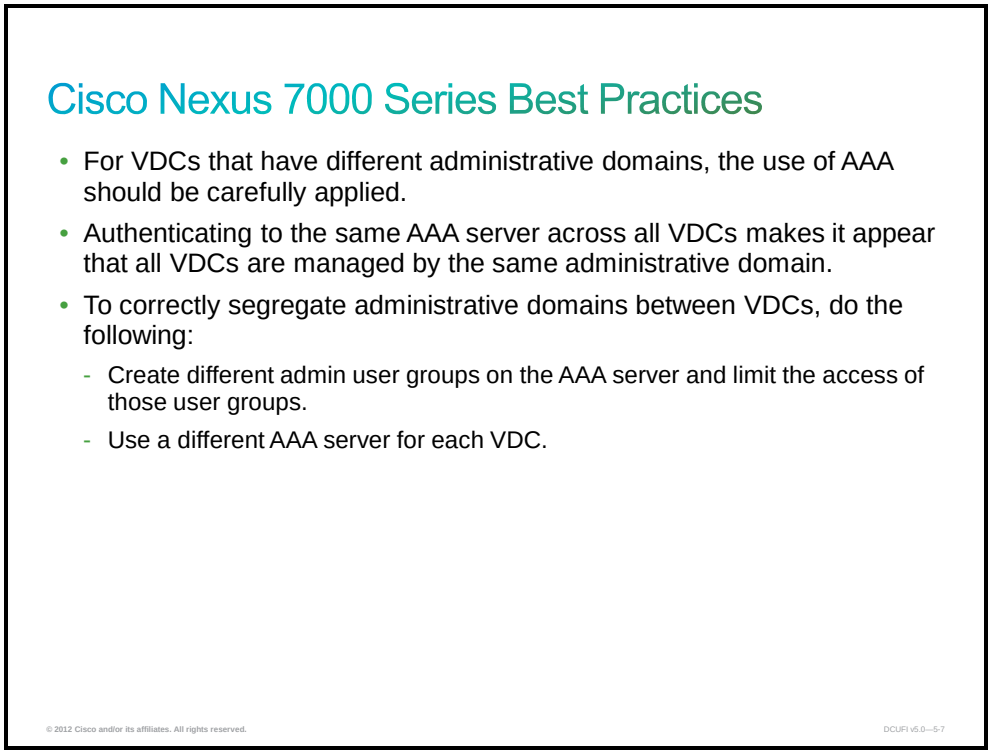
Accounting provides the method for collecting information, logging the information locally, and sending the information to the AAA server for billing, auditing, and reporting. The accounting feature tracks and maintains a log of every management session that is used to access the Cisco NX-OS device. You can use this information to generate reports for troubleshooting and auditing purposes. You can store accounting logs locally or send them to remote AAA servers.

AAA Server Groups

You can specify remote AAA servers for AAA using server groups. A server group is a set of remote AAA servers that implement the same AAA protocol. The purpose of a server group is to provide for failover servers in case a remote AAA server fails to respond. If the first remote server in the group fails to respond, the next remote server in the group is tried until one of the servers sends a response. If all the AAA servers in the server group fail to respond, then that server group option is considered a failure. If required, you can specify multiple server groups. If the Cisco NX-OS device encounters errors from the servers in the first group, it tries the servers in the next server group.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

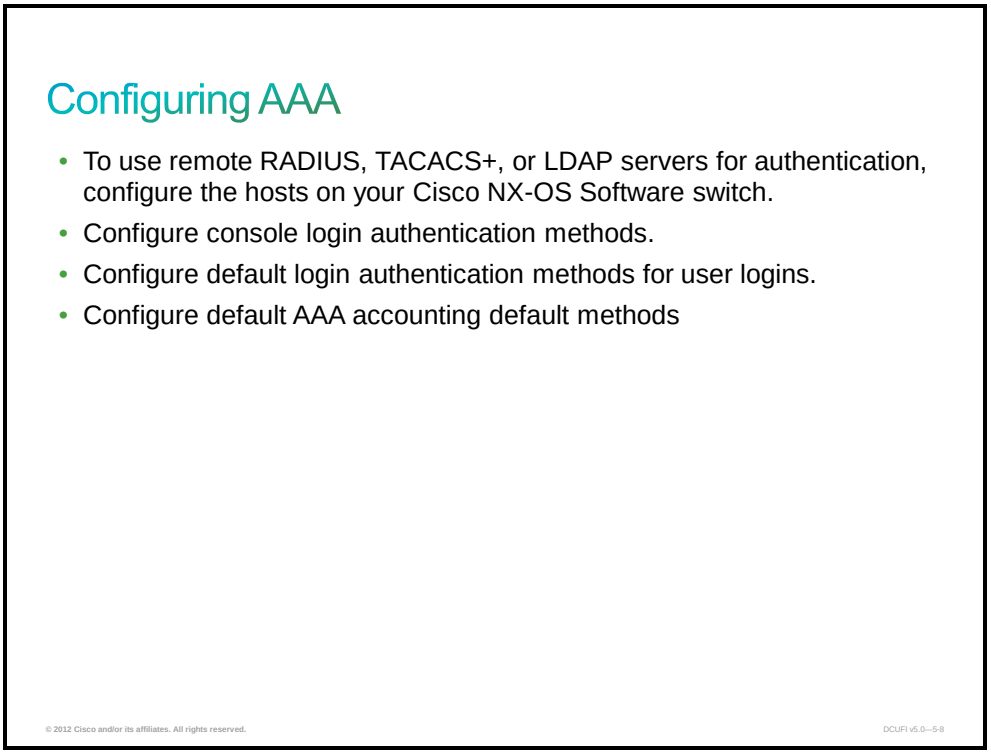


On the Cisco Nexus 7000 Series Switches, if VDCs have separate administrative domains (different VDC administrators), the use of AAA for authentication and authorization should be carefully applied.

Authenticating to the same AAA server across VDCs will implicitly apply authentication as if all VDCs were managed by a single administrative domain. To correctly segregate administrative domains between VDCs, and to prevent an administrative account from one VDC having access and privileges on another VDC, the following options are available:

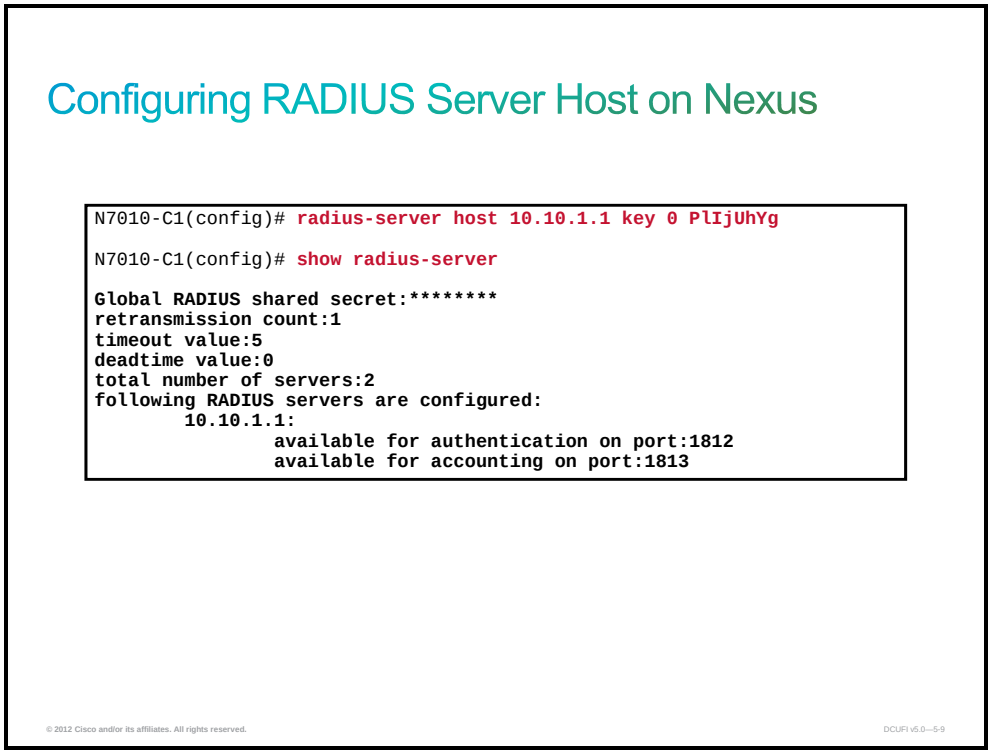
- Create different admin user groups on the AAA server and limit the access of those user groups, using a feature such as network access restriction (NAR) (on Cisco Secure Access Control Server [ACS]), to specific AAA client IP addresses or groups. Since each VDC sources its AAA traffic from its Local Management Interface IP address, this process will allow the distinction between VDCs. It will essentially limit the access of a given VDC admin group to a specific subset of VDCs. The AAA server will identify the requesting VDC client IP address and take that factor into account during the authorization decision process.
- Use a different AAA server for each VDC. VDCs allow the definition of discrete AAA servers per VDC. The administrative accounts should be configured such that the admin account for any given VDC does not exist in the user database of the other AAA servers.

https://t.me/learningnets



The figure describes the process for configuring AAA.

https://t.me/learningnets



RADIUS Server Configuration Process

Step 1 If needed, enable Cisco Fabric Services configuration distribution for RADIUS.

Step 2 Establish the RADIUS server connections to the Cisco NX-OS device.

Step 3 Configure the RADIUS secret keys for the RADIUS servers.

Step 4 If needed, configure RADIUS server groups with subsets of the RADIUS servers for AAA authentication methods.

Step 5 If needed, configure any of the following optional parameters:

Step 6 Dead-time interval

Step 7 RADIUS server specification that is allowed at user login

Step 8 Timeout interval

Step 9 TCP port

Step 10 (Optional) If RADIUS distribution is enabled, commit the RADIUS configuration to the fabric.

https://t.me/learningnets

Configuring RADIUS Server Groups

```
N7010-C1(config)# aaa group server radius RadServer
N7010-C1(config-radius)# server 10.10.1.1
N7010-C1(config-radius)# server 10.10.20.2
N7010-C1(config-radius)# deadtime 30
N7010-C1(config-radius)# use-vrf management

N7010-C1(config-radius)# show radius-server groups

total number of groups:2
following RADIUS server groups are configured:
        group radius:
                server: all configured radius servers
        group RadServer:
                deadtime is 30
                vrf is management
```

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

You can specify one or more remote AAA servers for authentication using server groups. All members of a group must belong to the RADIUS protocol. The servers are tried in the same order in which you configure them. You can configure up to 100 server groups. (On a Cisco Nexus 7000 Series switch, this is per VDC.)

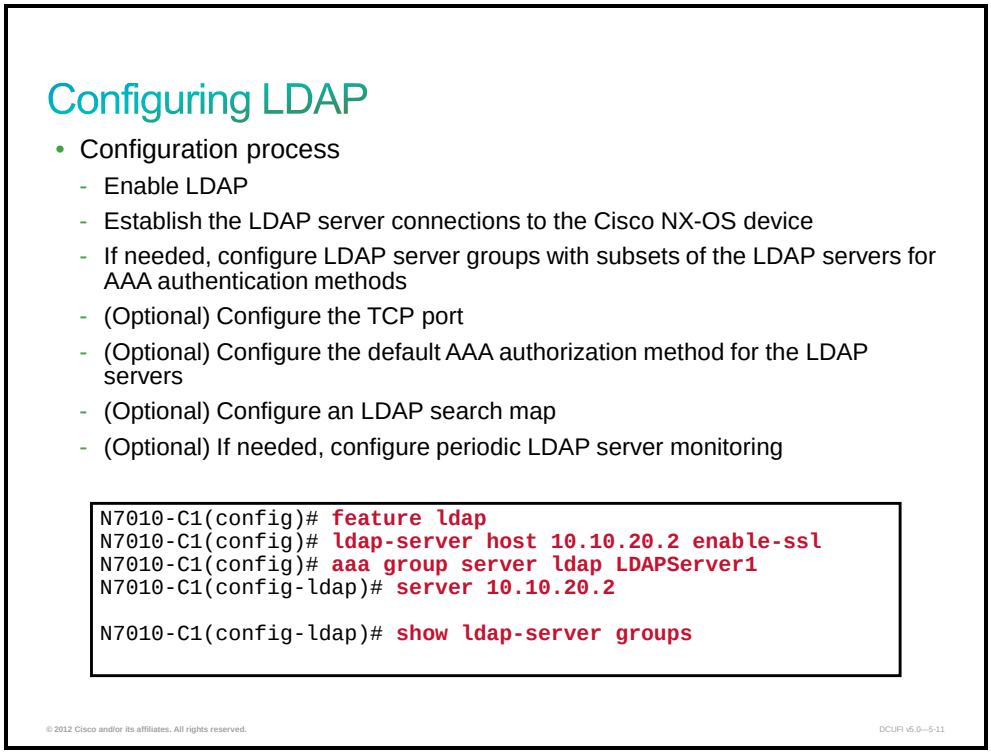
You can configure these server groups at any time but they only take effect when you apply them to an AAA service.

Note Cisco Fabric Services does not distribute RADIUS server group configurations.

Before you begin, ensure that all servers in the group are RADIUS servers.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets



The Lightweight Directory Access Protocol (LDAP) provides centralized validation of users attempting to gain access to the Cisco Nexus 7000 Series switch. LDAP services are maintained in a database on an LDAP daemon running on a UNIX or Windows NT workstation. Access is required to configure the LDAP server before configuring the LDAP features on the Cisco NX-OS device.

The LDAP client-server protocol uses TCP (TCP port 389) for transport requirements, and the Cisco NX-OS device provides centralized authentication using the LDAP protocol.

The figure details the steps that are required to configure LDAP on the Cisco NX-OS device.

Note For additional details about LDAP configuration on Cisco Nexus 7000 Series Switches, refer to the Configuring LDAP section of the *Cisco Nexus 7000 Series NX-OS Security Configuration Guide, Release 5.x* at: http://www.cisco.com/en/US/docs/switches/datacenter/sw/5_x/nx-os/security/configuration/guide/b_Cisco_Nexus_7000_NX-OS_Security_Configuration_Guide__Release_5.x_chapter_0111.html

https://t.me/learningnets

Configuring Default Login Authentication Methods

```
N7010-C1(config)# aaa authentication login default group radius

N7010-C1(config)# show aaa authentication

        default: group radius
        console: group radius
        dot1x: not configured
        eou: not configured
```

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

The login authentication methods include the following:

- Global pool of RADIUS servers
- Named subset of RADIUS, TACACS+, or LDAP servers
- Local database on the Cisco NX-OS Software device
- Username only

The default method is local, but you have the option to disable it.

Note	The configuration and operation of AAA for the console login apply only to the default VDC.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Configuring Console Login Authentication Methods

```
N7010-C1(config)# aaa authentication login console group RadServer

N7010-C1(config)# show aaa authentication

        default: local
        console: group RadServer
        dot1x: not configured
        eou: not configured
```

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

The console login authentication methods include the following:

- Global pool of RADIUS servers
- Named subset of RADIUS, TACACS+, or LDAP servers
- Local database on the Cisco NX-OS device
- Username only (none)

The default method is local, but you have the option to disable it.

The group-list argument consists of a space-delimited list of group names. The group names are the following:

- **radius:** Uses the global pool of RADIUS servers for authentication.
- **named-group:** Uses a named subset of RADIUS, TACACS+, or LDAP servers for authentication.
- **local:** The local method uses the local database for authentication,
- **none:** The none method specifies that no AAA authentication should be used.

The default console login method is local, which is used when no methods are configured or when all the configured methods fail to respond, unless fallback to local is disabled for the console login.

https://t.me/learningnets

Configuring Accounting Default Methods

```
N7010-C1(config)# aaa accounting default group radius
N7010-C1(config)# show aaa accounting
        default: group radius
```

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

Cisco NX-OS Software supports TACACS+ and RADIUS methods for accounting. Cisco NX-OS devices report user activity to TACACS+ or RADIUS security servers in the form of accounting records. Each accounting record contains accounting attribute-value pairs and is stored on the AAA server.

When you activate AAA accounting, the Cisco NX-OS device reports these attributes as accounting records, which are then stored in an accounting log on the security server.

You can create default method lists defining specific accounting methods, which include the following:

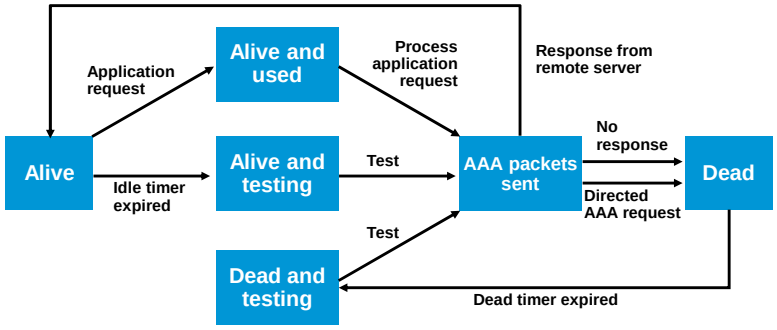
- **RADIUS server group:** Uses the global pool of RADIUS servers for accounting
- **Specified server group:** Uses a specified RADIUS or TACACS+ server group for accounting
- **Local:** Uses the local username or password database for accounting

Note	If you have configured server groups and the server groups do not respond, by default, the local database is used for authentication.

https://t.me/learningnets

AAA Server Monitoring

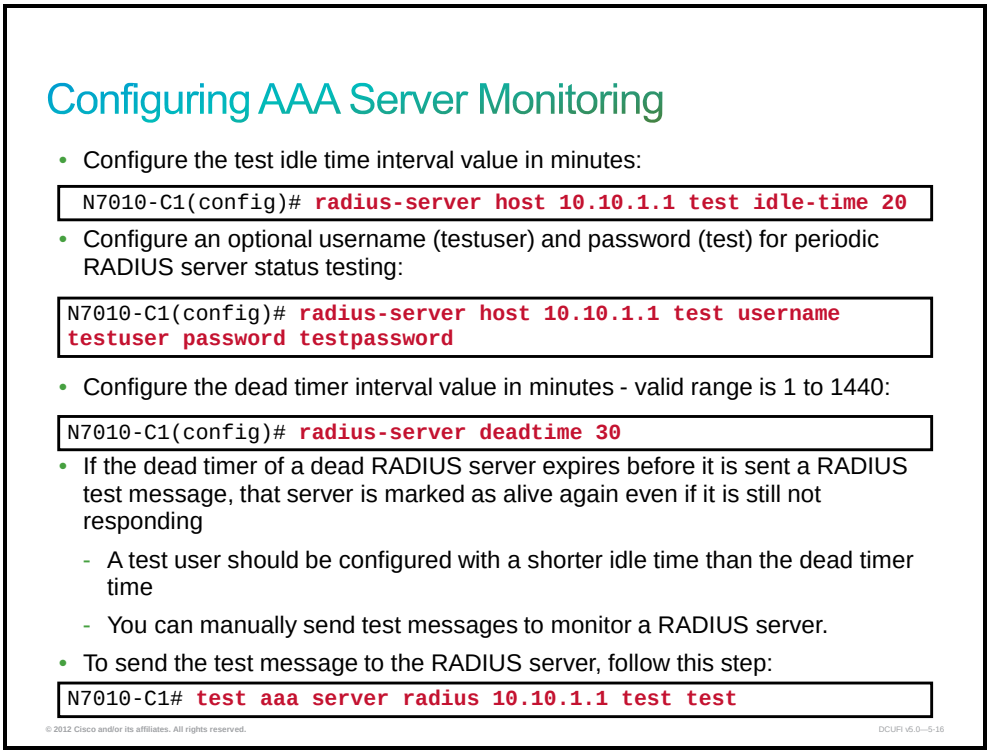
- Prevents delayed processing of AAA requests caused by unresponsive AAA servers
- The monitoring intervals for alive servers and dead servers are different and can be configured
- AAA server monitoring is performed by sending a test authentication request to the AAA server



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

An unresponsive AAA server introduces delay in the processing of AAA requests. A Cisco Nexus switch can periodically monitor an AAA server to check whether it is responding (or alive) to save time in processing AAA requests. The Cisco Nexus switch marks unresponsive AAA servers as dead and does not send AAA requests to any dead AAA servers. A Cisco Nexus switch periodically monitors dead AAA servers and brings them to the alive state once they are responding. This monitoring process verifies that an AAA server is in a working state before real AAA requests are sent its way. Whenever an AAA server changes to the dead or alive state, a Simple Network Management Protocol (SNMP) trap is generated and the Cisco Nexus switch warns the administrator that a failure is taking place before it can affect performance.

https://t.me/learningnets



To Configure AAA Server Monitoring follow these steps:

Step 1 Configure the test idle time interval value in minutes:

switch(config)# **radius-server host 10.10.1.1 test idle-time 20**

Step 2 Configure an optional username (testuser) and password (test) for periodic RADIUS server status testing:

switch(config)# **radius-server host 10.10.1.1 test username***testuser***password testpassword**

Step 3 Configure the dead timer interval value in minutes. The valid range is 1 to 1440:

switch(config)# **radius-server deadtime 30**

Note If the dead timer of a dead RADIUS server expires before it is sent a RADIUS test message, that server is marked as alive again even if it is still not responding. A test user should be configured with a shorter idle time than the dead timer time.

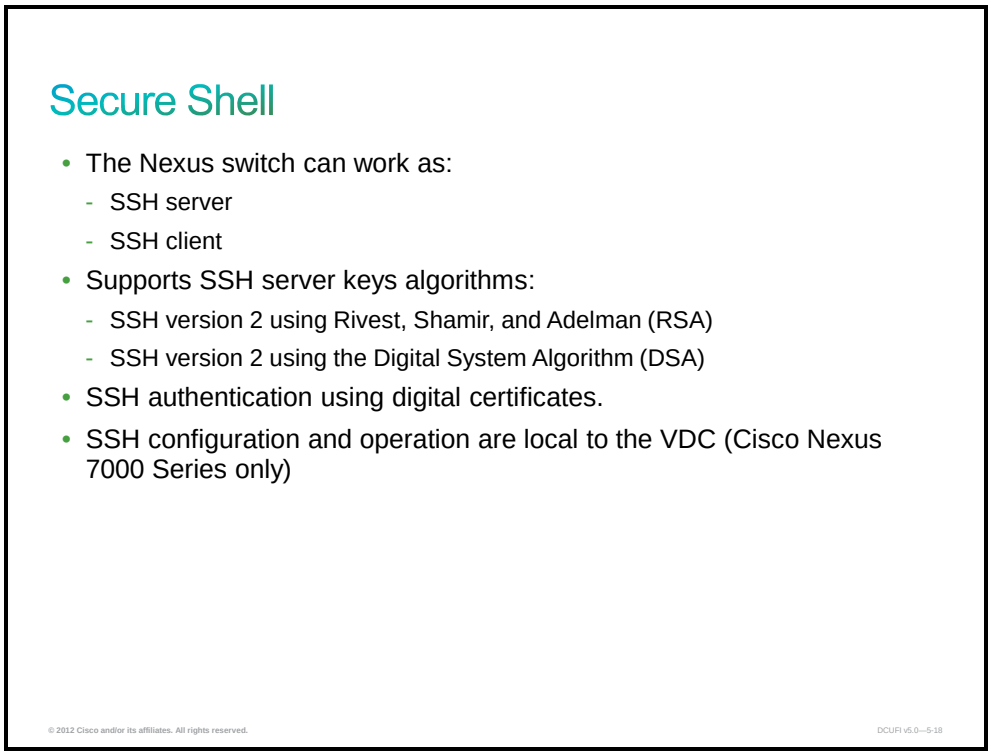
Step 4 Manually send test messages to monitor a RADIUS server:

switch#**test aaa server radius 10.10.1.1 test test**

© 2012 Cisco Systems, Inc.
https://t.me/learningnets

Secure Shell

Secure Shell (SSH) provides a secure management connection between the management station and the Cisco Nexus switch. This topic explains how to configure the Cisco Nexus Family of switches to support SSH management connections.



SSH Server

You can use the SSH server to enable an SSH client to make a secure, encrypted connection to a Cisco NX-OS device. SSH uses strong encryption for authentication. The SSH server in the Cisco NX-OS Software can interoperate with publicly and commercially available SSH clients. The user authentication mechanisms that are supported for SSH are RADIUS, TACACS+, and the use of locally stored usernames and passwords.

