DCUCD

Designing Cisco Data Center Unified Computing

Volume 2

Version 5.0

Student Guide

Text Part Number: 97-3177-01

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

Design Cisco Unified Computing Solutions ***4-1***

Overview 4-1
Module Objectives 4-1

Designing the Unified Computing Network **4-3**

Overview 4-3
Objectives 4-3
Define Fabric Interconnect Network Operational Modes 4-4
EHM Mode 4-4
Switching Mode 4-5
Understand Network Connectivity 4-7
Recognize Disjoint Layer 2 Domains 4-13
Define the Network High-Availability Aspects 4-17
Define VM-FEX Aspects 4-21
Summary 4-24

Designing Unified Computing Storage **4-25**

Overview 4-25
Objectives 4-25
Define Fabric Interconnect SAN Operational Modes 4-26
Understand SAN Connectivity 4-28
Define the SAN High-Availability Aspects 4-34
Summary 4-37

Designing the Virtual Access Layer **4-39**

Overview 4-39
Objectives 4-39
Cisco Nexus 1000V Overview 4-40
Cisco Nexus 1000V VMware Integration 4-53
Summary 4-67
Module Summary 4-69
Module Self-Check 4-71
Module Self-Check Answer Key 4-72

Design Cisco Unified Computing Solutions Server Deployment ***5-1***

Overview 5-1
Module Objectives 5-1

Designing Cisco UCS Server Deployment **5-3**

Overview 5-3
Objectives 5-3
Identify Server Deployment Aspects 5-4
Design Naming Convention 5-19
Design UUID Addressing 5-32
Design MAC Addressing 5-35
Design WWN Addressing 5-40
Design Common Policies 5-45
Summary 5-50

Designing Unified Computing Management **5-51**

Overview 5-51
Objectives 5-51
Design Management Access 5-52
Design Configuration Hierarchy 5-60
Design Remote Management Connectivity 5-64
Summary 5-67
Module Summary 5-69
Module Self-Check 5-71
Module Self-Check Answer Key 5-72

https://t.me/learningnets

Cisco Unified Computing Solution Applications **6-1**

Overview 6-1
Module Objectives 6-1

Designing Cisco Unified Communications on Cisco UCS **6-3**

Overview 6-3
Objectives 6-3
Cisco Unified Communications Overview 6-4
Cisco Unified Communications Characteristics 6-7
Selecting Cisco UCS Hardware for Cisco Unified Communications 6-15
Summary 6-24
References 6-24

Designing Distributed Computing on Cisco UCS **6-25**

Overview 6-25
Objectives 6-25
Distributed Computing Overview 6-26
Hadoop Architecture 6-31
Hadoop Performance Characteristics 6-35
Designing Cisco UCS for Greenplum MR Hadoop 6-42
Summary 6-48
References 6-48
Module Summary 6-49
Module Self-Check 6-51
Module Self-Check Answer Key 6-52

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

Module 4

Design Cisco Unified Computing Solutions

Overview

Network connectivity is vital to the Cisco Unified Computing System (UCS) solution design because it enables communication between clients and servers, different application instances, servers to storage, and so on.

This module introduces the network and storage connectivity and design aspects of Cisco UCS and also introduces the virtual access layer that is implemented with Cisco Nexus 1000V Series Switches.

Module Objectives

Upon completing this module, you will be able to employ design steps to create a scalable computing solution for a given set of requirements. This ability includes being able to meet these objectives:

- Design the LAN of the Cisco data center unified computing solution
- Design the SAN of the Cisco data center unified computing solution
- Recognize Cisco Nexus 1000V and aspects of Nexus 1000V design

https://t.me/learningnets

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

Lesson 1

Designing the Unified Computing Network

Overview

This lesson discusses the LAN design aspects of the Cisco data center unified computing solution.