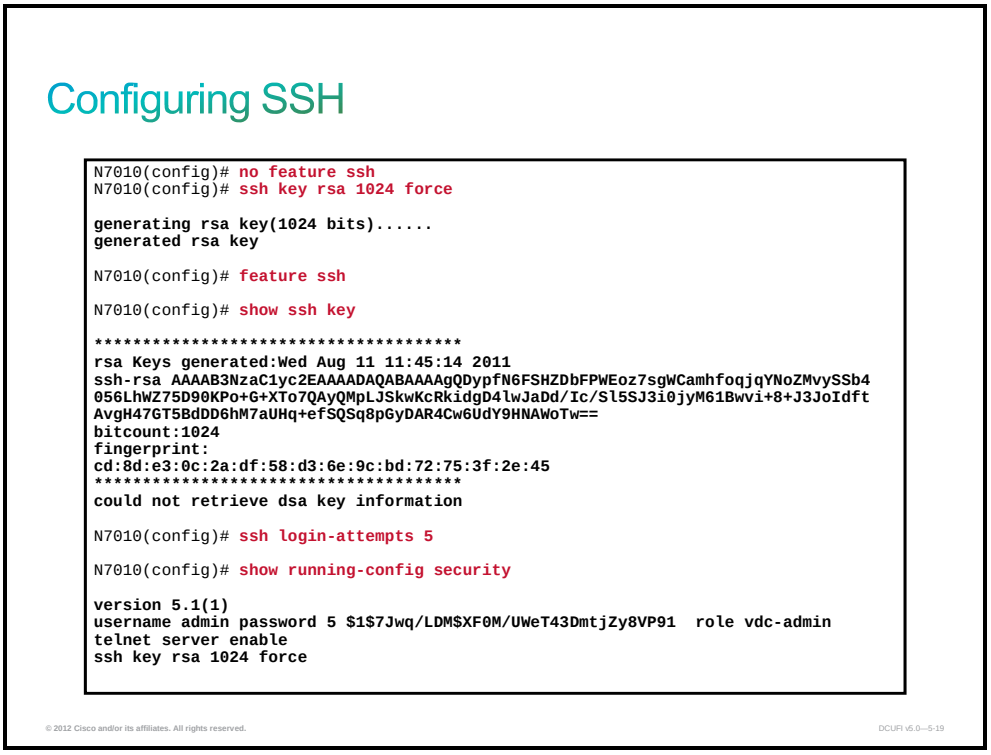
SSH Client

The SSH client feature is an application that runs over the SSH Protocol to provide device authentication and encryption. The SSH client enables a Cisco NX-OS device to make a secure, encrypted connection to another Cisco NX-OS device or to any other device that runs the SSH server. This connection provides an outbound connection that is encrypted. With authentication and encryption, the SSH client allows for a secure communication over an insecure network. The SSH client in the Cisco NX-OS Software works with publicly and commercially available SSH servers.

https://t.me/learningnets

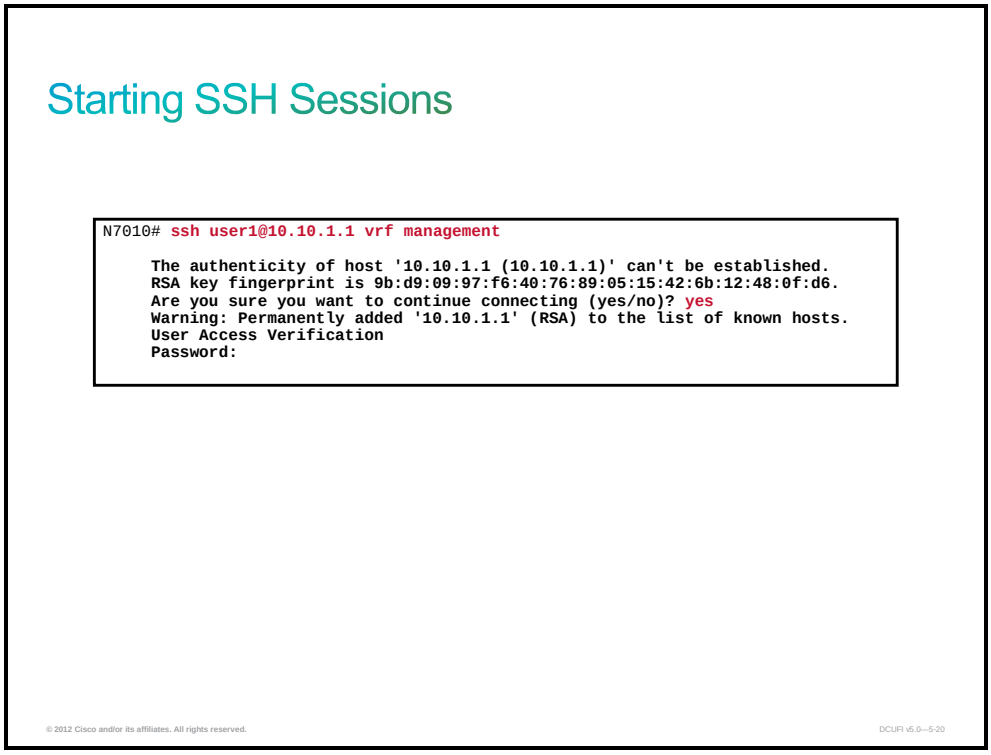
SSH has the following configuration guidelines and limitations:

- The Cisco NX-OS Software supports only SSH version 2 (SSHv2).
- You can configure your device for either SSH authentication using an X.509 certificate or SSH authentication using a public key certificate, but not both. If either of them is configured and the authentication fails, you are prompted for a password.
- Beginning in Cisco NX-OS Software Release 5.1, SSH runs in Federal Information Processing Standard (FIPS) mode.
- The Secure FTP (SFTP) server feature does not support the regular SFTP **chown** and **chgrp** commands.
- When the SFTP server is enabled, only the admin user can use SFTP to access the switch.



The figure describes how to configure SSH.

© 2012 Cisco Systems, Inc.
https://t.me/learningnets

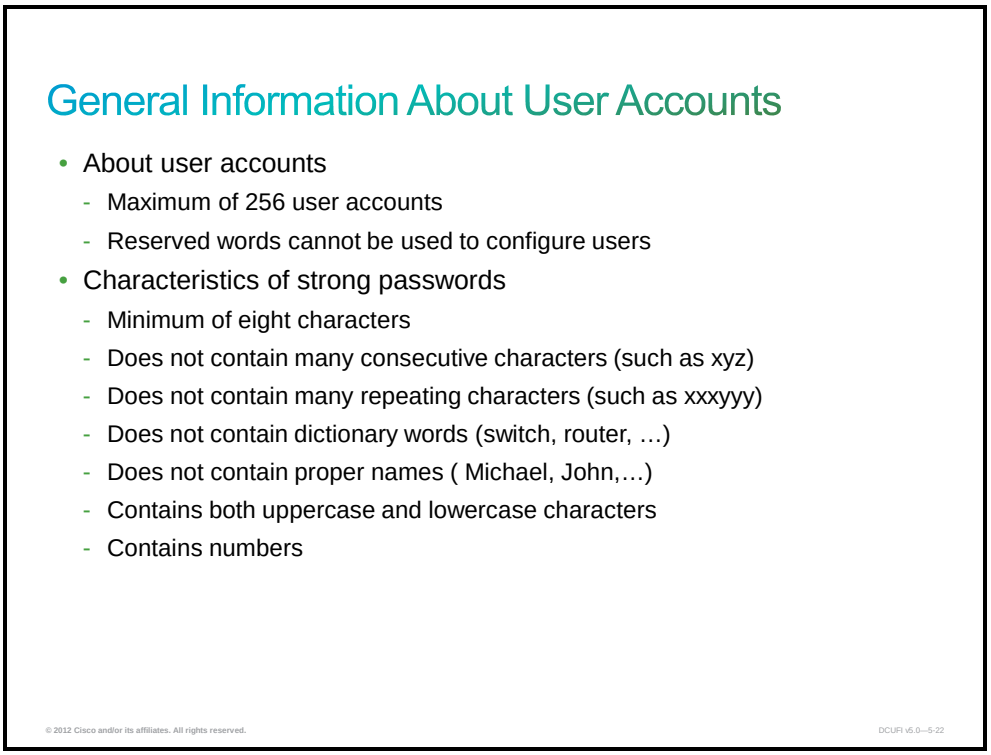


You can start SSH sessions using IP version 4 (IPv4) or IP version 6 (IPv6) to connect to remote devices from the Cisco NX-OS device.

https://t.me/learningnets

User Accounts and Roles

Users administrating the Cisco Nexus switch require a username and an associated role that specifies which tasks they are able to perform. This topic explains how to configure user accounts and roles for local authentication on the Cisco Nexus Family of switches.



You can configure a maximum of 256 user accounts. By default, the user account does not expire unless you explicitly configure it to expire. The expire option determines the date when the user account is disabled.

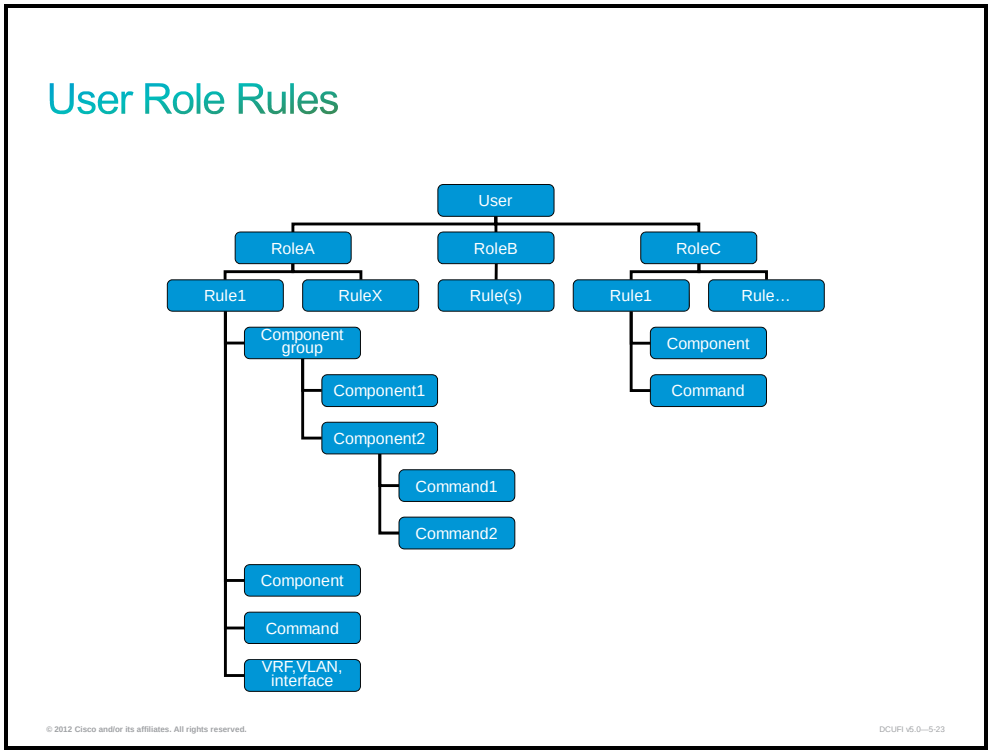
On the Cisco Nexus 7000 Series Switches, users can have user accounts on multiple VDCs. These users can move between VDCs after an initial connection to a VDC has been created.

The following words are reserved and cannot be used to configure users: bin, daemon, adm, lp, sync, shutdown, halt, mail, news, uucp, operator, games, gopher, ftp, nobody, nscd, mailnull, root, rpc, rpcuser, xfs, gdm, mtsuser, ftpuser, man, and sys.

Caution	Usernames must begin with an alphanumeric character and can contain only these special characters: (+= . _ \ -). The # and ! symbols are not supported. If the username contains characters that are not allowed, the specified user is unable to log in.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets



The rule is the basic element of a role. A rule defines what operations the role allows the user to perform. You can apply rules for the following parameters:

- **Command:** A command or group of commands that are defined in a regular expression
- **Feature:** A command or group of commands that are defined in a regular expression
- **Feature group:** Default or user-defined group of features

These parameters create a hierarchical relationship. The most basic control parameter is the command. The next control parameter is the feature, which represents all commands that are associated with the feature. The last control parameter is the feature group. The feature group combines related features and allows you to easily manage the rules. The Cisco NX-OS Software also supports the predefined feature group L3 that you can use.

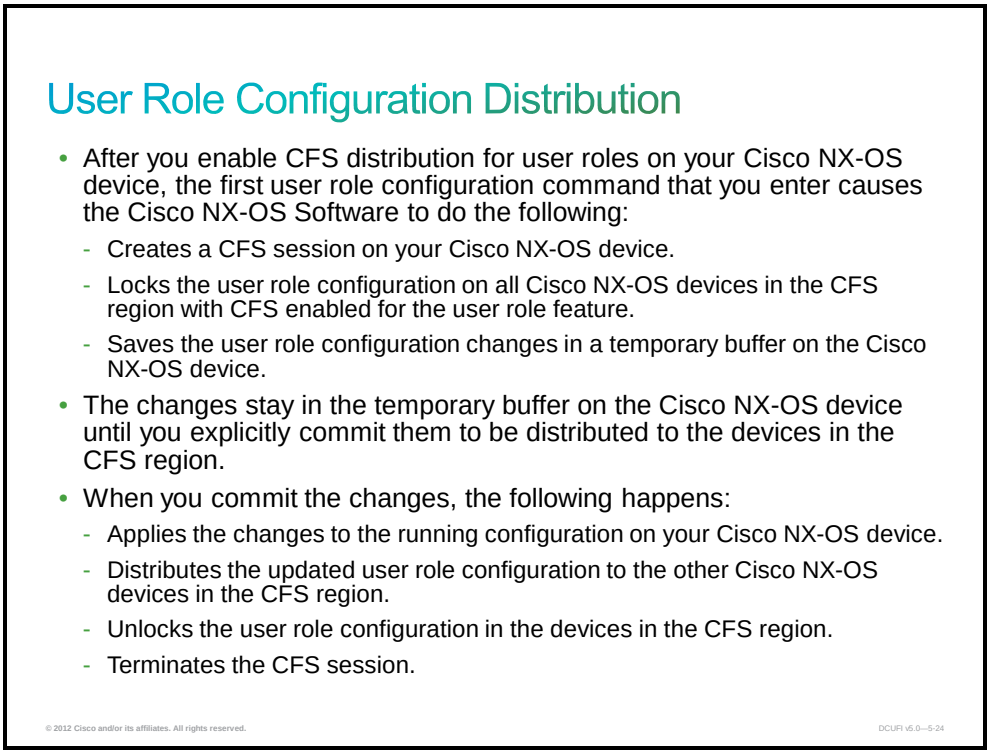
Each user role can contain multiple rules and each user can have multiple roles. For example, if role A allows access only to configuration operations, and role B allows access only to debug operations, then users who belong to both role A and role B can access configuration and debug operations. You can also limit access to specific VLANs, virtual routing and forwarding instances (VRFs), and interfaces.

User accounts and role-based access control (RBAC) have the following configuration guidelines and limitations:

- On the Cisco Nexus 7000 Series Switches, you can create up to 64 user-defined roles in a VDC in addition to the four default user roles in the default VDC and the two default user roles in the nondefault VDCs. On the Cisco Nexus 5000 and 5500 Platform switches, you can create up to 64 user-defined roles.
- You can add up to 256 rules to a user role.
- You can add up to 64 user-defined feature groups to a VDC in addition to the default feature group, L3. Feature groups can be configured on the Cisco Nexus 5000 and 5500 Platform switches.

https://t.me/learningnets

- You can configure up to 256 users in a VDC. The Cisco Nexus 5000 and 5500 Platform switches support up to 256 users.
- You can assign a maximum of 64 user roles to a user account.
- If you have a user account that is configured on the local Cisco NX-OS device that has the same name as a remote user account on an AAA server, the Cisco NX-OS Software applies the user roles for the local user account to the remote user, not the user roles that are configured on the AAA server.
- RBAC is not supported for traffic between Cisco Nexus 7000 F1-Series module ports and Cisco Nexus 7000 M-1 Series module ports in the same VLAN (Cisco Nexus 7000 Series Switches only).



Cisco Fabric Services allows the Cisco NX-OS device to distribute the user role configuration to other Cisco NX-OS devices in the network. When you enable Cisco Fabric Services distribution for a feature on your device, the device belongs to a Cisco Fabric Services region containing other devices in the network that you have also enabled for Cisco Fabric Services distribution for the feature. Cisco Fabric Services distribution for the user role feature is disabled by default.

Note	You must explicitly enable Cisco Fabric Services for user roles on each device to which you want to distribute configuration changes.

© 2012 Cisco Systems, Inc.
https://t.me/learningnets

User Role Configuration Distribution (Cont.)

```
N7010-C1(config)# role distribute

N7010-C1(config)# show role status
Distribution: Enabled
Session State: Unlocked

N7010-C1(config)# role commit
```

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

To distribute the user roles configuration to other Cisco NX-OS devices in the network, you must first enable Cisco Fabric Services distribution for user roles.

You can apply the user role global and, or, server configuration that is stored in the temporary buffer to the running configuration across all switches in the fabric (including the originating switch).

You can discard the temporary database of user role changes and end the Cisco Fabric Services distribution session.

https://t.me/learningnets

Configuring User Accounts

```
N7010-C1(config)# username admin password 4Ty18Rnt role network-admin

N7010-C1(config)# show user-account
user:admin
        this user account has no expiry date
        roles:network-admin


N7010-C1(config)# show role
role: network-admin
  description: Predefined network admin role has access to all commands
  on the switch
  ------------------------------------------------------------------
  Rule    Perm    Type        Scope               Entity
  ------------------------------------------------------------------
  1       permit  read-write
role: network-operator
```

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

You can create a maximum of 256 user accounts on a Cisco NX-OS device. User accounts have the following attributes:

- Username
- Password
- Expiry date
- User roles

You can enter the password in cleartext format or encrypted format. The Cisco NX-OS password encrypts cleartext passwords before saving them to the running configuration. Encrypted format passwords are saved to the running configuration without further encryption.

The expire date option format is YYYY-MM-DD. The default is no expiry date.

User accounts can have a maximum of 64 user roles. The user can determine what commands are available by using the CLI context-sensitive help utility.

Note	Changes to user account attributes do not take effect until the user logs in and creates a new session.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

Creating User Roles and Rules

```
N7010-C1(config)# role name RoleA
N7010-C1(config-role )# rule 1 rule 1 permit read-write feature-group L3
N7010-C1(config-role )# description This is VRF admin role
N7010-C1(config-role-vlan)# vlan policy deny
N7010-C1(config-role-vlan)# permit vlan 1-100
N7010-C1(config-role-vlan)# vrf policy deny (Cisco Nexus 7000 Series only)
N7010-C1(config-role-vrf)# permit vrf INTERSWITCH (Cisco Nexus 7000 Series only)
N7010-C1(config-role-vrf)# permit vrf INTERVDC (Cisco Nexus 7000 Series only)
N7010-C1(config-role-vrf)# interface policy deny
N7010-C1(config-role-interface)# permit interface ethernet 1/1-2
N7010-C1(config-role-interface)# permit interface port-channel 1

N7010-C1(config-role-interface)# show role name L3-ADMIN

Role: RoleA
  Description: This is VRF admin role
  Vlan policy: deny
  Permitted vlans: 1-100
  Interface policy: deny
  Permitted interfaces:
  port-channel1
  Ethernet1/1-2
  Vrf policy: deny
  Permitted vrfs: INTERSWITCH,INTERVDC
  ----------------------------------------------------------------------
  Rule    Perm    Type        Scope               Entity
  ----------------------------------------------------------------------
  1       permit  read-write  feature-group       L3
```

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

You can configure up to 256 rules for each role. The user-specified rule number determines the order in which the rules are applied. Rules are applied in descending order. For example, if a role has three rules, rule 3 is applied before rule 2, which is applied before rule 1.

Note Regardless of the read-write rule that is configured for a user role, some commands can be executed only through the predefined network-admin and vdc-admin roles.

You can create custom feature groups to add to the default list of features that are provided by the Cisco NX-OS Software. These groups contain one or more of the features. You can create up to 64 feature groups (per VDC on the Cisco Nexus 7000 Series Switches).

Note You cannot change the default feature group L3.

You can change a user role interface policy to limit the interfaces that the user can access. By default, a user role allows access to all interfaces. (On the Cisco Nexus 7000 Series Switches this is in the VDC.)

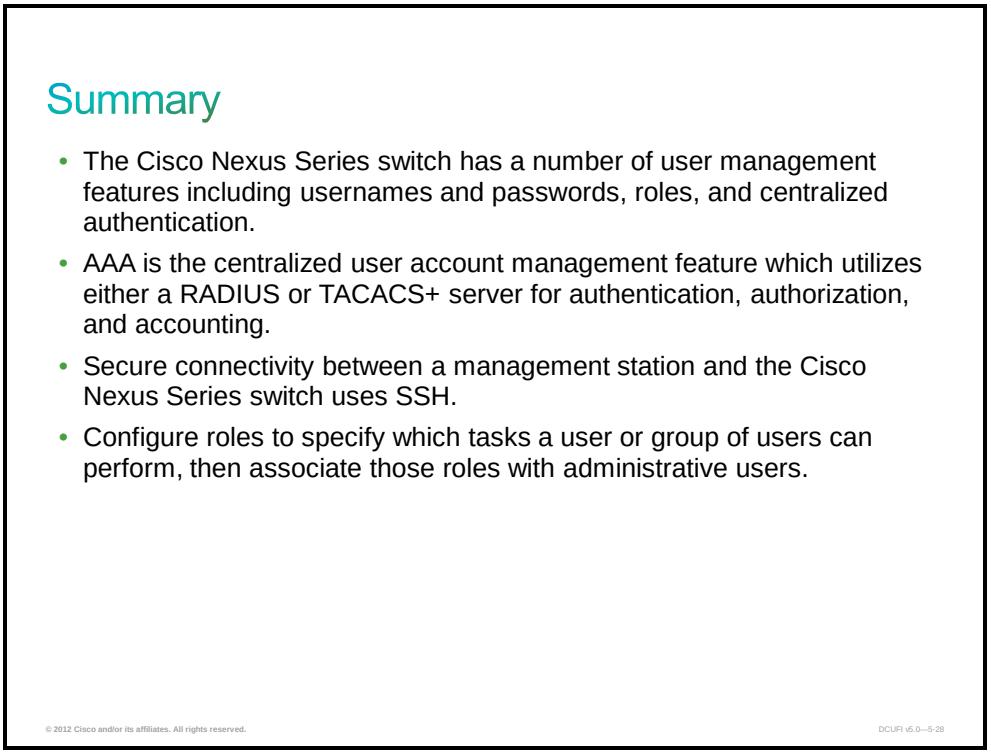
You can change a user role VLAN policy to limit the VLANs that the user can access. By default, a user role allows access to all VLANs. (On the Cisco Nexus 7000 Series Switches this is in the VDC).

You can change a user role VRF policy to limit the VRFs that the user can access. By default, a user role allows access to all VRFs in the VDC on the Cisco Nexus 7000 Series Switches.

https://t.me/learningnets

Summary

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



© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Lesson 3

Describing System Management

Overview

System management and troubleshooting are critical parts of data center maintenance. The Cisco Nexus Family of switches has a system management feature with many features—Cisco Fabric Services, Network Time Protocol (NTP), Cisco Discovery Protocol, system message logging, Cisco Smart Call Home, scheduler, and Simple Network Management Protocol (SNMP). In this lesson, you will learn to identify which system management features are available on the Cisco Nexus Family of switches and how to configure those features to support the management infrastructure requirements on the switches.

Objectives

Upon completing this lesson, you will be able to identify which system management features are available on the Cisco Nexus Family of switches and explain how to apply licensing and configure the switches to support the network requirements. This ability includes being able to meet these objectives:

- Describe which system management features are available on the Cisco Nexus Family of switches and where they would be used in managing the system
- Configure the Cisco Fabric Services on the Cisco Nexus Family of switches
- Configure the NTP and the PTP on the Cisco Nexus Family of switches
- Configure and use the EEM to respond to events that occur on the Cisco Nexus 7000 Series Switches
- Configure NetFlow on the Cisco Nexus Family of switches
- Configure Smart Call Home on the Cisco Nexus Family of switches
- Configure the scheduler on the Cisco Nexus Family of switches
- Use and configure the SPAN feature on the Cisco Nexus 7000 Series Switches
- Configure SMNP on the Cisco Nexus Family of switches
- Use the XML interface to manage the Cisco Nexus Family of switches
- Implement Cisco DCNM for LAN to enable monitoring and managing the Cisco Nexus Family of switches

https://t.me/learningnets

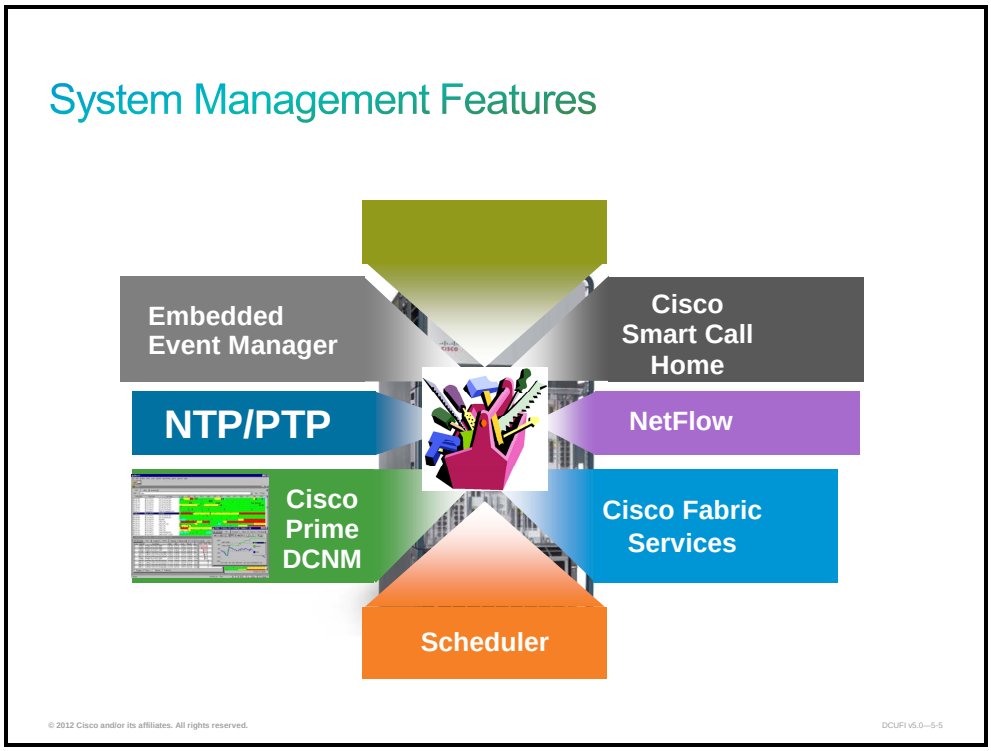
- Integrate Cisco Nexus with VMware vCenter Manager
- Apply and manage the licenses on the Cisco Nexus Family of switches
- Identify when the firmware and EPLDs should be upgraded and how to perform that task

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

System Management Features

This topic describes which system management features are available on the Cisco Nexus Family of switches and where they would be used in managing the system.



Cisco Fabric Services

Cisco Fabric Services is a Cisco proprietary feature that distributes data, including configuration changes, to all Cisco Nexus Operating System (NX-OS) devices in a network.

Network Time Protocol

NTP synchronizes the time of day among a set of distributed time servers and clients. This feature correlates time-specific information, such as system logs, received from the devices in your network.

Cisco Smart Call Home

Cisco Smart Call Home provides an email-based notification of critical system policies. The Cisco NX-OS Software provides a range of message formats for optimal compatibility with pager services, standard email, or XML-based automated parsing applications. You can use this feature to page a network support engineer, email a network operations center, or you can use Smart Call Home services to automatically generate a case with the Cisco Technical Assistance Center (TAC).

Scheduler

The scheduler allows you to create and manage jobs such as routinely backing up data or making quality of service (QoS) policy changes. The scheduler can start a job according to your needs—only once at a specified time or at periodic intervals.

https://t.me/learningnets

Simple Network Management Protocol

SNMP is an application-layer protocol that provides a message format for communication between SNMP managers and agents. SNMP provides a standardized framework and a common language that is used for the monitoring and management of devices in a network.

Cisco Prime Data Center Network Manager

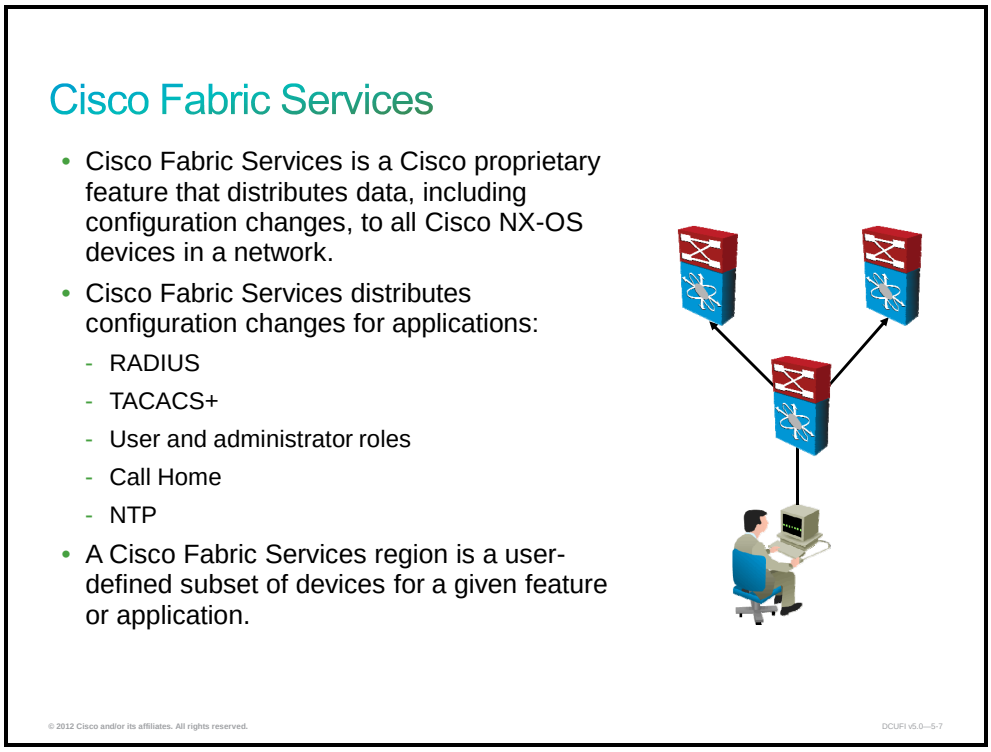
Cisco Prime Data Center Network Manager (DCNM) is a management solution that maximizes overall data center infrastructure uptime and reliability, which improves business continuity.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Configuring Cisco Fabric Services

Cisco Fabric Services is a Cisco proprietary feature that distributes data, including configuration changes, to all Cisco NX-OS devices including Cisco Nexus 7000 Series Switches and Cisco Nexus 5000 Series Switches (which include the 5000 and 5500 Platform switches) that can be found in a network. This topic discusses configuring this feature.



Cisco Fabric Services

Cisco Fabric Services is used to distribute and synchronize a configuration on one Cisco device or with all other Cisco devices in the network. The distribution of the configuration is performed using Cisco Fabric Services over IP, which provides consistent, and in most cases identical, configurations and behavior in the network.

When a new device is detected in the network, Cisco Fabric Services manages the merging, or synchronizing, of its configuration with that of the other devices. Cisco Fabric Services will coordinate and minimize the number of merges by designating one device to manage merges per application per region.

During a merger of two networks, the designated managers exchange configuration databases. The application on one of them merges the database, decides if the merger is successful, and notifies all other devices.

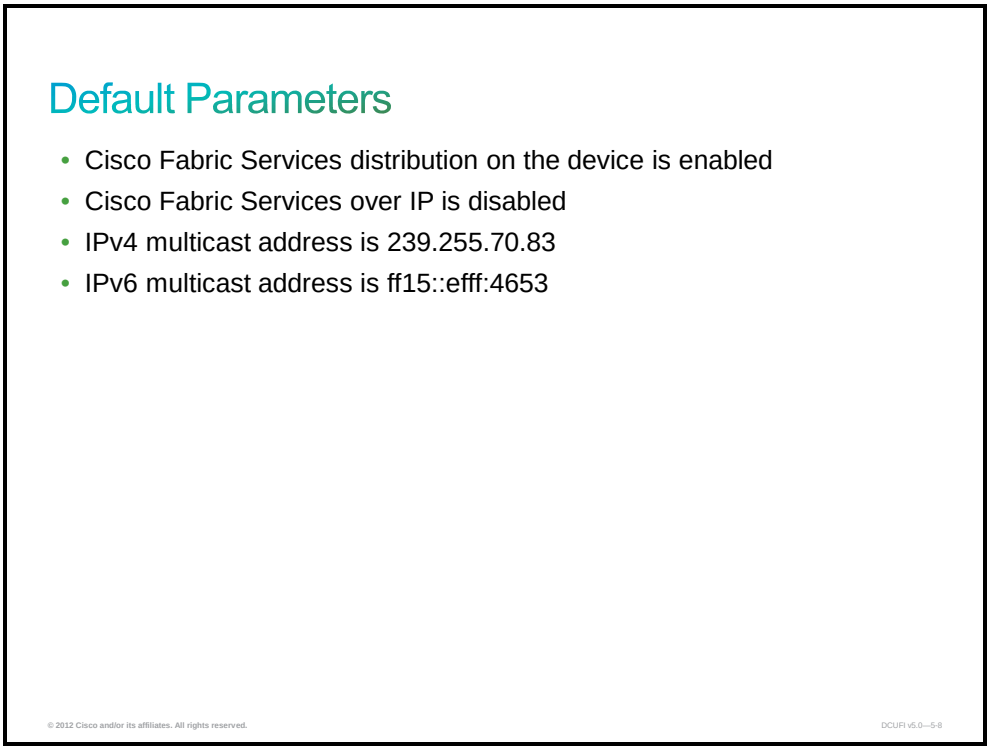
If the merge is successful, the merged database is distributed to all devices in the combined fabric and the entire new fabric emerges in a consistent state.

If the merge fails, you can start a distribution from any device in the new fabric. This distribution will restore all peers in the fabric to the same configuration database.

Cisco Fabric Services Region

A Cisco Fabric Services region is a user-defined subset of devices for a given feature or application. You usually define regions to localize or restrict distribution that is based on devices that are close to one another. When a network covers many geographies with many different administrators who are responsible for subsets of devices, you can manage the physical scope of an application by setting up a Cisco Fabric Services region.

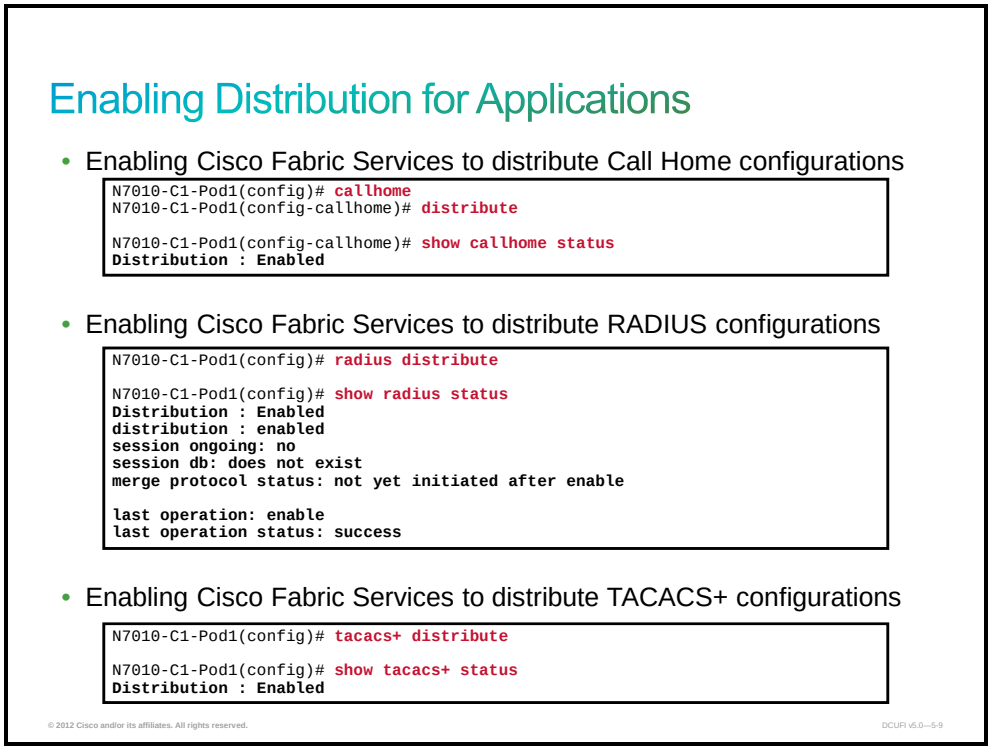
https://t.me/learningnets



The figure lists the default settings for Cisco Fabric Services parameters.

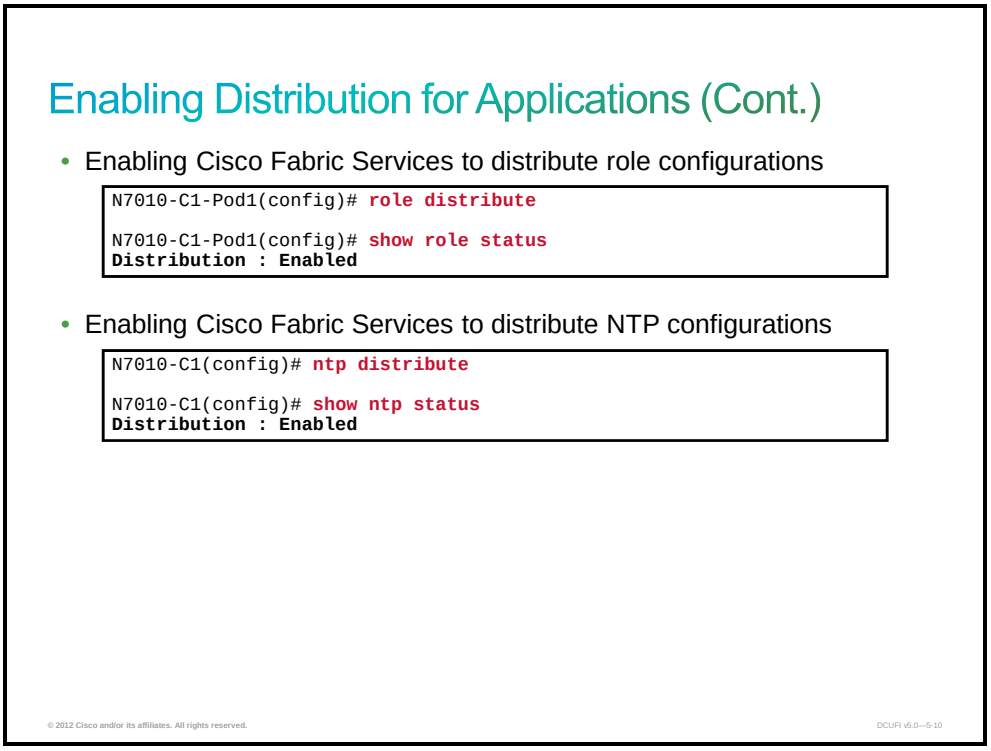
When Cisco Fabric Services performs the distribution and synchronization over IP, a multicast address is used, which is recognized by all Cisco Fabric Services-enabled devices. Cisco Fabric Services over IP is supported over IP version 4 (IPv4) and IP version 6 (IPv6).

© 2012 Cisco Systems, Inc.
https://t.me/learningnets



The figure describes the configuration steps that are used to perform an initial setup for Cisco Fabric Services over IP for Cisco Smart Call Home, RADIUS, and TACACS+ configurations. Cisco Fabric Services over IP distributes and synchronizes a configuration on one Cisco device or with all other Cisco devices in the network. The use of Cisco Fabric Services over IP provides consistent, and in most cases identical, configurations and behavior in the network.

https://t.me/learningnets



The figure describes the configuration steps that should be used to perform an initial setup for Cisco Fabric Services over IP for role and NTP configurations.

The **show callhome status**, **show radius status**, **show tacacs+ status**, **show role status**, and **show ntp status** commands validate the Cisco Fabric Services distribution for applications.

The following steps provide an overview of how Cisco Fabric Services distributes application configurations.

Step 1 Enable Cisco Fabric Services to distribute configurations for an application, such as Smart Call Home.

Step 2 Enter a command to change the configuration for a Cisco Fabric Services application, such as Smart Call Home.

Step 3 Cisco Fabric Services checks to determine if an active fabric lock indicates that a configuration change is already in progress for this application.

Note Only one Cisco Fabric Services session for an application can be active at a time. Cisco Fabric Services uses locks to enforce this restriction. Distribution is not allowed to start if locks are in place for the application anywhere else in the fabric.

Step 4 One of the following occurs:

- If an active fabric lock exists for this application, Cisco Fabric Services rejects the command. No changes are permitted until the existing fabric lock is released.
- If there is not an active fabric lock for this application, then Cisco Fabric Services starts a session and locks the fabric for this application.

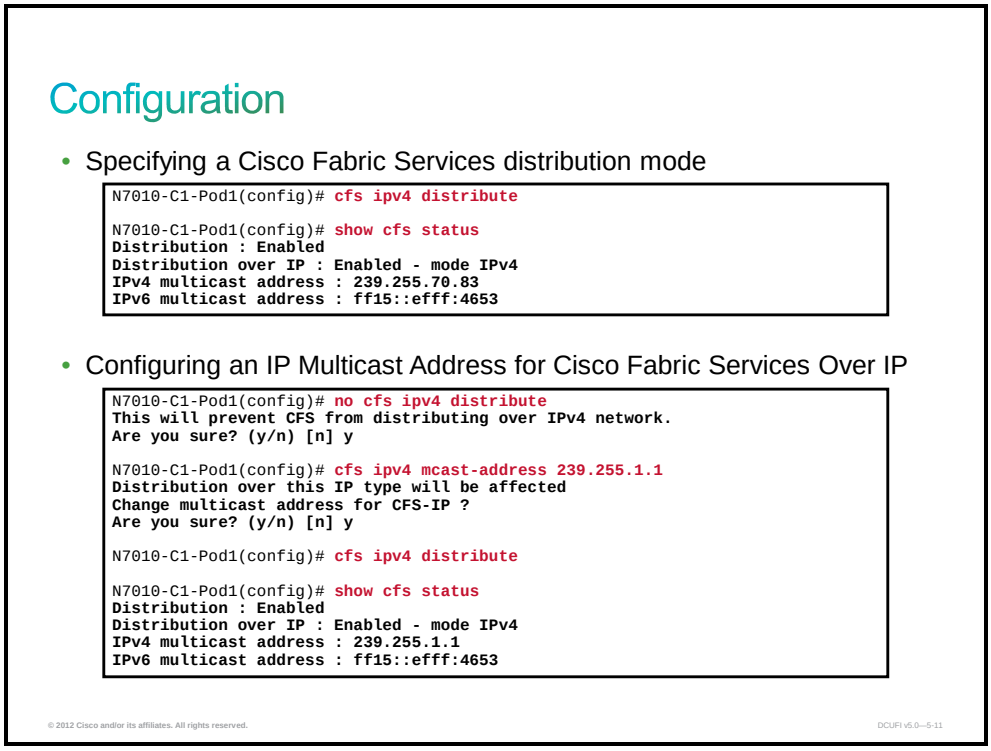
Step 5 You enter the remaining configuration commands for the application.

Step 6 You commit the configuration by using the **commit** command.

Cisco Fabric Services distributes the configuration and releases the lock.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets



The figure describes the steps that are necessary to specify and enable a Cisco Fabric Services distribution mode (Ethernet, IPv4, or IPv6) and to modify an IP multicast address for Cisco Fabric Services over IP.

For Cisco Fabric Services protocol-specific distributions, such as the keepalive mechanism for detecting network topology changes, use the IP multicast address to send and receive information.

The **show cfs status** command output validates the Cisco Fabric Services distribution method and the IP multicast address for Cisco Fabric Services over IP.

Note	Cisco Fabric Services over IP must be disabled before you can change the multicast address.

https://t.me/learningnets

Configuration (Cont.)

- Clearing a locked session

```
N7010-C1-Pod1(config)# show ntp status
Distribution : Enabled
Last operational state: Fabric Locked

N7010-C1-Pod1(config)# clear ntp session

N7010-C1-Pod1(config)# show ntp status
Distribution : Enabled
Last operational state: No session
```

- Discarding a configuration

```
N7010-C1-Pod1(config)# ntp abort
This will prevent CFS from distributing the
configuration to other switches.
Are you sure? (y/n)  [n] y

N7010-C1-Pod1(config)# show ntp session status
Last Action Time Stamp : Wed Nov 12 16:07:25 2008
Last Action : Abort
Last Action Result : Success
Last Action Failure Reason : none
```

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

The figure describes how to clear a lock that is held by an application from any device in the fabric. The second figure describes how to discard the Cisco Fabric Services configuration.

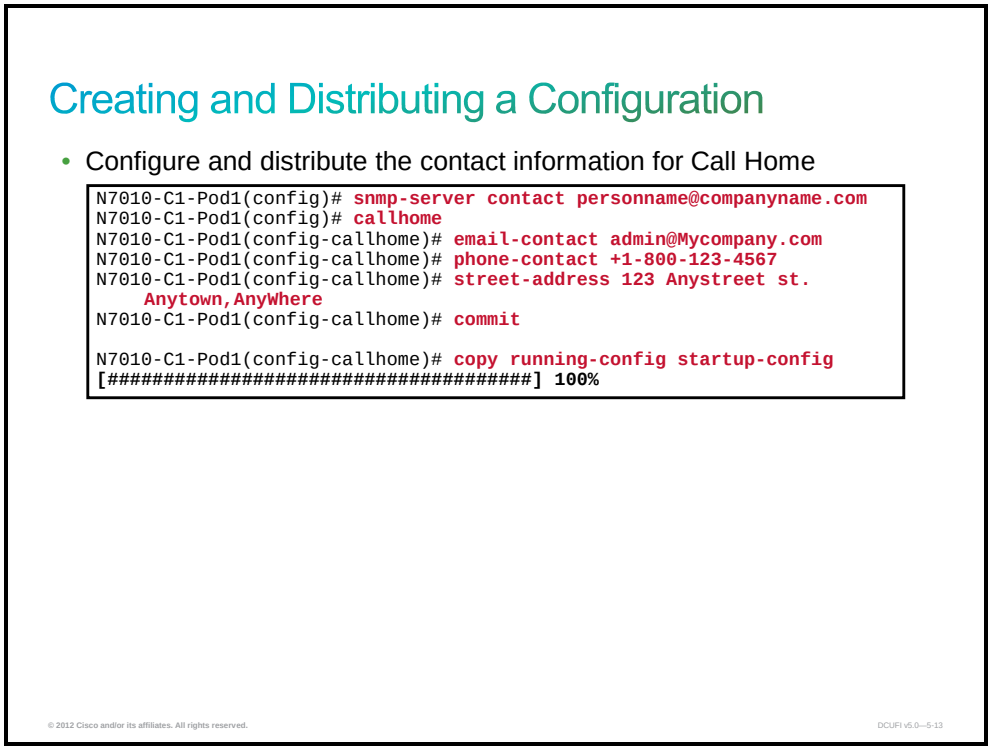
The **show***application name***status** command verifies Cisco Fabric Services status.

You must have administrative permissions to release a lock.

When you clear a lock in the fabric, any pending configurations in any device in the fabric are discarded.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets



The figure describes how to create a configuration change for an application and then distribute it to its application peers.

Caution	If you do not commit the changes, they are not distributed and saved in the running configuration of application peer devices.
Caution	If you do not save the changes to the startup configuration in every application peer device where distributed, then changes are retained only in their running configurations.

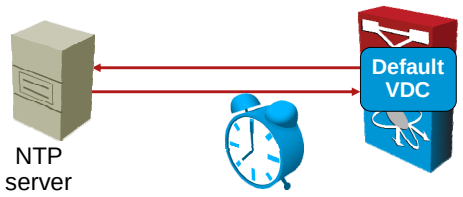
https://t.me/learningnets

Configuring NTP and PTP

This topic identifies how to configure the NTP on Cisco Nexus 7000 and 5000 Series Switches and the Precision Time Protocol (PTP) on the Cisco Nexus 7000 Series Switches.

Network Time Protocol

- NTP synchronizes the time of day among a set of distributed time servers and clients.
- NTP must be configured in the default VDC on the Cisco Nexus 7000 Series
- NTP uses the default VRF if you do not configure a specific VRF for the NTP server and NTP peer.