Objectives

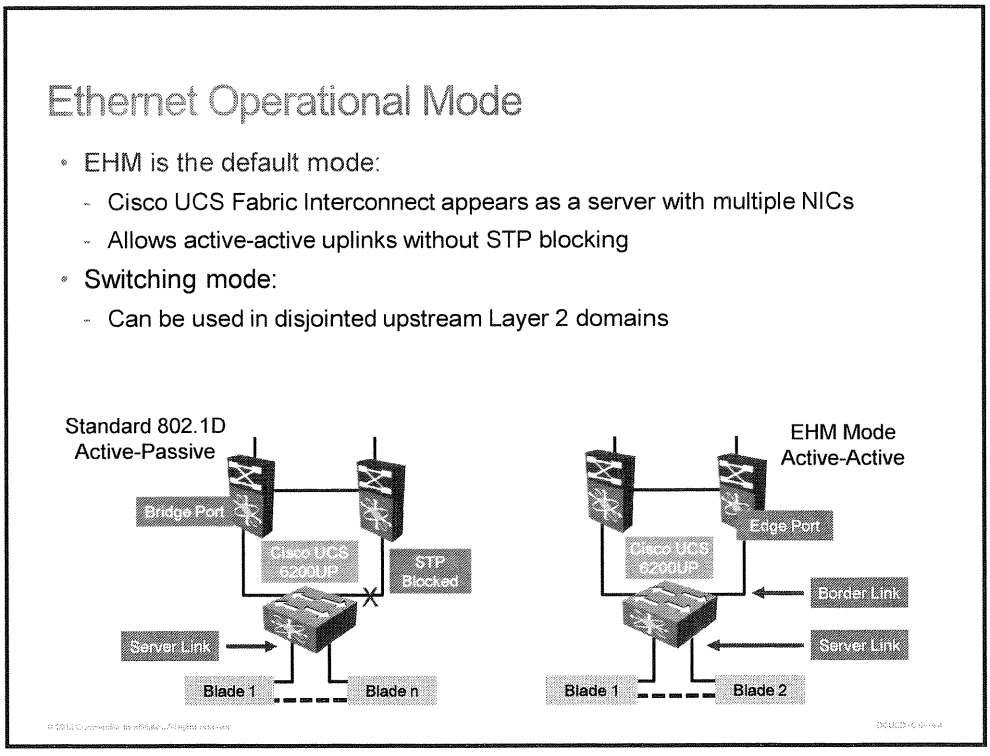
Upon completing this lesson, you will be able to design the LAN of the Cisco data center unified computing solution. You will be able to meet these objectives:

- Recognize the network operational modes of the Cisco UCS 6200 Series Fabric Interconnects
- Understand Cisco UCS network connectivity
- Recognize Layer 2 disjoint domain concepts and implications
- Define the network high-availability mechanisms for Cisco UCS network connectivity
- Define the VM-FEX requirements for the Cisco UCS B- and C-Series

https://t.me/learningnets

Define Fabric Interconnect Network Operational Modes

This topic describes how to set network operational modes of the Cisco Unified Computing System (UCS) 6200 Series Fabric Interconnects.



EHM Mode

The Cisco UCS Fabric Interconnect operates in an end-host mode (EHM) mode by default (which is also known as end-host virtualization). In the EHM mode, Cisco UCS appears to the external LAN as an end station with multiple adapters.

There are two types of links in the EHM operational mode:

- Server links
- Border links

Border links are Cisco UCS uplinks and can be in the form of a single link or aggregated in a channel. When operating in the EHM operational mode, Cisco UCS Fabric Interconnect does not participate in a Spanning Tree Protocol (STP) topology. Instead, the following is used to achieve loop-free topology:

- Border links must connect to Layer 2 network
- Traffic forwarding between border links is denied

The benefit of running a Cisco UCS Fabric Interconnect in the EHM mode is that the LAN STP topology is simplified and the STP domain size is reduced. Also, because no links are blocked by the STP, the active-active approach utilizes all redundant links to a Layer 2 network.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