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

NTP synchronizes the time of day among a set of distributed time servers and clients so that you can correlate events when you receive system logs and other time-specific events from multiple network devices.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Guidelines and Limitations

- You should have a peer association with another device only when you are sure that your clock is reliable (which means that you are a client of a reliable NTP server).
- A peer configured alone takes on the role of a server and should be used as a backup. If you have two servers, you can configure several devices to point to one server and the remaining devices to point to the other server. You can then configure a peer association between these two servers to create a more reliable NTP configuration.
- If you have only one server, you should configure all the devices as clients to that server.
- You can configure up to 64 NTP entities (servers and peers).
- If you configure NTP in a VRF, ensure that the NTP server and peers can reach each other through the configured VRFs.

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

The figure lists guidelines and limitations for when you configure NTP on a Cisco Nexus switch.

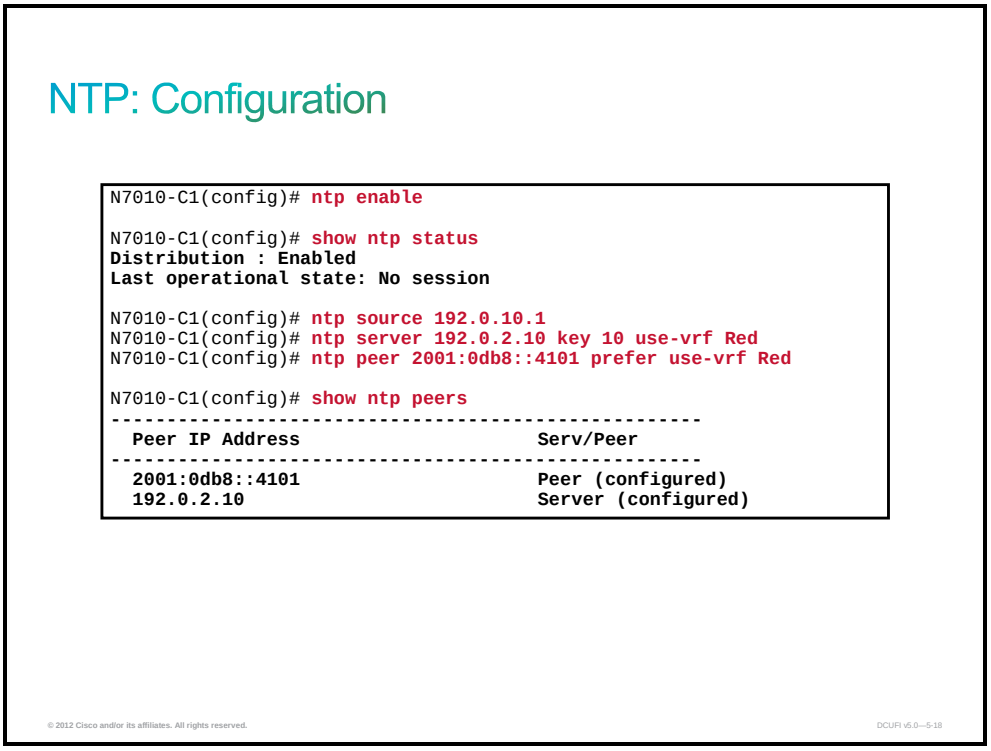
Guidelines and Limitations (Cont.)

- If Cisco Fabric Services is disabled for NTP, then NTP does not distribute any configuration and does not accept a distribution from other devices in the network.
- After Cisco Fabric Services distribution is enabled for NTP, the entry of an NTP configuration command locks the network for NTP configuration until a commit command is entered. During the lock, no changes can be made to the NTP configuration by any other device in the network except the device that initiated the lock.
- If you use Cisco Fabric Services to distribute NTP, all devices in the network should have the same VRFs configured as you use for NTP.
- You must manually distribute NTP authentication keys on the NTP server and Cisco NX-OS devices across the network.

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

The figure lists more guidelines and limitations for when you configure NTP on the Cisco Nexus Series switch.

https://t.me/learningnets



The figure describes the steps that are necessary for configuring NTP. NTP is enabled by default. You can disable NTP and then reenable it. You can configure an NTP server and peer.

Use the **key** keyword to configure a key to be used while communicating with the NTP peer. The range for the key-id argument is from 1 to 65535.

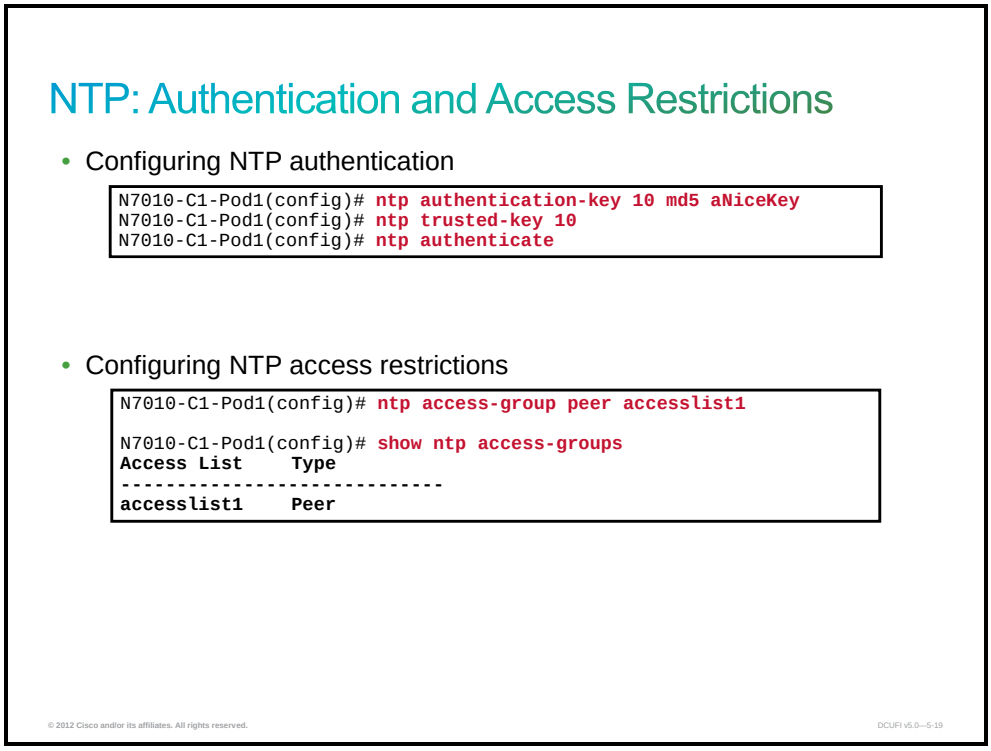
Use the **prefer** keyword to make this the preferred NTP server for the device.

On the Cisco Nexus 7000 Series Switches use the **use-vrf** keyword to configure the NTP server to communicate over the specified virtual routing and forwarding (VRF). The vrf-name argument can be default, management, or any case-sensitive alphanumeric string up to 32 characters.

NTP sets the source IP address for all NTP packets based on the address of the interface through which the NTP packets are sent. You can configure NTP to use a specific source IP address.

The **show ntp peers** command verifies the status of configured NTP peers.

© 2012 Cisco Systems, Inc.
https://t.me/learningnets



The figure describes the configuration steps that should be used to perform NTP authentication and apply access restrictions.

Before you begin, make sure that you configured the NTP server with the authentication keys that you plan to specify in this procedure.

Command **ntp authentication-key** defines the authentication keys. The device does not synchronize to a time source unless the source has one of these authentication keys and the key number is specified by the **ntp trusted-key** number command.

The range for authentication keys is from 1 to 65535. For the MD5 string, you can enter up to eight alphanumeric characters.

You can control access to NTP services by using access groups. Specifically, you can specify the types of requests that the device allows and the peers from which it accepts responses.

The **peer** keyword allows time requests and NTP control queries and allows the Cisco NX-OS device to synchronize only to a remote device whose IP address passes the access list criteria.

If you do not configure any access groups, NTP access is granted to all devices. If you configure any access groups, NTP access is granted only to the remote device whose source IP address passes the access list criteria.

The **show ntp access-group** command verifies NTP access restrictions.

https://t.me/learningnets

Precision Time Protocol

- PTP is a time synchronization protocol for distribution of time information through entire network
- Provides greater accuracy than other time synchronization protocols
- PTP can work on both PTP and non-PTP enable devices.
- PTP devices are:
  - Ordinary clock
  - Boundary clock and
  - Transparent clock
- Non-PTP devices are:
  - Ordinary network switches
  - Routers
  - Any other infrastructure equipment
- Cisco NX-OS supports multiple instances of PTP, one instance per virtual device context

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

PTP is a time synchronization protocol for nodes that are distributed across a network. Its hardware time-stamp feature provides greater accuracy than other time synchronization protocols such as NTP.

A PTP system can consist of a combination of PTP and non-PTP devices. PTP devices include ordinary clocks, boundary clocks, and transparent clocks. Non-PTP devices include ordinary network switches, routers, and other infrastructure devices.

PTP is a distributed protocol that specifies how real-time PTP clocks in the system synchronize with each other. These clocks are organized into a master-member synchronization hierarchy with the grand master clock, the clock at the top of the hierarchy, determining the reference time for the entire system. Synchronization is achieved by exchanging PTP timing messages, with the members using the timing information to adjust their clocks to the time of their master in the hierarchy. PTP operates within a logical scope that is called a PTP domain.

The following clocks are common PTP devices:

- **Ordinary clock:** Communicates with the network based on a single physical port, similar to an end host. An ordinary clock can function as a grand master clock.
- **Boundary clock:** Typically has several physical ports, with each port behaving like a port of an ordinary clock. However, each port shares the local clock, and the clock data sets are common to all ports. Each port decides its individual state, either master (synchronizing other ports that are connected to it) or member (synchronizing to a downstream port), based on the best clock that is available to it through all of the other ports on the boundary clock. Messages that are related to synchronization and establishing the master-member hierarchy terminate in the protocol engine of a boundary clock and are not forwarded.
- **Transparent clock:** Forwards all PTP messages like an ordinary switch or router but measures the residence time of a packet in the switch (the time that the packet takes to traverse the transparent clock) and in some cases the link delay of the ingress port for the packet. The ports have no state because the transparent clock does not need to synchronize to the grand master clock.

https://t.me/learningnets

Guidelines and Limitations

- In Cisco NX-OS Release 5.2, PTP operates only in boundary clock mode; end-to-end and peer-to-peer transparent clock are not supported
- Only one PTP process can control all of the port clocks through the clock manager
- PTP supports only multicast communication. Negotiated unicast communication is not supported.
- PTP is limited to a single domain per network.
- PTP can be enabled only on F1 and F2 Series module ports.
- All management messages are forwarded on ports on which PTP is enabled. Handling management is not supported.
- PTP supports transport over UDP. Transport over Ethernet is not supported

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

The figure lists guidelines and limitations for when you configure PTP on the Cisco Nexus 7000 Series Switches.

PTP: Configuration

```
N7010-C1(config)# feature ptp

N7010-C1(config)# ptp source 192.0.10.1
N7010-C1(config)# ptp priority1 1
N7010-C1(config)# ptp priority2 1

N7010-C1(config)# show ptp clock
PTP Device Type: Boundary clock
Clock Identity : 0:22:55:ff:ff:79:a4:c1
Clock Domain: 0
Number of PTP ports: 0
Priority1 : 1
Priority2 : 1
Clock Quality:
Class : 248
Accuracy : 254
Offset (log variance) : 65535
Offset From Master : 0
Mean Path Delay : 0
Steps removed : 0
Local clock time: Thr May 3 14:13:24 2012
```

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

The figure describes the steps that are necessary to configure PTP. PTP is not enabled by default. You can enable PTP using the command **feature ptp**.

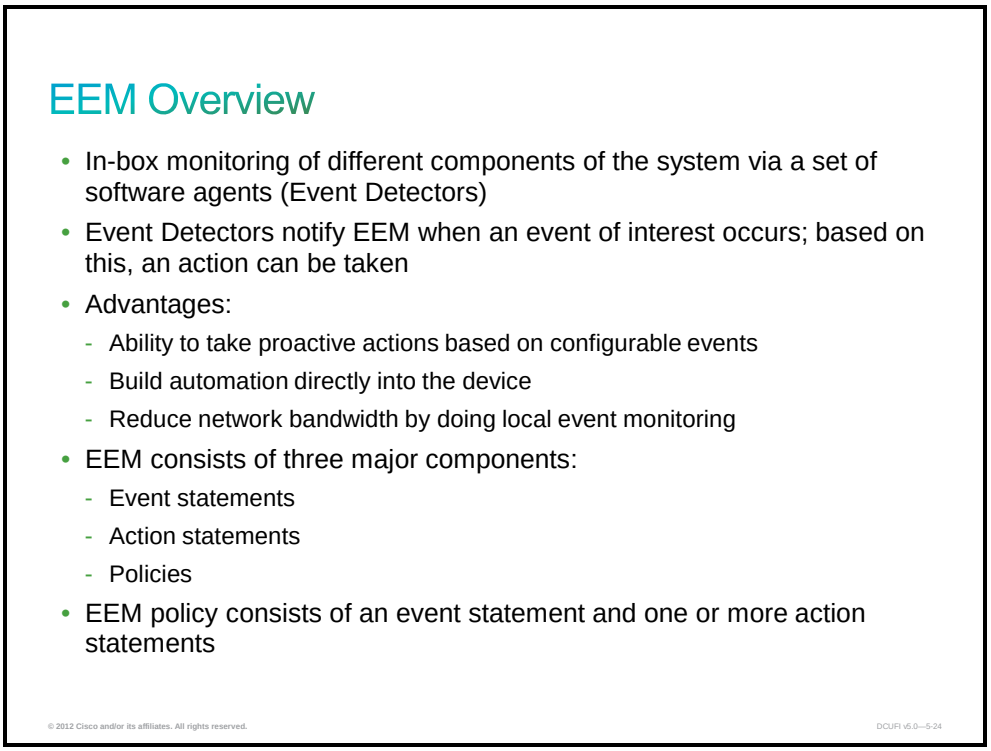
Use the **ptp source** keyword to configure the source IP address for all PTP packets. The IP address can be in IPv4 or IPv6 format.

Use the **ptp priority1** and **ptppriority2** keywords to override the default criteria (clock quality, clock class, and so on) for the best master clock selection for your situation.

https://t.me/learningnets

Configuring EEM

This topic explains how to use and configure the Embedded Event Management (EEM) to respond to events that occur on the Cisco Nexus 7000 Series Switches.



EEM monitors events that occur on your device and takes action to recover or troubleshoot these events, which are based on your configuration. The monitoring is done using different components of the system via a set of software agents (Event Detectors). Event Detectors notify EEM when an event of interest occurs; based on this notification, an action can be taken.

EEM consists of three major components:

- **Event statements:** Events to monitor from another Cisco NX-OS component that may require some action, workaround, or notification
- **Action statements:** An action that EEM can take, such as sending an email, or disabling an interface, to recover from an event
- **Policies:** An event paired with one or more actions to troubleshoot or recover from the event

An EEM policy consists of an event statement and one or more action statements. The event statement defines the event to look for as well as the filtering characteristics for the event. The action statement defines the action EEM takes when the event occurs.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Guidelines and Limitations

- The maximum number of configurable EEM policies is 500
- Action statements within your user policy or overriding policy should not negate each other or adversely affect the associated system policy
- An override policy that consists of an event statement and no action statement triggers no action and no notification of failures
- An override policy without an event statement overrides all possible events in the system policy
- EEM event correlation is supported only on the supervisor module, not on individual line cards
- EEM event correlation is not supported across different modules within a single policy
- EEM event correlation does not override the system default policies
- Default action execution is not supported for policies that are configured with tagged events.

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

The figure lists guidelines and limitations for when you configure EEM on a Cisco Nexus switch.

EEM Actions

- Execute any CLI commands
- Update a counter
- Log an exception.
- Force the shutdown of any module.
- Reload the device.
- Shut down specified modules because the power is over budget.
- Generate a syslog message.
- Generate a Call Home event.
- Generate an SNMP notification.
- Use the default action for the system policy.

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

Action statements describe the action that is triggered by a policy. Each policy can have multiple action statements. If no action is associated with a policy, EEM still observes events but takes no actions.

https://t.me/learningnets

EEM supports the following actions in action statements:

- Execute any CLI commands.
- Update a counter.
- Log an exception.
- Force the shutdown of any module.
- Reload the device.
- Shut down specified modules because the power is over budget.
- Generate a syslog message.
- Generate a Smart Call Home event.
- Generate an SNMP notification.
- Use the default action for the system policy.

EEM Configuration Examples

- Monitoring Interface Shutdown

```
N7010-C1(config)# event manager applet monitorShutdown

N7010-C1(config-applet)# description “Monitors interface shutdown.”
N7010-C1(config-applet)# event cli match “conf t; interface *; shutdown”
N7010-C1(config-applet)# action 1.0 cli show interface e 3/1
N7010-C1(config-applet)#
```

- Monitoring Module Powerdown

```
N7010-C1(config)# event manager applet monitorPoweroff

N7010-C1(config-applet)# description “Monitors module power down.”
N7010-C1(config-applet)# event cli match “conf t; poweroff *”
N7010-C1(config-applet)# action 1.0 cli show module
N7010-C1(config-applet)#
```

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

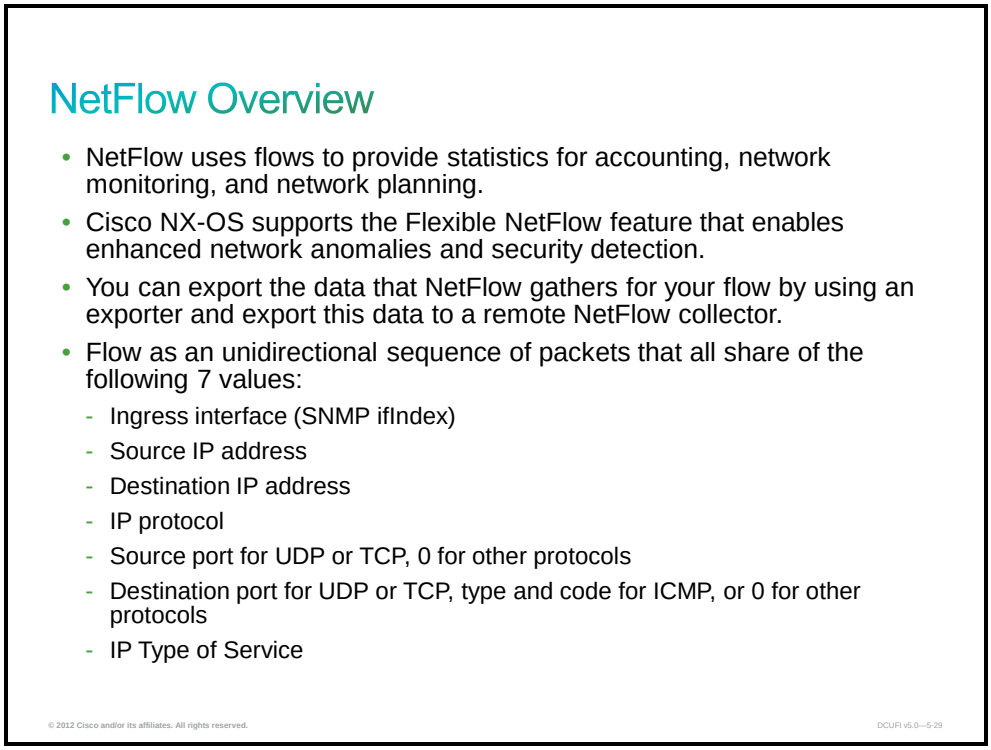
This example shows how to monitor an interface shutdown and module powerdown.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

Configuring NetFlow

This topic describes how to configure NetFlow on the Cisco Nexus 7000 Series Switches.



NetFlow uses flows to provide statistics for accounting, network monitoring, and network planning. A flow is a unidirectional stream of packets that arrives on a source interface (or VLAN) and has the same values for the keys. A key is an identified value for a field within the packet. You create a flow using a flow record to define the unique keys for your flow.

Cisco NX-OS supports the Flexible NetFlow feature that enables enhanced network anomalies and security detection. Flexible NetFlow allows you to define an optimal flow record for a particular application by selecting the keys from a large collection of predefined fields.

Cisco NX-OS supports the version 5 and version 9 export formats. We recommend that you use the version 9 export format for the following reasons:

- Variable field specification format
- Support for IPv6, Layer 2, and Multiprotocol Label Switching (MPLS) fields
- More efficient network utilization

If you configure the version 5 export format, you have these limitations:

- Fixed field specifications
- A 16-bit representation of the 32-bit interface index used in Cisco NX-OS
- No support for IPv6, Layer 2, or MPLS fields

https://t.me/learningnets

Guidelines and Limitations

- You must configure a valid record name for every flow monitor
- Use v9 export to see the full 32-bit SNMP ifIndex values at the NetFlow connector
- The maximum number of supported NetFlow entries is 512K.
- The Cisco Nexus 2000 Series Fabric Extender supports bridged NetFlow (for flows within a VLAN).
- F1 Series ports do not support bridged NetFlow
- NetFlow is not supported on F2 Series modules.
- Only Layer 2 NetFlow is applied on Layer 2 interfaces, and only Layer 3 NetFlow is applied on Layer 3 interfaces.
- A rollback will fail if you try to modify a record that is programmed in the hardware during a rollback.
- You must configure a source interface. If you do not configure a source interface, the exporter will remain in a disabled state.

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

The figure lists guidelines and limitations for when you configure the NetFlow feature on a Cisco Nexus switch.

NetFlow Configuration

```
N7010-C1(config)# feature netflow

N7010-C1(config)# flow exporter nexusexample
N7010-C1(config-flow-exporter)# version 9
N7010-C1(config)# flow record RDR
N7010-C1(config-flow-record)# match ipv4 source address
N7010-C1(config-flow-record)# match ipv4 destination address
N7010-C1(config-flow-record)# collect counter bytes
N7010-C1(config-flow-record)# collect counter packets
N7010-C1(config)# flow monitor Monitor
N7010-C1(config-flow-monitor)# record RDR
N7010-C1(config-flow-monitor)# exporter nexusexample
N7010-C1(config)# interface Ethernet 2/45
N7010-C1(config-if)# ip flow monitor Monitor output
```

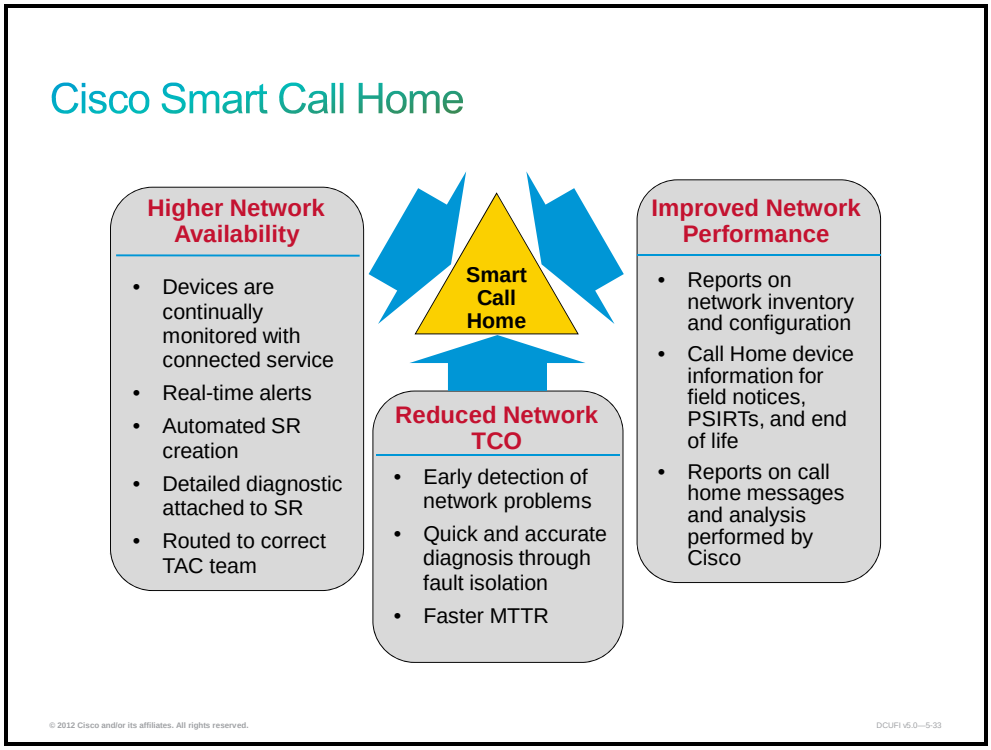
© 2012 Cisco and/or its affiliates. All rights reserved.

The figure shows how to enable NetFlow feature on Cisco Nexus switches. It also shows how to choose an exporter version and what type of records you want to have as statistics and, at the end, how to apply the NetFlow feature on appropriate interface.

https://t.me/learningnets

Configuring Smart Call Home

Cisco Smart Call Home provides an email-based notification for critical system policies. A range of message formats are available for compatibility and pager services, standard email, or XML-based automated parsing applications. This topic discusses configuring this feature on the Cisco Nexus Family of switches.

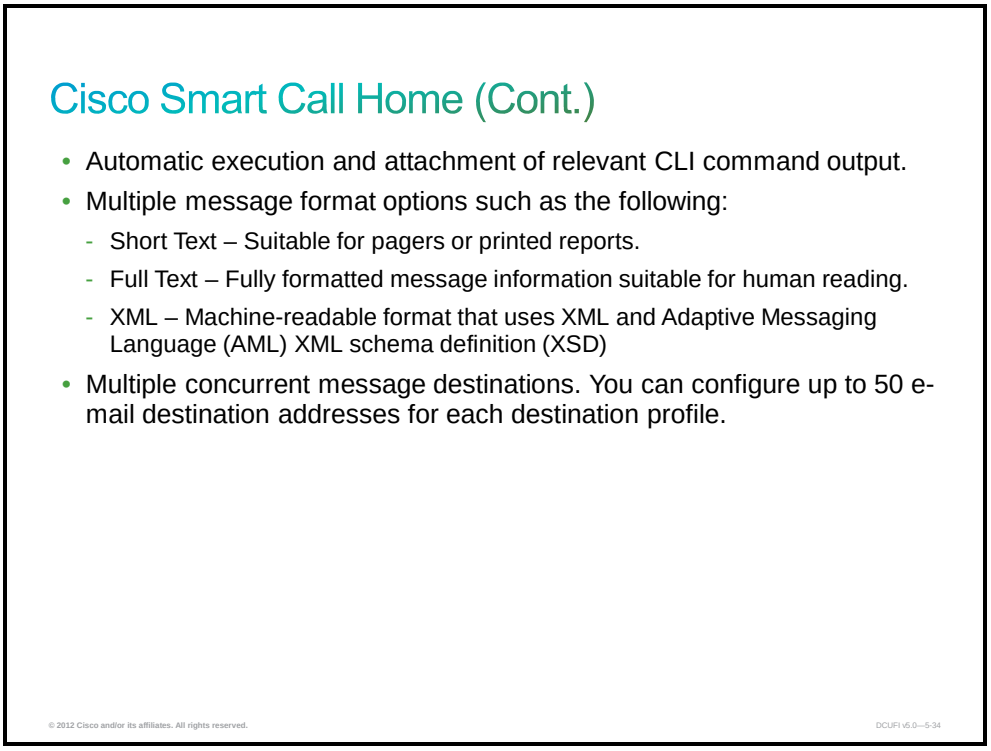


The Cisco Smart Call Home feature continuously monitors hardware and software components to provide interactive technical support. A range of message formats are available for optimal compatibility with pager services, standard email, and XML-based automated parsing applications. Cisco Smart Call Home combines Cisco Generic Online Diagnostics (GOLD) and Cisco IOS Embedded Event Manager (EEM) capabilities to offer detailed diagnostics and real-time alerts, leading to faster resolution of problems.

The Cisco NX-OS Software provides a range of message formats for optimal compatibility with pager services, standard email, or XML-based automated parsing applications that can be used to provide notification of critical system events.

.

https://t.me/learningnets



The figure provides an overview of the Cisco Smart Call Home features.

https://t.me/learningnets

Cisco Smart Call Home: Destination Profiles

- One or more alert groups – The group of alerts that trigger a specific Cisco Smart Call Home message if the alert occurs.
- One or more e-mail destinations – The list of recipients for the Cisco Smart Call Home messages generated by alert groups assigned to this destination profile.
- Message format – The format for the Cisco Smart Call Home message (short text, full text, or XML).
- Message severity level – The Smart Call Home severity level that the alert must meet before Cisco NX-OS generates a Cisco Smart Call Home message to all email addresses in the destination profile.

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

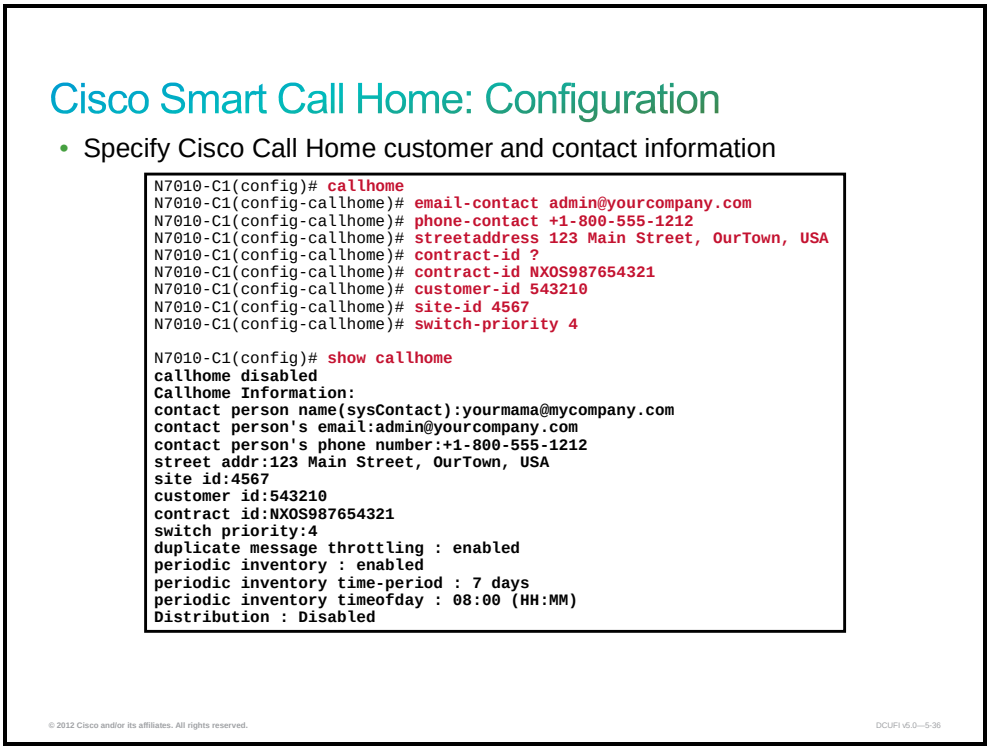
Cisco Smart Call Home alert groups are a predefined subset of Cisco Smart Call Home alerts that are supported in all Cisco NX-OS Software switches. Alert groups permit the selection of Cisco Smart Call Home alerts that you can send to a predefined or custom destination profile.

The destination profiles provide flexibility when you are configuring alert groups, email destinations, message format, and message severity level.

The supported alert groups are as follows:

- Cisco TAC
- Configuration
- Diagnostic
- EEM
- Environmental
- Inventory
- License
- Line module hardware
- Supervisor hardware
- Syslog port group
- System
- Test

https://t.me/learningnets



The figure describes the configuration steps used to perform an initial setup for Cisco Smart Call Home. The details that are specified include email and phone contacts, mailing address, contract, customer, and site IDs.

The **show call home** command provides validation for the Cisco Smart Call Home setup parameters.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Cisco Smart Call Home: Configuration (Cont.)

- Create and validate a Cisco Call Home destination profile

```
N7010-C1(config)# callhome
N7010-C1(config-callhome)# destination-profile bigdatacenter
N7010-C1(config-callhome)# destination-profile bigdatacenter format full-txt

N7010-C1(config)# callhome
N7010-C1(config-callhome)# transport email smtp-server 192.168.200.1 use-vrf ha-maintain
N7010-C1(config-callhome)# transport email from bigdaddy@yourcompany.com
N7010-C1(config-callhome)# transport email reply-to bigtech@yourcompany.com

N7010-C1(config-callhome)# show callhome transport-email
from email addr:bigdaddy@yourcompany.com
reply to email addr:bigtech@yourcompany.com
smtp server:192.168.200.1
smtp server port:25
```

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

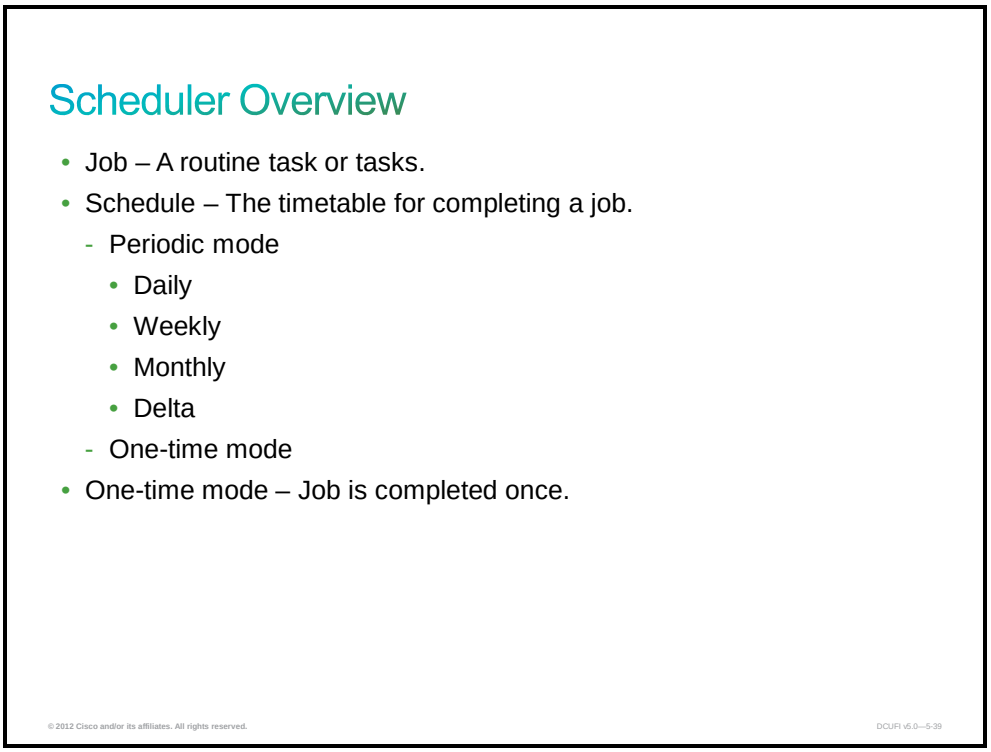
The figure describes the steps that are necessary to create a Cisco Smart Call Home Destination Profile. It specifies the output format, email Simple Mail Transfer Protocol (SMTP) server, and reply-to address.

The **show callhome transport-email** command validates the Cisco Smart Call Home destination profile.

https://t.me/learningnets

Configuring the Scheduler

The scheduler allows you to define and set a timetable for maintenance activities such as QoS policy changes, data backup, and saving a configuration. This topic discusses configuring this feature on the Cisco Nexus Family of switches.



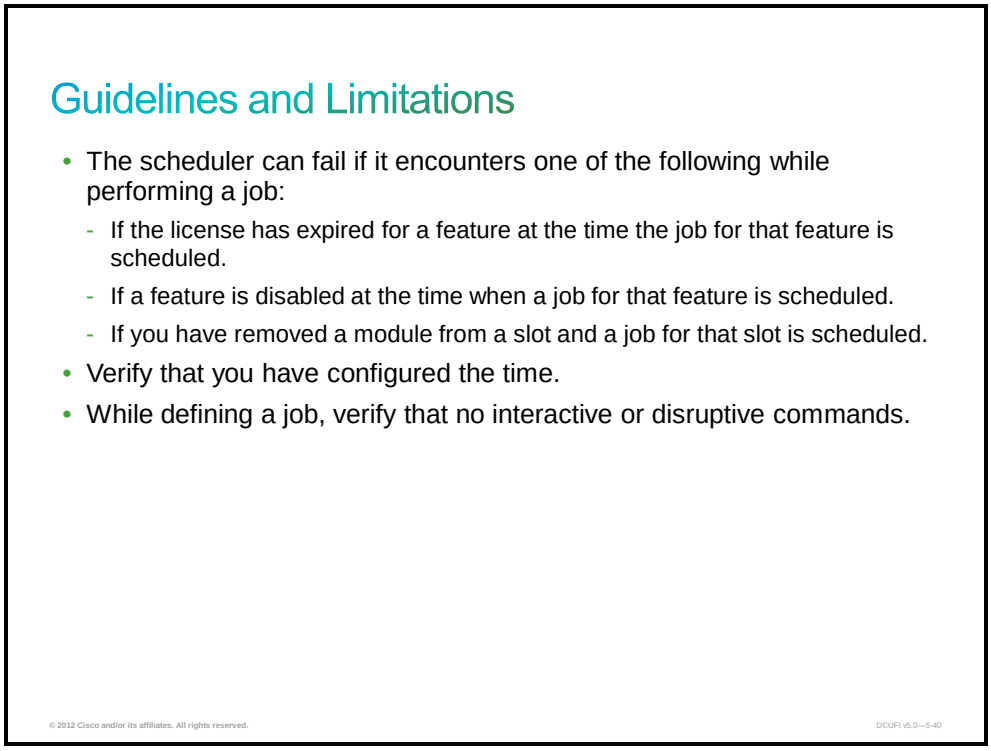
The scheduler defines a job and its timetable as follows:

- **Job:** A routine task or tasks that are defined as a command list and completed according to a specified schedule.
- **Schedule:** The timetable for completing a job. Multiple jobs can be assigned to a schedule, which is defined as either periodic or one-time only.
  - **Periodic mode:** A recurring interval that continues until you delete a job. The following types of interval can be configured:
    - **Daily:** Job is completed once a day
    - **Weekly:** Job is completed once a week
    - **Monthly:** Job is completed once a month
    - **Delta:** Job begins at a specified start time and then at specified intervals (days:hours:minutes)
  - **One-time mode:** Job is completed once at a specified time

.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

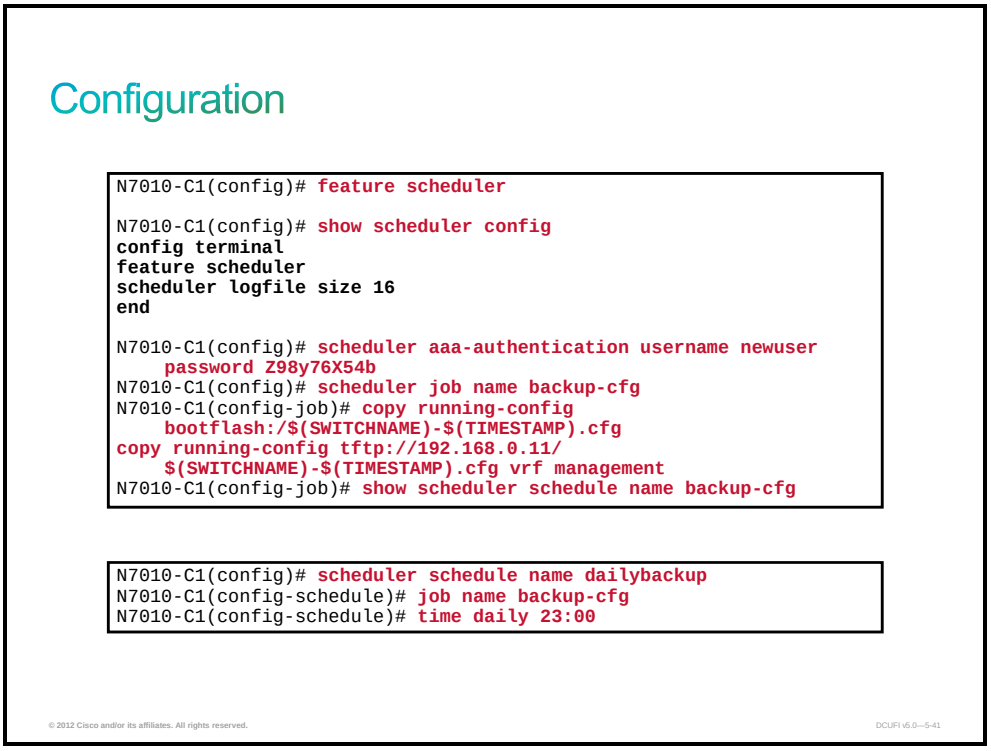


The figure shows the guidelines and limitations for when you configure the scheduler.

If a job fails to verify that the time is configured, the scheduler does not apply a default timetable. If a schedule is created with assigned jobs, but there is no configured time, the job will not start.

When you define a job, there should be no interactive or disruptive commands requiring user intervention because the job must be conducted noninteractively.

https://t.me/learningnets



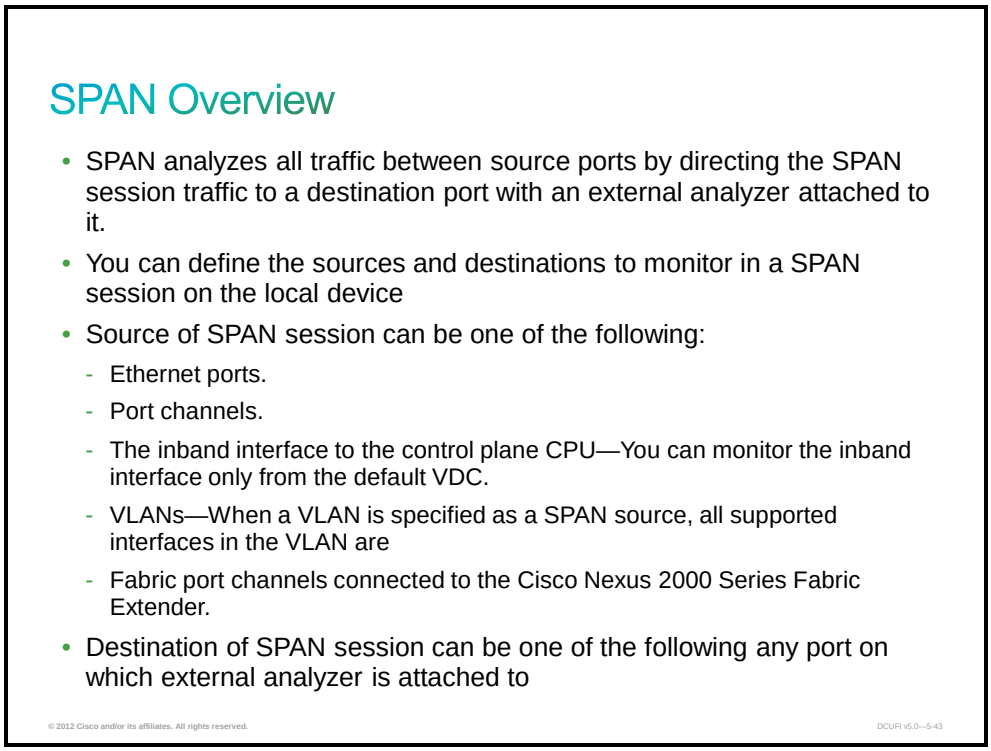
The figure describes the configuration steps you need to perform the configuration on the Cisco Nexus switch.

The **show scheduler name***name* command validates the configuration.

© 2012 Cisco Systems, Inc.
https://t.me/learningnets

Configuring SPAN and ERSPAN

This topic explains how to use and configure the Switched Port Analyzer (SPAN) feature on the Cisco Nexus 7000 and 5000 Series Switches.



SPAN analyzes all traffic between source ports by directing the SPAN session traffic to a destination port with an external analyzer attached to it.

You can define the sources and destinations that you want to monitor in a SPAN session on the local device.

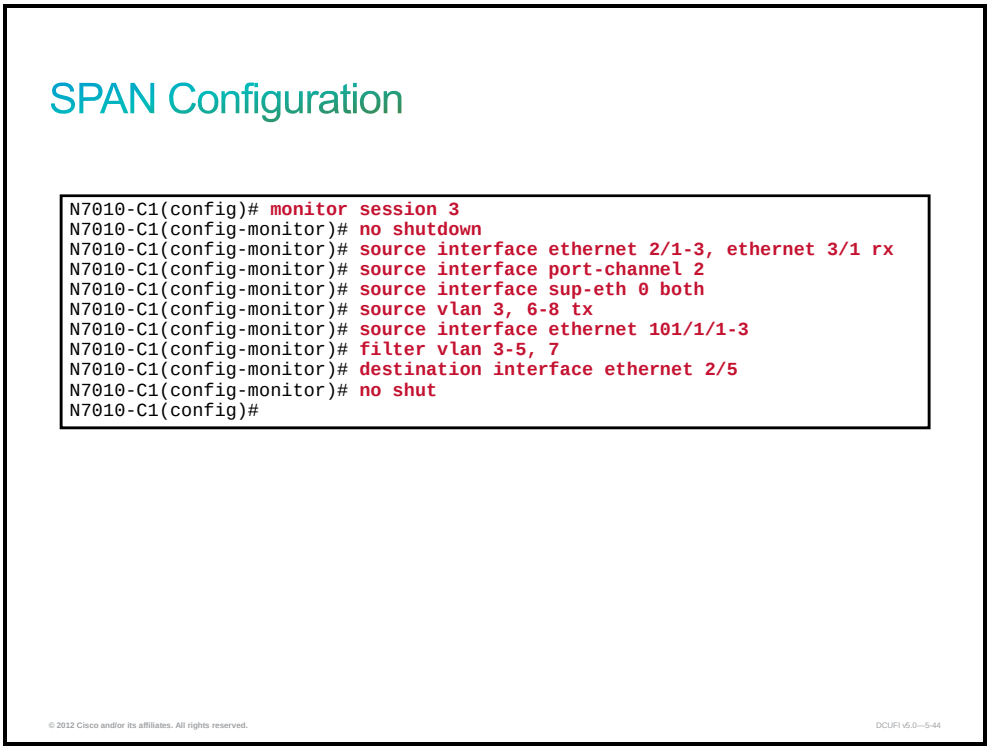
The interfaces from which traffic can be monitored are called SPAN sources. Sources designate the traffic to monitor and whether to copy ingress, egress, or both directions of traffic. SPAN sources include the following:

- Ethernet ports.
- Port channels.
- The in-band interface to the control plane CPU: You can monitor the in-band interface only from the default virtual device context (VDC). In-band traffic from all VDCs is monitored.
- VLANs: When a VLAN is specified as a SPAN source, all supported interfaces in the VLAN are SPAN sources.
- Remote SPAN (RSPAN) VLANs.
- Fabric port channels connected to the Cisco Nexus 2000 Series Fabric Extender.

Note	Layer 3 subinterfaces are not supported.

SPAN destinations refer to the interfaces that monitor source ports. Destination ports receive the copied traffic from SPAN sources.

https://t.me/learningnets



The figure shows the procedure you should follow to configure SPAN on the local switch. For the source, you can choose different types of interfaces and types of traffic.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

ERSPAN Overview

- ERSPAN enables remote monitoring of multiple switches across your network.
- ERSPAN transports mirrored traffic from source ports of different switches to the destination port, where the network analyzer has connected.
- The traffic is encapsulated at the source switch and is transferred to the destination switch, where the packet is decapsulated and then sent to the destination port.
- ERSPAN consists of :
  - ERSPAN source session
  - routable ERSPAN generic routing encapsulation (GRE)-encapsulated traffic,
  - ERSPAN destination session.
- You can configure ERSPAN source sessions and destination sessions on different switches separately.
- ERSPAN does not monitor any packets that are generated by the supervisor, regardless of their source

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

Encapsulated Remote Switched Port Analyzer (ERSPAN) transports mirrored traffic over an IP network. The traffic is encapsulated at the source router and is transferred across the network. The packet is de-encapsulated at the destination router and then sent to the destination interface.

ERSPAN consists of an ERSPAN source session, routable ERSPAN Generic Routing Encapsulation (GRE)-encapsulated traffic, and an ERSPAN destination session. You separately configure ERSPAN source sessions and destination sessions on different switches.

The ERSPAN feature supports stateless and stateful restarts. After a reboot or supervisor switchover, the running configuration is applied.

ERSPAN does not monitor any packets that are generated by the supervisor, regardless of their source.

https://t.me/learningnets

ERSPAN Sources and Destinations

- ERSPAN sources could be one of the following:
  - Ethernet ports and port channels.
  - The inband interface to the control plane CPU—You can monitor the inband interface only from the default VDC. Inband traffic from all VDCs is monitored.
  - VLANs—When a VLAN is specified as an ERSPAN source, all supported interfaces in the VLAN are ERSPAN sources.
  - Fabric port channels connected to the Cisco Nexus 2000 Series Fabric Extender (FEX).
  - Satellite ports and host interface port channels on the Cisco Nexus 2000 Series Fabric Extender— These interfaces are supported in Layer 2 access mode, Layer 2 trunk mode, and Layer 3 mode.
- ERSPAN destinations could be one of the following:
  - Destinations for an ERSPAN session include Ethernet ports or port-channel interfaces in either access or trunk mode.
  - A port configured as a destination port cannot also be configured as a source port.
  - A destination port can be configured in only one ERSPAN session at a time.
  - Destination ports do not participate in any spanning tree instance or any Layer 3 protocols
  - Ingress and ingress learning options are not supported on monitor destination ports.
  - F1 and F2 Series module core ports, Fabric Extender host interface (HIF) ports, HIF port channels

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

The figure shows what interfaces can be a source or destination for the ERSPAN feature.

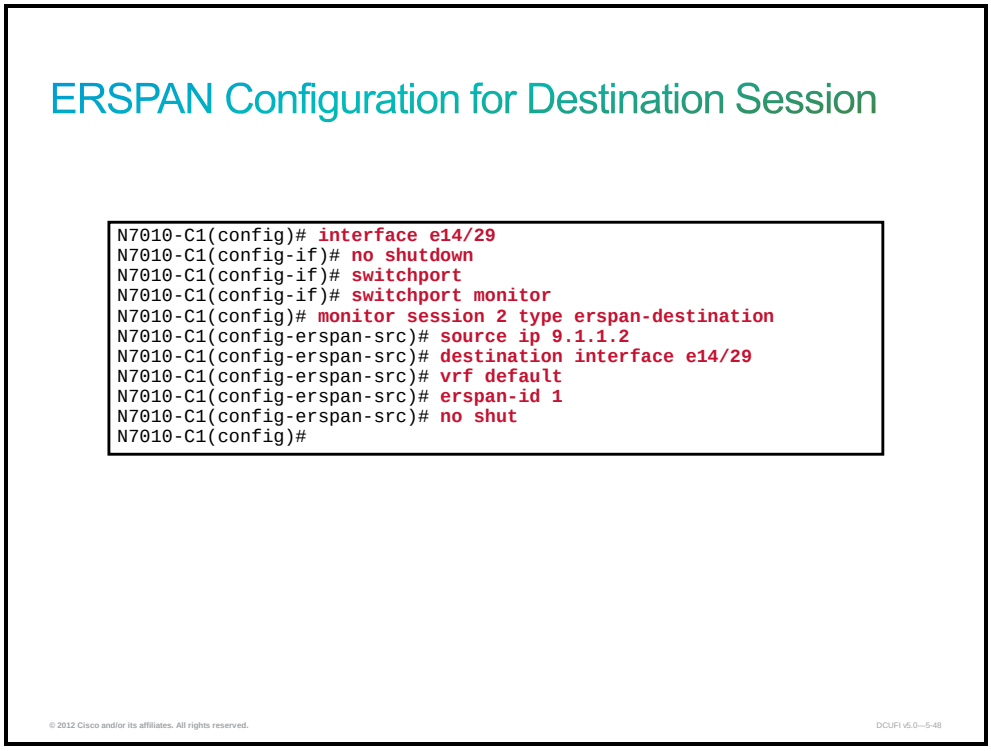
ERSPAN Configuration for Source Session

```
N7010-C1(config)# interface e14/30
N7010-C1(config-if)# no shutdown
N7010-C1(config)# monitor erspan origin ip-address 3.3.3.3 global
N7010-C1(config)# monitor session 1 type erspan-source
N7010-C1(config-erspan-src)# source interface e14/30
N7010-C1(config-erspan-src)# erspan-id 1
N7010-C1(config-erspan-src)# ip ttl 16
N7010-C1(config-erspan-src)# ip dscp 5
N7010-C1(config-erspan-src)# vrf default
N7010-C1(config-erspan-src)# destination ip 9.1.1.2
N7010-C1(config-erspan-src)# no shut
N7010-C1(config)#
```

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

The figure shows the procedure to follow to configure ERSPAN as a source session.

https://t.me/learningnets

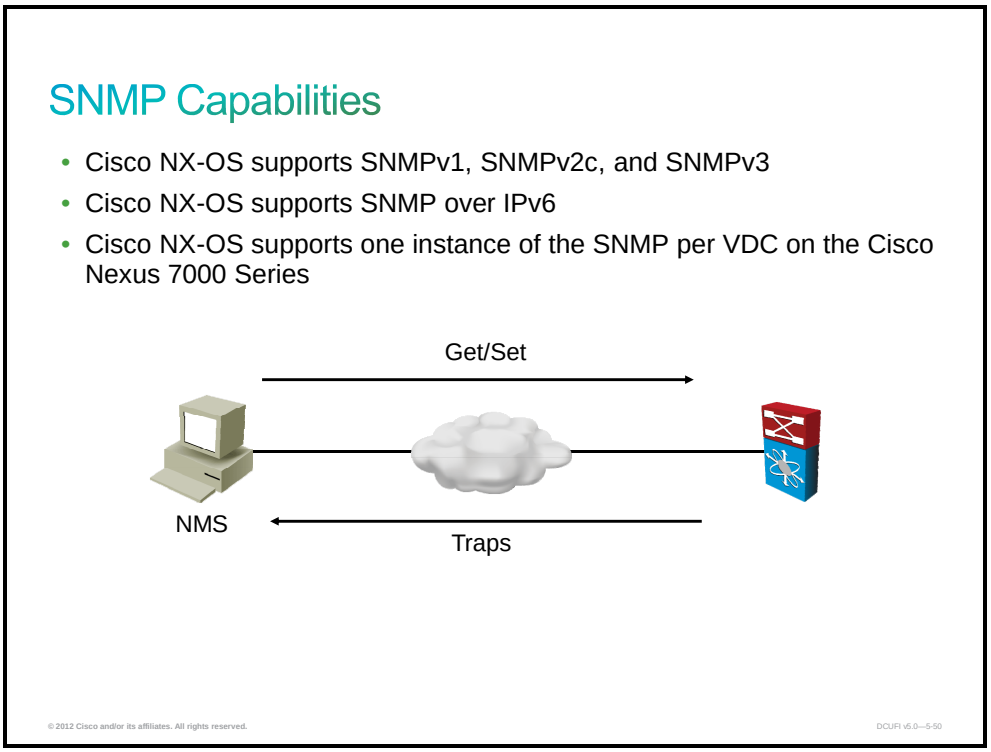


The figure shows the procedure you should follow to configure ERSPAN as a destination session.

https://t.me/learningnets

Configuring SNMP

SNMP is an application layer that provides a message format for communication between SNMP managers and agents. This topic discusses configuring the SNMP feature on a Cisco Nexus switch.



SNMP provides a standard framework and a common language that is used for monitoring and management of devices in the network.

The SNMP framework consists of three parts:

- **SNMP manager:** Used to control and monitor the activities of network devices using SNMP.
- **SNMP agent:** A software component within the managed device that maintains the data for the device and reports the data to managing systems. The Cisco NX-OS Software supports the agent and MIB, enabling the SNMP agent relationship between the manager and the agent, both of which must be defined.
- **MIB:** Collection of managed objects on the SNMP agent.

SNMP notifications are used to indicate improper user authentication, restarts, the closing of the connection, loss of the connection to a neighbor router, or other significant event. These notifications are generated by the SNMP agent and sent to the SNMP manager. The Cisco NX-OS Software generates SNMP notifications as either traps or informs. Traps are less reliable than inform messages because the SNMP manager does not send any acknowledgments of receipt of the trap. SNMP managers that receive an inform request acknowledge receipt; if the Cisco NX-OS Software does not receive that receipt, it sends the inform message again.

© 2012 Cisco Systems, Inc.

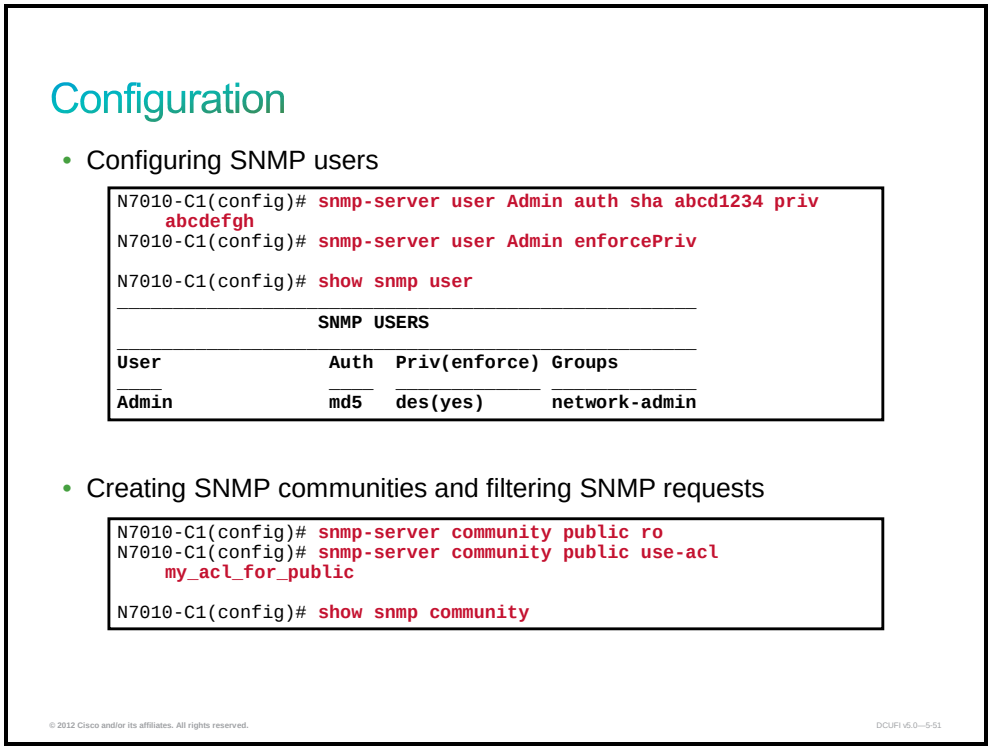
https://t.me/learningnets

SNMPv3 provides secure access to devices by authenticating and encrypting frames over the network. The security features provided include the following:

- **Message integrity:** Ensures that a message has not been tampered with in-transit
- **Authentication:** Ensures that the message is from a valid source
- **Encryption:** Prevents the message from being seen by unauthorized sources

SNMP version 3 (SNMPv3) provides for security models and security levels. A security model is an authentication strategy that is set up for users, while the security level is the permitted level of security inside a security model. The combination of the security model and security level determines which mechanism is employed when processing SNMP packets.

The Cisco NX-OS Software supports one instance of SNMP per VDC on the Cisco Nexus 7000 Series Switches. SNMP is VRF-aware, and you can configure SNMP to use a particular VRF to reach the SNMP notification host receiver. The SNMP notification filter can configure the SNMP host receiver based on the VRF where the notification occurred.

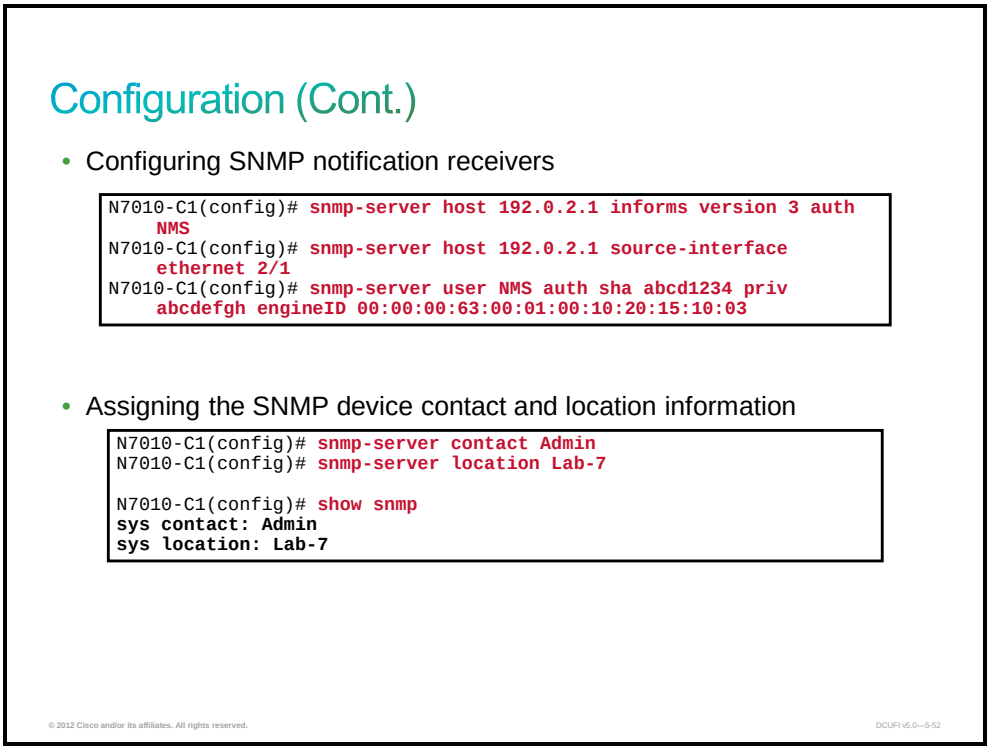


The figure describes how to configure a user for SNMPv3 and how to create SNMP communities for SNMP version 1 (SNMPv1) or community-based SNMP version 2 (SNMPv2).

The **show snmp user** command validates SNMP users.

The **show snmp community** command displays configured SNMP communities.

https://t.me/learningnets



The first part of the figure describes the proper steps that are required to configure SNMP notification receivers. The second part of the figure describes how to set SNMP device contact and location information.

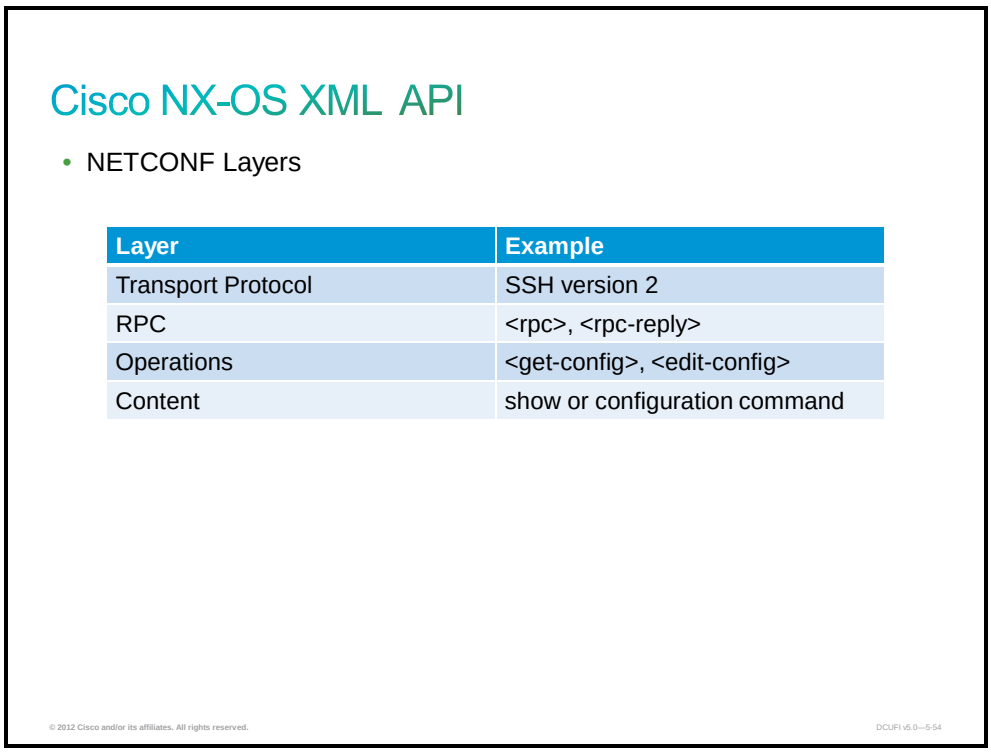
The **show snmp** command displays configured SNMP device contact and location information.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Using XML Interface

The XML interface on the Cisco Nexus Family of switches is an alternative interface that can be used to configure the device. The interface uses the XML-based Network Configuration Protocol (NETCONF) protocol that allows you to manage devices and communicate over the interface using an XML management tool. This topic discusses the use of the XML interface to manage the Cisco Nexus switches.



NETCONF allows you to manage devices and communicate over the interface with an XML management tool or program.

NETCONF is implemented with an XML schema that permits the enclosing of device configuration elements within a remote procedure call (RPC) message. Within the RPC message, choose one of the NETCONF operations that match the type of command you want the device to execute. The entire set of CLI commands can be configured using NETCONF.

There are four NETCONF layers:

- **Secure Shell (SSH) transport protocol:** Provides a secure, encrypted connection between the client and server
- **RPC tag:** Introduces a configuration command from the requestor and the corresponding reply from the XML server
- **NETCONF operation tag:** Indicates the type of configuration command
- **Content:** Indicates the XML representation of the feature to be configured

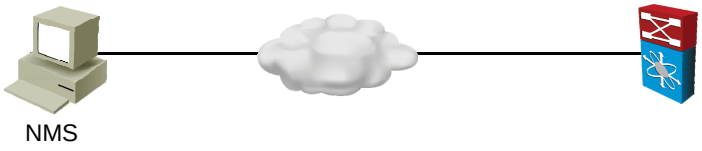
The Cisco NX-OS Software provides an SSH service that is called “xml agent” that uses NETCONF over SSH version 2. NETCONF over SSH is initiated by the exchange of a hello message between the client and the XML server. After the initial exchange, the client sends XML requests to which the server responds with XML responses. The client and server terminate requests and responses with the character sequence]]>]]>. Because this character sequence is not valid in XML, the client and server interpret them when the messages end to ensure synchronization of the communication.

https://t.me/learningnets

Cisco NX-OS XML API (Cont.)

Reasons Why

- SNMP shortcomings
  - %SNMP-3-CPUHOG: Processing GetNext of dot1dTpFdbEntry.1.8.0.32.178.224.223
  - %SYS-3-CPUHOG: Task ran for 3748 msec (13539/12), process = IP SNMP, PC = 801D1C9C.
  - Traceback= 801D1CA4 80076610
  - Google cpuhog snmp returns 10K hits



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

The XML-based NETCONF is the future direction for Cisco regarding network management stations (NMSs). It provides a solution for SNMP shortcomings.

- CLI scraping has been necessary until now because many statistics and counters were not available in MIBs.
- SNMP does not always scale well. The UDP-based “get next request-response,” even with bulk gets, can cause an excessive CPU load. It can be hard for the client to collect statistics simultaneously.
- SNMP is difficult to read and understand:
  - %SNMP-3-CPUHOG: Processing GetNext of dot1dTpFdbEntry.1.8.0.32.178.224.223
  - %SYS-3-CPUHOG: Task ran for 3748 msec (13539/12), process = IP SNMP, PC = 801D1C9C
  - -Traceback= 801D1CA4 80076610
  - Google of **cpuhogsnmp** returns 10,000 responses
- Reading MIBs is not always an easy task.
- Many MIBs suffer from the 32-bit counter rollover.
  - At 10 Gigabit Ethernet, it takes less than 4 seconds to roll over.
  - Binary protocols (even ASN.1 notation ones) are difficult to extend.
- There are no guarantees that interfaces will always keep their index (ifindex persistence).
  - ifTable.ifEntry.ifIndex = 100
    ifTable.ifEntry.ifDesc = “Ethernet1/1”

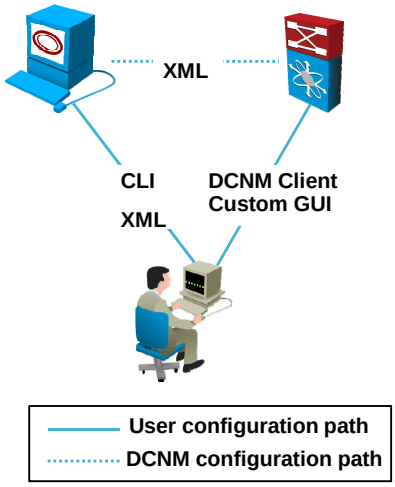
.

https://t.me/learningnets

Configuration Methods

Configuration method

- CLI
- XML API management interface
- Cisco DCNM client
- User-defined GUI



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

You can configure Cisco NX-OS devices using the CLI or the XML management interface over Secure Shell (SSH) as follows:

- **CLI from an SSH session, a Telnet session, or the console port:** You can configure devices using the CLI from an SSH session, a Telnet session, or the console port. SSH provides a secure connection to the device.
- **XML management interface over SSH:** You can configure devices using the XML management interface, which is a programmatic method that is based on the NETCONF protocol that complements the CLI functionality.

You can configure Cisco NX-OS devices using the Cisco DCNM client or from your own GUI as follows:

- **Cisco DCNM client:** You can configure devices using the Cisco DCNM client, which runs on your local PC and uses Web Services on the Cisco DCNM server. The Cisco DCNM server configures the device over the XML management interface.
- **Custom GUI:** You can create your own GUI to configure devices using the Cisco DCNM Web Services application program interface (API) on the Cisco DCNM server. You use the Simple Object Access Protocol (SOAP) to exchange XML-based configuration messages with the Cisco DCNM server. The Cisco DCNM server configures the device over the XML management interface.

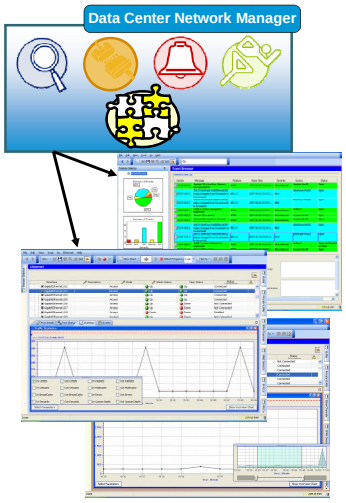
https://t.me/learningnets

Implementing Cisco Prime DCNM for LAN

This topic discusses how to implement Cisco Prime DCNM for LAN to enable monitoring and managing the Cisco Nexus Family of switches.

Cisco DCNM for LAN

- Operational monitoring of data center infrastructure
  - Proactive monitoring
  - Performance and capacity
  - Topology views
- Data center resource management
  - Automated discovery
  - Configuration and change management
  - Template-based provisioning
- Image management
  - Integration with enterprise systems
  - Web services APIs
  - Event forwarding



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

Cisco Prime DCNM is a management solution that maximizes overall data center infrastructure uptime and reliability, which improves business continuity. Focused on the management requirements of the data center network, Cisco DCNM provides a robust framework and rich feature set that fulfills the switching needs of present and future data centers. In particular, Cisco DCNM automates the provisioning process.

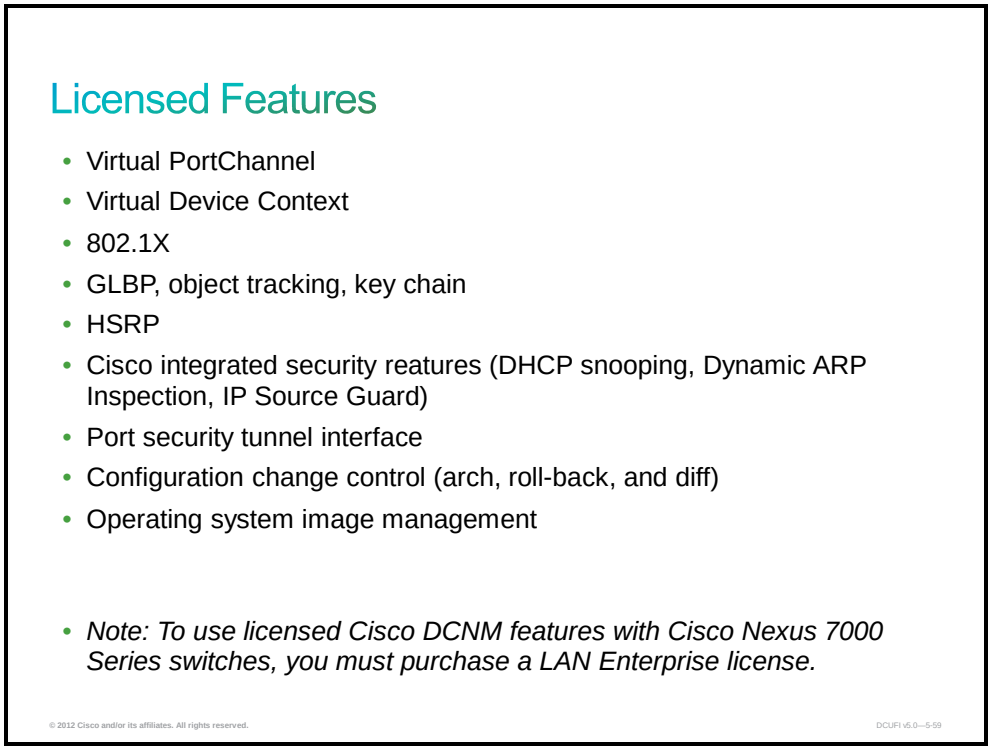
Web Services APIs provide easy integration with third-party applications and allow accurate flow-through provisioning and data mining.

Event forwarding integrates Cisco DCNM with the enterprise network operations center (NOC) for alerts and events. Cisco DCNM uses email and alerts to notify operations staff of critical outages that may be service-impacting.

Cisco DCNM supports the entire Cisco Nexus Family of hardware platforms:

- Cisco Nexus 7000 Series Switches
- Cisco Nexus 5000 Series Switches
- Cisco Nexus 4000 Series Switches
- Cisco Nexus 2000 Series Fabric Extenders
- Cisco Nexus 1000V Series Switches and Cisco Nexus 1010 Virtual Service Appliance

https://t.me/learningnets



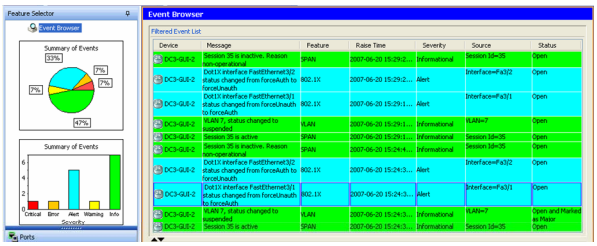
The figure lists the licensed Cisco DCNM features that are used with the Cisco Nexus 7000 Series.

The following features work without a license:

- Port/port channel
- Fabric extender port pinning
- Link Aggregation Control Protocol (LACP)
- VLAN 802.1Q and private VLAN (PVLAN)
- Spanning Tree Protocol (STP) (Multiple Spanning Tree [MST] and Rapid Per VLAN Spanning Tree[Rapid PVST])
- ACL (Mac, IP, VLAN)
- Traffic storm control
- Authentication, authorization, and accounting (AAA)
- Environmental (hardware resource utilization with ternary content addressable memory [TCAM] statistics)
- Module temperature
- SPAN
- Discovery and inventory
- Layer 1 and Layer 2 topology map
- Fault management
- Traffic statistics reports
- Role-based access control (RBAC)
- Web services
- Java API

https://t.me/learningnets

Fault Management



- Industry standard event browser
- Event collection and normalization
- Per network feature correlation
- Noise filtering for root cause isolation
- Event propagation

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

The Cisco DCNM client includes the Event Browser and feature-specific Events tabs that appears in the Details pane for features that can have events. The Event Browser shows all recent status events while a feature-specific Events tab shows recent status events that pertain to the feature.

The Cisco DCNM client updates the Event Browser and Events tabs dynamically when it receives new events from the server. You can use the Event Browser to view recent events and a summary chart of those events.

You can filter events in the Event Browser by the following criteria:

- **Event date and time:** By default, the Cisco DCNM client displays all events that are received after you started the Cisco DCNM client and for a configurable number of hours before starting the Cisco DCNM client
- **Event severity:** By default, the Cisco DCNM client displays events of all severities.

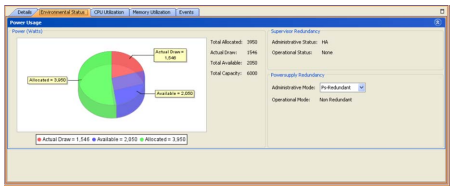
You can change the status of an event to one of the following statuses:

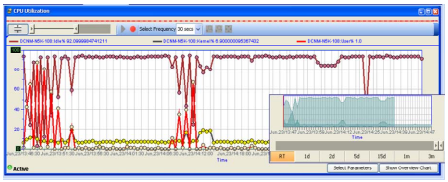
- **Acknowledged:** Shown as a green check mark.
- **Closed:** Shown as a yellow folder.

https://t.me/learningnets

Performance Monitoring

- Real-time traffic statistics
  - Port utilization
  - Port error
  - Packet loss
  - GRE tunnel
  - Port security
  - PortChannel
  - VLAN traffic
  - ARP packets
  - Storm control statistic
- Environmental status and resource utilization
- Historical reports
- Visuals and exportable reports

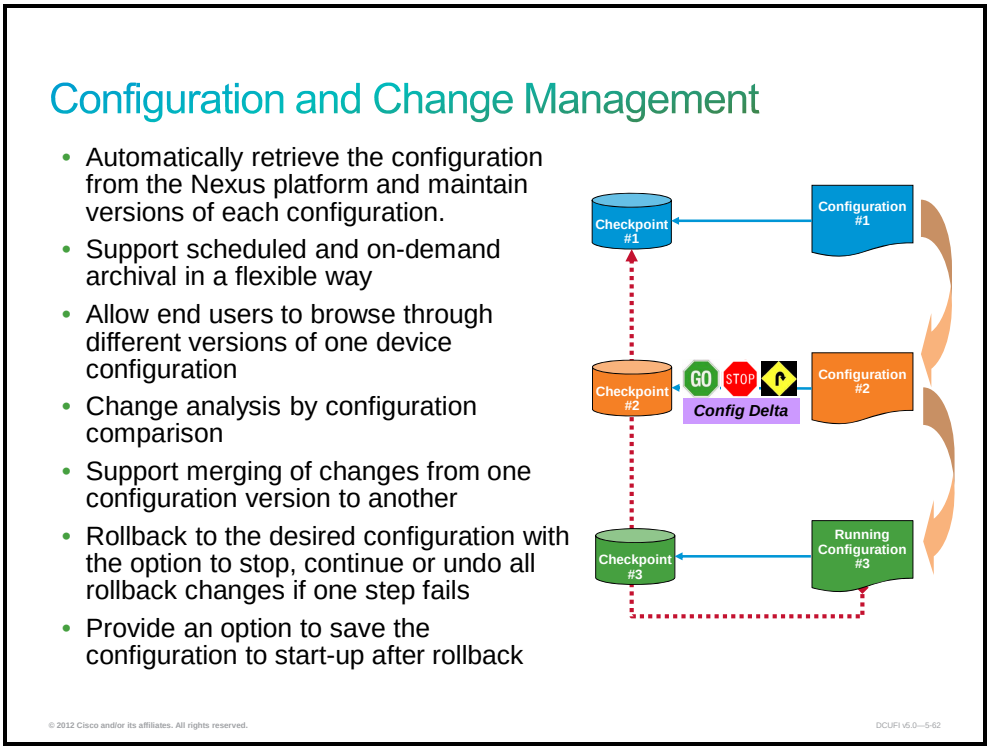




© 2012 Cisco and/or its affiliates. All rights reserved.

Cisco DCNM provides various statistics such as real-time interface statistics, power usage, environmental status, CPU utilization, memory utilization, and chassis events.

https://t.me/learningnets



The Configuration Change Management feature allows you to keep an archive of configurations from managed devices. You can view and compare archived configurations. You can roll back the running configuration of a managed device to any archived configuration version that is available for the device in Cisco DCNM.

Version Browser

The Version Browser feature allows you to see information about archived configurations, view and compare specific configuration versions, and merge changes from one configuration version to another. From the Version Browser, you can initiate a configuration rollback for a managed Cisco Nexus Series device, using any of the archived configurations available in Cisco DCNM for the device. Cisco DCNM uses the rollback feature available in Cisco NX-OS Software.

Archival Jobs

The Archival Jobs feature allows you to control the automated archival of the running configuration on managed devices.

Archival Settings

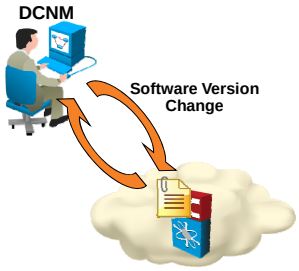
The Archival Settings feature allows you to configure settings that are related to configuration change management, including the number of configuration versions that Cisco DCNM stores for each managed device, how many rollback and archival history entries Cisco DCNM stores for each managed device, and which file server Cisco DCNM uses during a configuration rollback.

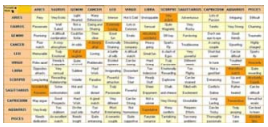
© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Cisco NX-OS Image Management

- Wizard-based installation of Cisco NX-OS images on multiple devices simultaneously
- Performs validations before installation:
  - Verifies the switch's flash memory space availability for the image to download
  - Verifies compatibility between currently running network services and the new image
- Allows for time based deployment i.e. on-demand or scheduled
- Fully leverages Cisco NX-OS ISSU transparent software upgrade that has no impact to the network traffic (no service disruption, zero packet loss)
- Detects installation failure and automatically initiates recovery action
- Images can be installed from external servers using TFTP/FTP/SFTP





Compatibility Matrix

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

The Device OS Management screen allows you to view information about the software images that are used by a managed device. You can also start the Software Installation Wizard from the Device OS Management Summary pane.

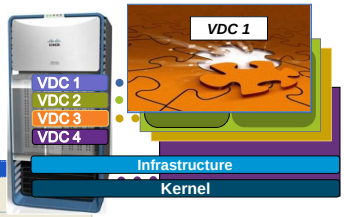
The Cisco DCNM supports file servers that use the following protocols:

- FTP
- Secure FTP (SFTP)
- TFTP

https://t.me/learningnets

VDC Management

- VDC are transparently handled throughout the application wizard-based configuration
  - Interface allocation across VDC
  - Resource limit enforcement with templates
  - Resource consumption monitoring
  - IPv4- and IPv6-capable
- VDC-aware fault and performance monitoring
- VDC-aware RBAC
- Topology representation
- VDC per chassis
- VDC-to-VDC connectivity
- Real-time or delayed discovery



Resources

Refresh

Name	Allocation		Current Usage		
	Minimum	Maximum	Used	Available	Used Percent
Port Channels	0	768	2	758	0%
IPV4 Route Memory Limit	8	8	1	7	12%
VRF	2	1000	2	993	0%
M4 Route Memory Limit	8	8	2	6	25%
M6 Route Memory Limit	2	2	1	1	50%
IPV6 Route Memory Limit	4	4	1	3	25%
SPAN Sessions	0	2	0	1	0%
VLAN	16	4094	21	4073	0%

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

The figure describes managing VDCs with Cisco DCNM.

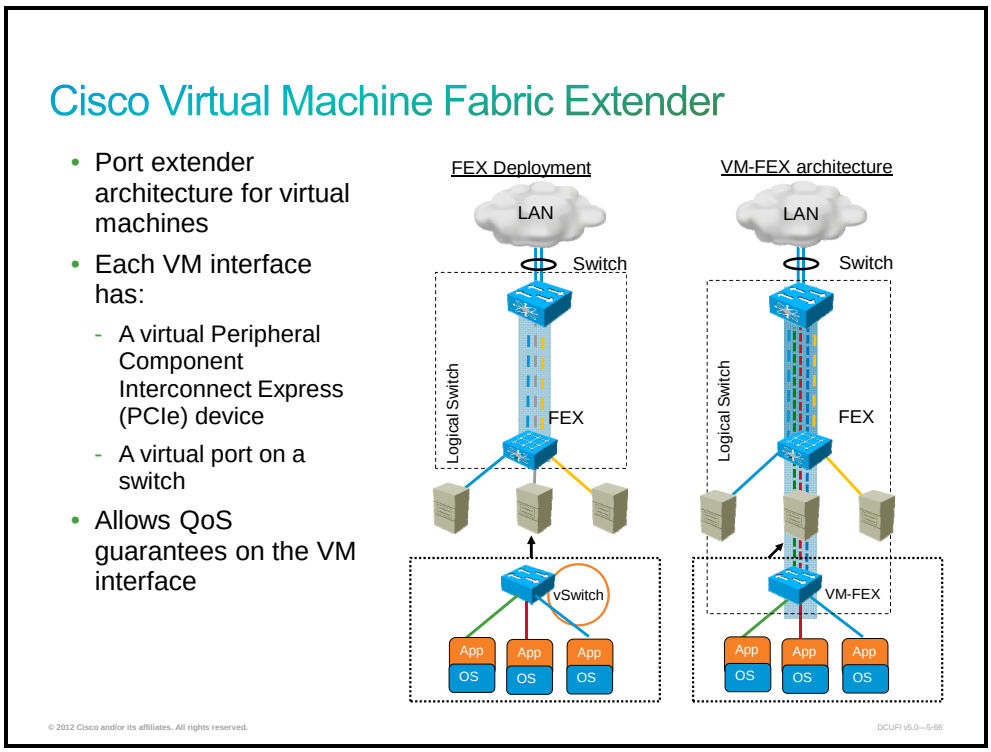
VDC Setup Wizard is used to create a new VDC.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

Integrating VMware vCenter Manager

This topic identifies how to integrate Cisco Nexus with the VMware vCenter Manager.



Virtualization allows the creation of multiple virtual machines (VMs) to run in isolation, side by side, on the same physical machine.

Each VM has its own set of virtual hardware (RAM, CPU, network interface card [NIC]) upon which an operating system and fully configured applications are loaded. The operating system sees a consistent, normalized set of hardware regardless of the actual physical hardware components.

In a VM, both hardware and software are encapsulated in a single file for rapid copying, provisioning, and moving between physical servers. You can move a virtual machine, within seconds, from one physical server to another for zero-downtime maintenance and continuous workload consolidation.

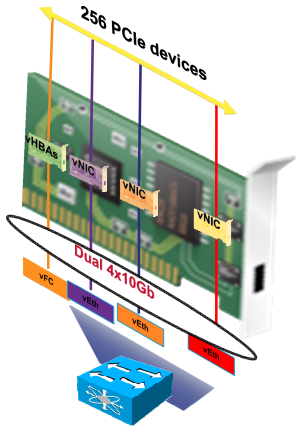
Cisco Data Center Virtual Machine Fabric Extender (Cisco VM-FEX) extends the (prestandard) IEEE 802.1Qbh port extender architecture to virtual machines. In this architecture, each VM interface is provided with a virtual Peripheral Component Interconnect Express (PCIe) device and a virtual port on a switch. This solution allows precise rate limiting and quality of service (QoS) guarantees on the VM interface.

Cisco VM-FEX works with a Cisco virtual interface card (VIC) adapter to replace software-based switching on the server with external hardware-based switching in the fabric interconnect.

https://t.me/learningnets

Cisco VIC Adapter

- CNA designed for both single-OS and VM-based deployments
- Replaces software-based switching on the server with hardware switching in the fabric interconnect
- Supports static or dynamic virtualized interfaces
- Offers up to 128 vNICs
- Example: Cisco UCS M81KR VIC



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

A Cisco VIC adapter, such as the Cisco UCS M81KR Virtual Interface Card, is a converged network adapter (CNA) that is designed for both single operating system and VM-based deployments. The VIC adapter supports static or dynamic virtualized interfaces, including up to 128 virtual network interface cards (vNICs).

VIC adapters support VM-FEX to provide hardware-based switching of traffic to and from VM interfaces.

In a VMware environment, VM-FEX supports the standard VMware integration with VMware ESX hypervisors installed on the server and all virtual machine management that is performed through the VMware vCenter.

NIC and converged network adapters support virtualized environments with the standard VMware integration with VMware ESX installed on the server and all VM management performed through the virtual center.

If you implement service profiles, you retain the ability to easily move a server identity from one server to another. After you image the new server, the VMware ESX treats that server as if it were the original.

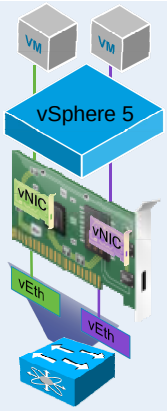
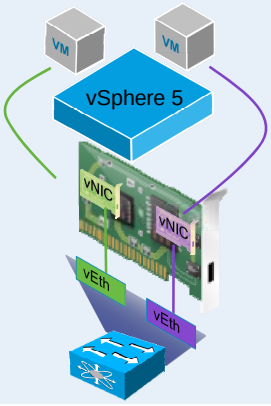
These adapters implement the standard communications between VMs on the same server. If a VMware ESX host includes multiple VMs, all communications must go through the virtual switch on the server.

If the system uses the native VMware drivers, the virtual switch is out of the domain of the network administrator and is not subject to any network policies. As a result, for example, QoS policies on the network are not applied to any data packets that are traveling from VM1 to VM2 through the virtual switch.

If the system includes another virtual switch, such as the Cisco Nexus 1000V Series Switches, that virtual switch is subject to the network policies that are configured on that switch by the network administrator.

https://t.me/learningnets

Cisco VM-FEX Modes

Mode	Standard (Emulated) Mode	High Performance Mode (VMDirect Path)
Description	• 12%-15% CPU performance improvement • vMotion supported	• Co-exists with Standard mode • Bypasses Hypervisor layer • 30% improvement in I/O performance • vMotion supported with ESX 5.0
Operational model		

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

Cisco VM-FEX ports can operate in standard mode or high-performance mode.

In standard mode, traffic to and from a VM passes through the distributed virtual switch (DVS) and the hypervisor. The figure in the left table column shows the traffic paths taken by VM traffic on a Cisco Unified Computing System (UCS) server with a VIC adapter that has Cisco VM-FEX ports in standard mode.

In high-performance mode, traffic to and from a VM bypasses the DVS and hypervisor. Traffic travels directly between VMs and the VIC adapter.

The benefits of high-performance mode are as follows:

- Increases I/O performance and throughput
- Decreases I/O latency
- Improves CPU utilization for virtualized I/O-intensive applications

With VMware, high-performance mode also supports VMware vMotion. During the vMotion migration, the hypervisor reconfigures links in high-performance mode to be in standard mode, transitions the link to the new hypervisor, and then reconfigures the link to be in high-performance mode. The figure in the right table column shows how Cisco VM-FEX operates in high-performance mode with vMotion.

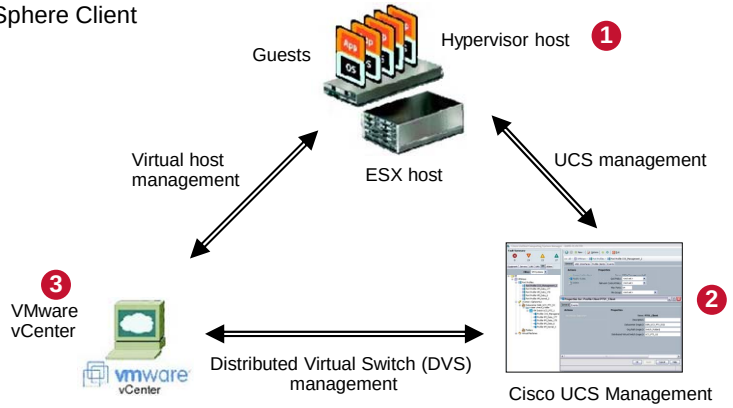
Specifically, the high-performance mode operation involves these stages:

1. Two VMs are attached to a VIC in high-performance mode.
2. When the vMotion migration begins on one VM, that VM transitions to standard mode.
3. The VM migrates to the other host, and standard mode is established.
4. The VM transitions back to high-performance mode.

https://t.me/learningnets

Cisco VM-FEX Components

1. Hypervisor host
2. Cisco UCS Manager or C-series UCS C and Nexus 5500
3. VMware vCenter virtualization management software
   - vCenter Server
   - vSphere Client



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

At a high level, Cisco VM-FEX for VMware requires a hypervisor host, Cisco UCS Manager, and VMware vCenter virtualization management software. The figure shows these three main components and the methods by which they are connected.

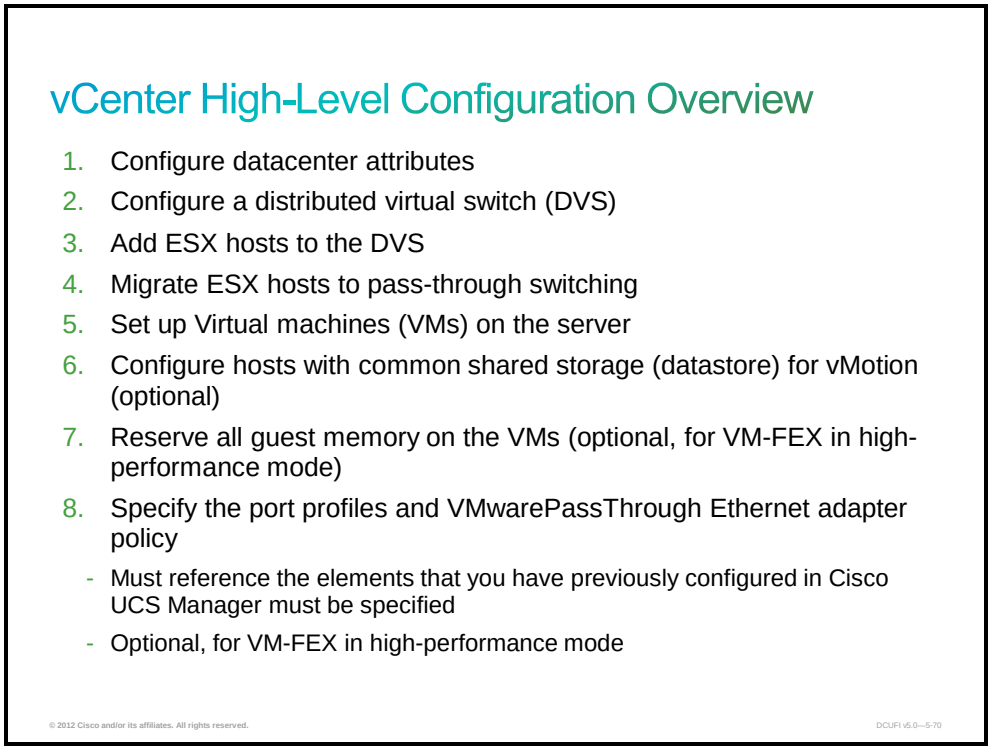
You must install a Cisco UCS M81KR Virtual Interface Card in the server that you intend to use as the hypervisor host.

You need VMware vCenter (vCenter Server and vSphere Client) for Cisco VM-FEX for VMware.

Note	It can also be used with the Cisco Nexus 5500 Platform switches and Cisco UCS C-Series Rack Server with P81E Virtual Interface card.
Note	For additional information visit: http://www.cisco.com/en/US/prod/collateral/switches/ps9441/ps9670/data_sheet_c78-657397.html

© 2012 Cisco Systems, Inc.

https://t.me/learningnets



You must configure VMware vCenter with the following parameters:

- A data center.
- A DVS.
- ESX hosts added to the DVS and configured to migrate to pass-through switching (PTS)/DVS.
- VMs that are required for the VMs on the server.
- (For VMware vMotion) Hosts with common shared storage (datastore) that are properly configured for vMotion.
- (For VM-FEX in high-performance mode) All guest memory on the VMs must be reserved.
- (For VM-FEX in high-performance mode) The port profiles and VMware PassThrough Ethernet adapter policy that you have previously configured in Cisco UCS Manager must be specified.

For information about how to configure these required components in VMware vCenter, see the VMware product documentation.

https://t.me/learningnets

Managing Licensing

Licensing provides access to specific features on the Cisco Nexus Series switch after you install the appropriate license for the feature. This topic discusses the licensing management for the Cisco Nexus Family of switches.

Licensing: Cisco Nexus 7000

Package	Content
Enterprise Services Package	• OSPF • BGP • EIGRP • IS-IS • PIM and SSM • MSDP • Policy-Based Routing • GRE
Advanced Service Package	• VDCs • Cisco TrustSec solution
Transport Services Package	• OTV • LISP
Scalable Feature Package	Provides chassis license for XL feature modules (one per chassis) – different license for each chassis model
Enhanced Layer 2 Package	FabricPath support on the F Series module
MPLS Services Package	Multiprotocol Label Switching (MPLS)
Storage Enterprise Package	• Inter-VSAN routing over FC and FCoE • IVR NAT over FC • VSAN-based Access Control • Fabric Bindings for open systems

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

The licensing model for the Cisco NX-OS Software is feature-based. Feature-based licenses make features available to the entire physical device. Each license only supports the listed features.

Note Any feature that is not included in a license package is bundled with the Cisco NX-OS Software and is provided at no extra charge.

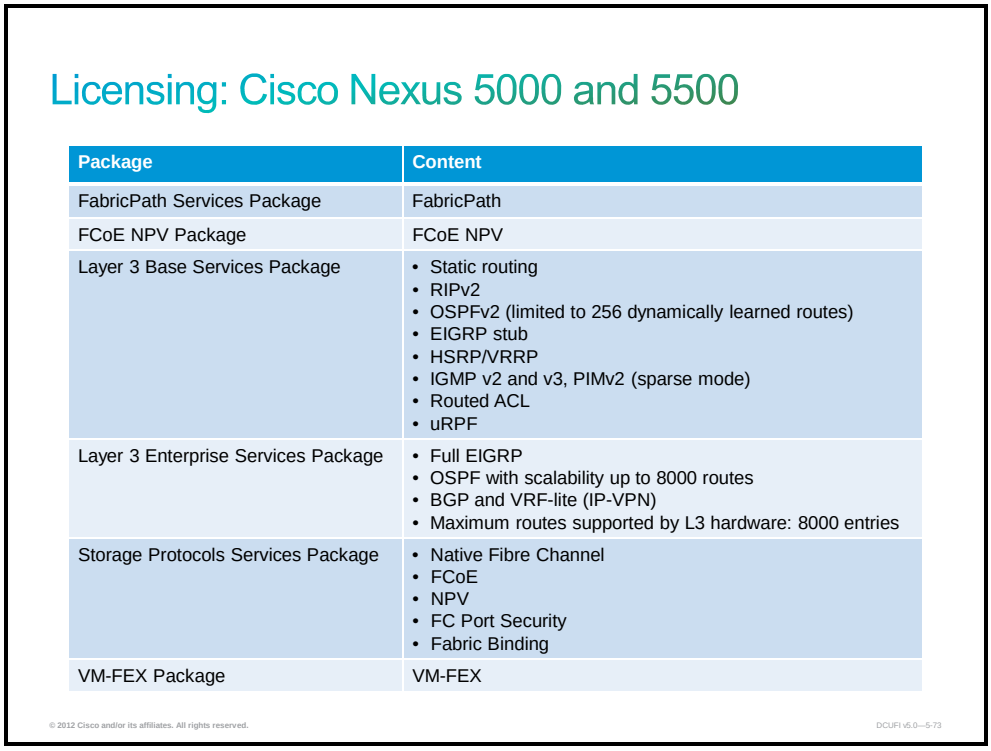
Note Licenses are independent of each other. If you want to use features that are covered by different licenses you must install all the appropriate licenses.

Installing any license in the device is a nondisruptive process, and will automatically save a copy of the permanent license to the chassis.

If you have enabled the grace period feature, enabling a licensed feature that does not have a license key starts a counter on the grace period. You then have 120 days to install the appropriate license keys, disable the use of that feature, or disable the grace period feature. If at the end of the 120-day grace period, the device does not have a valid license key for the feature, the Cisco NX-OS Software automatically disables the feature and removes the configuration from the device.

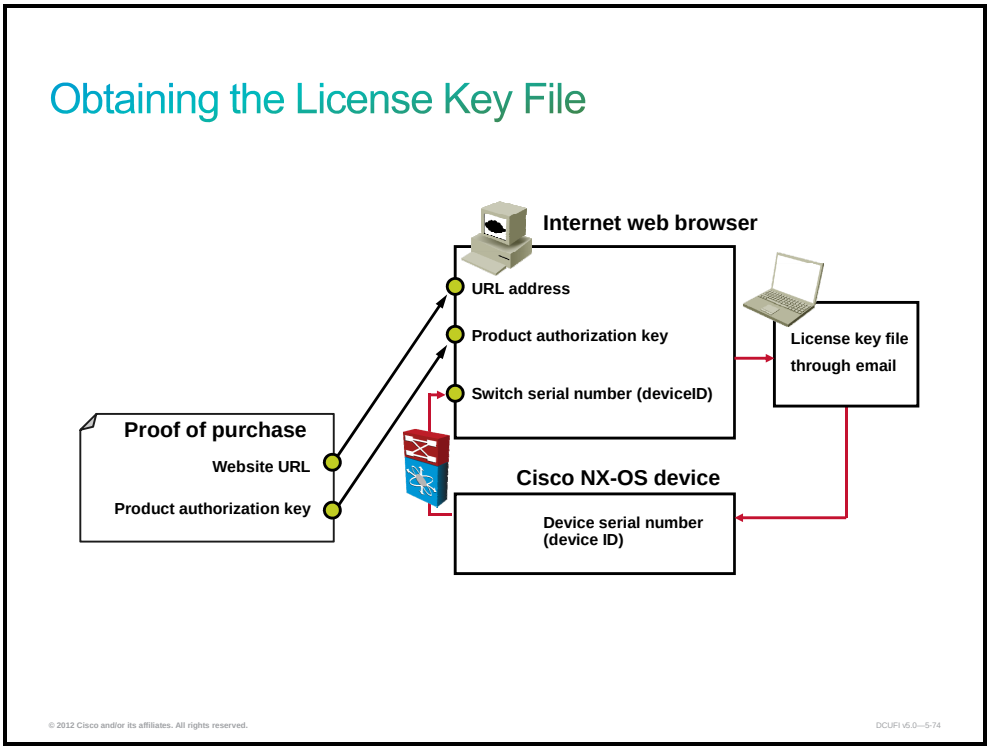
Note Cisco TrustSec solution does not have a grace period.

https://t.me/learningnets



The figure shows the licensing options that are available for the Cisco Nexus 5000 and 5500 Platform switches.

https://t.me/learningnets



Licenses can be obtained as factory-installed licenses for new Cisco NX-OS devices.

If you have an existing device, or if you wish to install the license yourself, you must first obtain the license key file and then install that file in the device.

The figure shows how to obtain a license key file.

To obtain the serial number for a device use the **show license host-id** command.

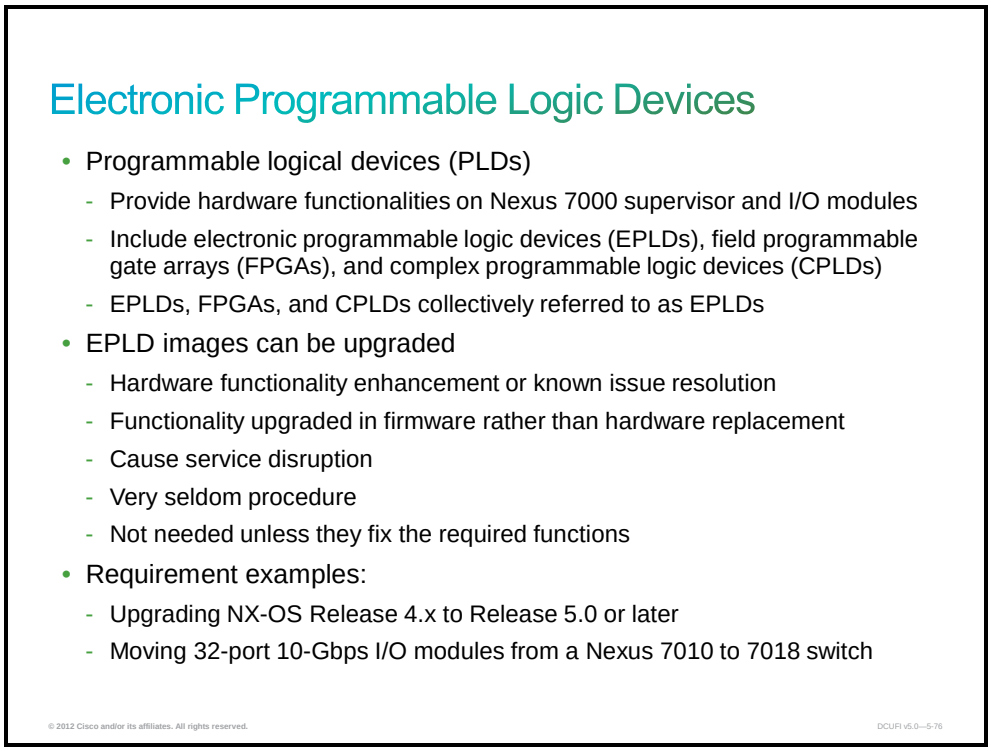
To install the license use the **install license bootflash:<*file_name*>**command.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Upgrading the Firmware and EPLDs

This topic identifies when the firmware and electronic programmable logic device (EPLDs) should be upgraded and how to perform that task.



The Cisco Nexus 7000 Series Switches contain several programmable logical devices (PLDs) that provide hardware functionalities in all modules. Cisco provides EPLD image upgrades to enhance hardware functionality or to resolve known issues. PLDs include EPLDs, field programmable gate arrays (FPGAs), and complex programmable logic devices (CPLDs), but they do not include ASICs. In this document, the term EPLD is used for FPGA and CPLDs.

The advantage of having EPLDs for some module functions is that when you need to upgrade those functions, you just upgrade their software images instead of replacing their hardware.

EPLD image upgrades for an I/O module disrupt the traffic that is going through the module because the module must power down briefly during the upgrade. The system performs EPLD upgrades on one module at a time, so at any one time the upgrade disrupts only the traffic going through one module.

Cisco does not provide upgrade EPLD images very frequently, and you do not have to upgrade your EPLD images unless they fix the functions for the hardware that you are using in your Cisco Nexus 7000 Series switch. The EPLD image upgrades are independent from the Cisco NX-OS In Service Software Upgrade (ISSU) process, which upgrades the system and kickstart images with no impact on the network environment.

https://t.me/learningnets

EPLD Upgrade Procedure

1. Determine whether to upgrade EPLD images
2. Download the EPLD image(s)
   - Go to cisco.com and choose the Products & Services tab
   - Navigate to Switches > Data Center > Nexus 7000 > Support > Download Software > NX-OS EPLD Updates
3. Upgrade the EPLD image(s) using one of the methods:
   - All of the modules installed in your switch
   - Specific modules installed in your switch
4. Verify the EPLD upgrades

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

You do not always need to upgrade EPLD images; however, when new EPLD images are available, the upgrades are always recommended if your network environment allows for a maintenance period in which some level of traffic disruption is acceptable. If such a disruption is not acceptable at this time, then you might consider postponing the upgrade until a better time.

In general, the EPLD upgrade procedure consists of these tasks:

1. Determine whether to upgrade EPLD images.
2. Download the EPLD images. You should go to Cisco.com and choose the Products & Services tab. Then navigate to Switches > Data Center > Nexus 7000 > Support > Download Software > NX-OS EPLD Updates.
3. Upgrade the EPLD images on all of the modules that are installed in your switch, or only on specific modules that are installed in your switch.
4. Verify the EPLD upgrades.

https://t.me/learningnets

Task 1: Determine Whether to Upgrade EPLD Images

Modules to Verify EPLD Status	Command
All modules on the switch	**show install all impact epld bootflash:***filename*
I/O and supervisor modules	**show install module** *slot_number* **impact epld bootflash:***filename*
Fabric modules	**show install xbar-module** *slot_number* **impact epld bootflash:***filename*
Fan-tray modules	**show install fan-module** *slot_number* **impact epld bootflash:***filename*

```
Nexus-7k# show version module 7 epld

EPLD Device                          Version
---------------------------------------
Power Manager                          4.008
IO                                     1.016
Forwarding Engine                      1.006
FE Bridge(1)                         186.008
FE Bridge(2)                         186.008
Linksec Engine(1)                      2.007
<output omitted>
```

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

You can use various show commands to determine whether the EPLDs can be upgraded for all the modules or for specific modules on a switch. These commands indicate the current EPLD images, new EPLD images, and whether the upgrades would be disruptive to switch operations.

The output of the **show version module 7 epld** command shows how to determine the EPLD version for the N7K-M132XP-12 module that is installed in a specific customer scenario. If the line FE Bridge(x) version displays a version earlier than 186.7, you should schedule an EPLD upgrade to a version that is compatible with the target Cisco NX-OS release. For example, if you want to run Cisco NX-OS Release 6.0(1), you should choose Release 6.0(1) EPLDs. In this example, the module runs Release 186.8 on the FE Bridge line, which is a correct EPLD version.

https://t.me/learningnets

Task 3: Upgrade EPLD Images

Modules Upgraded	Command
All installed modules with one module upgraded at a time	**install all epld** *epld_image*
All installed modules with the I/O modules upgraded in parallel	**install all epld** *epld_image* parallel
One or more I/O and supervisor modules with the I/O modules upgraded in parallel	**install all epld** *epld_image* parallel **module {all |** *slot_numbers***}**
One or more I/O and supervisor modules with the I/O modules upgraded in parallel and one or more fan-tray modules	**install all epld** *epld_image* parallel **module {all |** *slot_numbers***} fan-module {all |** *slot_numbers***}**
One or more I/O and supervisor modules with the I/O modules upgraded in parallel and one or more fabric (xbar) modules	**install all epld** *epld_image* parallel **module {all |** *slot_numbers***} xbar-module {all |** *slot_numbers***}**
One or more fan-tray modules and one or more fabric (xbar) modules	**install all epld** *epld_image* parallel **fan-module {all |** *slot_numbers***} xbar-module {all |** *slot_numbers***}**
One I/O or supervisor module	**install module** *slot_number* epld *epld_image*
One fan module	**install fan-module** *slot_number* epld *epld_image*
One fabric module	**install xbar-module** *slot_number* epld *epld_image*

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

To upgrade the EPLD images for a Cisco Nexus 7000 Series switch, you use one of the install commands that are listed in the table. These commands enable you to upgrade the EPLD images for all of the modules on the switch, multiple modules of one or two types, or single modules. When you specify a slot number, use one number. When you specify slot numbers, you can specify **all** for all slots, multiple slots that are separated by commas (x,y,z) or a range of slot numbers (x-y).

When you upgrade both supervisor modules in a switch, Cisco NX-OS upgrades the EPLD images for the standby supervisor module and then upgrades the active supervisor module. This action enables the upgrade of supervisor modules to be nondisruptive to switch operations.

When you upgrade a supervisor module in a single-supervisor switch, the operation is disruptive to switch operations if the switch is active.

https://t.me/learningnets

Task 4: Verify EPLD Upgrades

Command	Modules verified
show version module *slot_number* **epld**	I/O and supervisor modules
show version fan *slot_number* **epld**	Fan-tray modules
show version xbar *slot_number* **epld**	Fabric modules

```
switch# show install epld status
1) Xbar Module 4 upgraded on Wed Oct 26 16:36:27 2011 (524778 us)

Status: EPLD Upgrade was Successful
EPLD                                Curr Ver    Old Ver
-----------------------------------------------------
Power Manager                        1.003         1.003

2) Module 14 upgraded on Mon May 23 19:45:55 2011 (835895 us)

Status: EPLD Upgrade was Successful
Linksec Engine(4)                   2.007
<output omitted>
```

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

You can verify the EPLD upgrades for each slot in the switch by using the commands that are listed in the table.

The example below the table displays how you can use the **show install epld status** command to display the status of EPLD upgrades on the switch.

https://t.me/learningnets

Summary

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

Summary

- System management is the process of actively managing the system and being able to identify issues as soon as they arrive so they can be resolved quickly
- Cisco Fabric Services is a Cisco proprietary feature that distributes configuration data and changes to all Cisco NX-OS devices in a network
- The Network Time Protocol synchronizes the clocks of computer systems over packet-switched and variable-latency data networks
- The Cisco EEM is used to trigger actions based on events that occurr on the switch
- NetFlow is a technology that allows you to gather data about network traffic for the purpose of monitoring
- Cisco Smart Call Home provides an email-based notification for critical system policies
- The scheduler allows the administrator to define and set up a timetable for maintenance activities that should be performed regularly or once-off

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

Summary (Cont.)

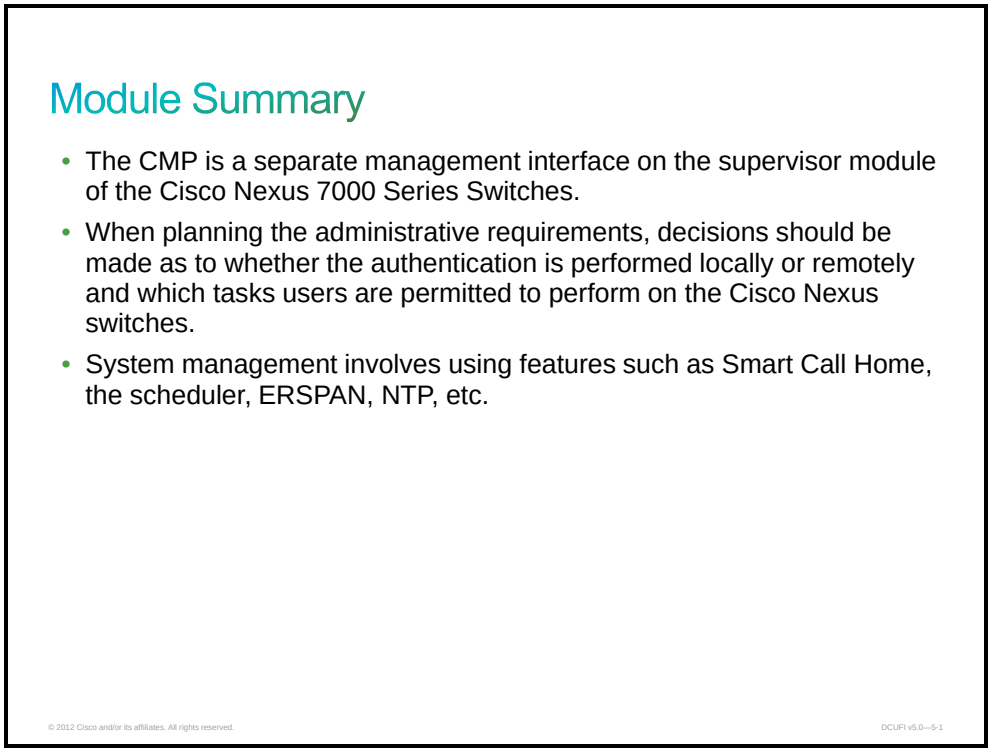
- SPAN is a technology that allows you to copy and monitor all data that travels across a switch port
- SNMP provides a message format for communication between SNMP managers and agents
- The XML interface is an alternative interface that can be used to configure the Cisco NX-OS device
- Cisco DCNM is a management solution focused on the management requirements of the data center network
- The three major VM-FEX components are Hypervisor host, Cisco UCS Manager, and VMware vCenter software consisting of the vCenter Server and vSphere Client
- Certain features on the Cisco Nexus Series are licensed
- The images of the electronic programmable logic devices (EPLDs) may have to be upgraded to enable new features on Cisco Nexus 7000 platforms

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

https://t.me/learningnets

Module Summary

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



The connectivity management processor (CMP) provides an out-of-band (OOB) management and monitoring capability that is independent of the primary operating system on the Cisco Nexus 7000 Series Switches. The CMP enables lights-out Remote Monitoring (RMON) and management of the supervisor module, all other modules, and the Cisco Nexus 7000 Series system without the need for separate terminal servers. In a dual-supervisor system, each CMP has its own IP address that is separate from the management interface on the main operating system.

When planning which user management features to use on the Cisco Nexus Family of switches, you should decide on local or remote authentication, and which secure connection should be used for administrative purposes. Users are configured and associated with roles. The roles contain rules detailing which tasks users are able to perform on the switch.

The Cisco Nexus Family of switches contains a plentiful set of system management features. These include Cisco Fabric Services for synchronization of configuration databases across the fabric, the scheduler for automating regularly performed tasks, and Cisco Smart Call Home as an email-based notification system for monitoring events that have occurred on the switch.

This module discussed the management tools that are available on the Cisco Nexus Family of switches.

https://t.me/learningnets

References

For additional information, refer to these resources:

- Cisco Systems, Inc. *Cisco Nexus 5000 Series NX-OS Security Configuration Guide, Release 5.x.* San Jose, California, July 2012. http://www.cisco.com/en/US/docs/switches/datacenter/nexus5000/sw/security/521_n1_1/b_5k_Security_Config_521N11.html
- Cisco Systems, Inc. *Cisco Nexus 5000 Series NX-OS System Management Configuration Guide, Release 5.x.* San Jose, California, July 2012. http://www.cisco.com/en/US/docs/switches/datacenter/nexus5000/sw/system_management/521_n1_1/b_5k_System_Mgmt_Config_521N11_preface_00.html
- Cisco Systems, Inc. *Cisco Nexus 7000 Series Connectivity Management Processor Configuration Guide* San Jose, California, August 2012. http://www.cisco.com/en/US/docs/switches/datacenter/hw/nexus7000/cmp/configuration/guide/n7cmp_preface.html
- Cisco Systems, Inc. *Cisco Nexus 7000 Series NX-OS Security Configuration Guide, Release 6.x.* San Jose, California, August 2012. http://www.cisco.com/en/US/docs/switches/datacenter/sw/6_x/nx-os/security/configuration/guide/b_Cisco_Nexus_7000_NX-OS_Security_Configuration_Guide__Release_6.x.html
- Cisco Systems, Inc. *Cisco Nexus 7000 Series NX-OS System Management Configuration Guide* San Jose, California, May 2010. http://www.cisco.com/en/US/docs/switches/datacenter/sw/6_x/nx-os/system_management/configuration/guide/sm_nx_os_cg.html

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Module Self-Check

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

Q1) Which three features are the important features of the CMP? (Choose three.) (Source: Describing the Connectivity Management Processor)

A) dedicated operating environment
B) shared operating environment
C) login authentication
D) control capability:

Q2) Which two protocols can you use for access to CMP? (Choose two.) (Source: Describing the Connectivity Management Processor)

A) Telnet
B) Rlogin
C) HTTP
D) SSH

Q3) Which three types of servers can you use for user authentication? (Choose three.) (Source: Configuring User Management)

A) RADIUS
B) Web
C) LDAP
D) TACACS+

Q4) Write down the command that is used for establishing the SSH server on Cisco Nexus switches. (Source: : Configuring User Management)

Q5) How many users can be configured on Cisco Nexus switches? (Source: Configuring User Management)

A) 16
B) 32
C) 256
D) 128

Q6) Name at least three attributes a user account can have. (Source: Configuring User Management)

Q7) Which four configuration methods are available for configuring Cisco Nexus switches? (Choose four.) (Source: Describing System Management)

A) CLI
B) ASDM
C) XML
D) DCNM
E) RLOGIN
F) User defined GUI

https://t.me/learningnets

Q8) Name at least three system management features that you can find on Cisco Nexus switches. (Source: Describing System Management)

Q9) Which three types of clocks are common PTP devices? (Choose three.) (Source: Describing System Management)

A) boundary
B) limited
C) ordinary
D) transparent

Q10) Cisco SPAN supports Layer 3subinterfaces. (Source: Describing System Management)

A) true
B) false

Q11) What is ERSPAN? (Source: Describing System Management)

A) port mirroring traffic forwarded using IP network
B) production traffic forwarded over MPLS network
C) control traffic forwarded over ATM network
D) routing protocol running over IPX network

Q12) Cisco Smart Call Home provides an email-based notification for critical system policies. (Source: Describing System Management)

A) true
B) false

Q13) Which command validates the Cisco Smart Call Home destination profile? (Source: Understanding System Management)

A) **show smart-call**
B) **show callhome transport-email**
C) **show e-mail**
D) **show smartnet**

Q14) Which two Embedded Event Management actions are supported on Cisco Nexus switches? (Choose two.) (Source: Describing System Management)

A) device reload
B) generate a syslog message
C) running routing protocol
D) enabling interface

Q15) Which three SNMP versions are supported on Cisco NX-OS? (Choose three.) (Source: Describing System Management)

A) SNMPv1
B) SNMPv1c
C) SNMPv3
D) SNMPv2c

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Q16) Which task is required when configuring the vCenter solution? (Source: Understanding System Management)

A) configure interface mode
B) configure a distributed virtual switch (DVS)
C) enable high performance mode
D) configure the CNA for single-OS deployment

Q17) What should you do when upgrading the Cisco NX-OS Software release on a Cisco Nexus 7000 switch? (Source: Describe System Management)

A) Upgrade the EPLDs.
B) Perform the EPLDs upgrade in a nondisruptive fashion.
C) Check to see if the EPLD image upgrades are available on Cisco.com.
D) Verify if you can upgrade Cisco NX-OS Software without an EPLD upgrade.

https://t.me/learningnets

Module Self-Check Answer Key

Q1) A, B, D

Q2) A, D

Q3) A, C, D

Q4) ssh key rsa 1024 force
feature ssh

Q5) C

Q6) username, password, expiry date, user roles

Q7) A, C, D, F

Q8) Scheduler, SNMP, NTP and PTP, ERSPAN, SPAN, XML….

Q9) A, C, D

Q10) B

Q11) A

Q12) A

Q13) B

Q14) A, B

Q15) A, C, D

Q16) B

Q17) D

© 2012 Cisco Systems, Inc.

https://t.me/learningnets