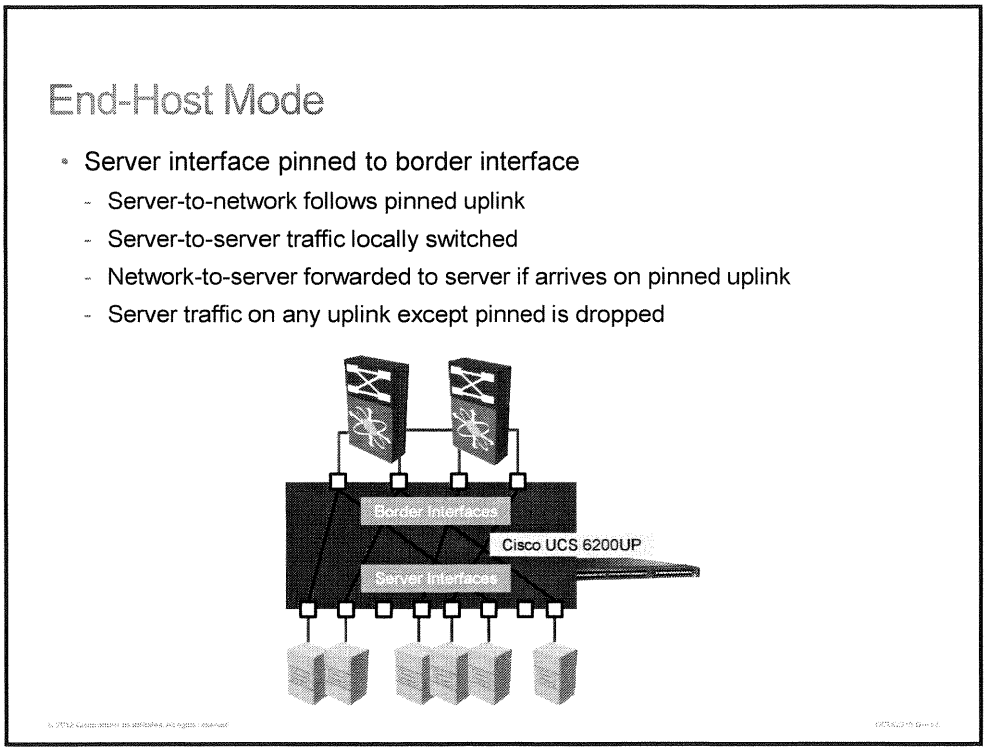
Switching Mode

In a normal LAN topology, STP protocols take care of the loops in the topology by disabling some of the links. Thus, the underlying network infrastructure is not fully utilized.

If desired, the fabric interconnects can be set to operate in traditional Ethernet switching mode.

This method introduces the need for STP to avoid broadcast loops. Despite operating in this manner, not all typical Cisco switch options are available. Thus, this deployment method is typically not recommended.

The switching mode is used in case of disjointed upstream Layer 2 domains at the outset.



In the EHM mode, each server link is pinned to exactly one border link. The pinning logic equally distributes server links to various border links.

Note that the server-to-server traffic is locally switched, while server-to-network traffic goes out on the pinned border link. To achieve local switching, the MAC addresses within the chassis are learned.

Network-to-server unicast traffic is forwarded to the server only if it arrives on a pinned border link, which is verified by the Reverse Path Forwarding (RPF) check. Server traffic that is received on any border link except the pinned border link is dropped (déjà vu check).

https://t.me/learningnets

End-Host Mode (Cont.)

- MAC address learning
  - Internal MAC addresses are learned
  - Learning is disabled on border interfaces
- Traffic forwarding
  - To server based on destination MAC
  - Learned MAC addresses age out when server link goes down or server is moved (repinned)

Traffic	Received on	Sent to
Learned Unicast	Server or uplink port	Server
Unknown Unicast	Server port	Pinned uplink and all active server ports
Unknown Unicast	Uplink port	Dropped
Broadcast/Unknown Multicast	Uplink port	All server ports
Broadcast/Unknown Multicast	Server port	Pinned uplink and all active server ports
Known Multicast	Server or uplink port	As per IGMP snooping

The MAC address learning in EHM mode is as follows:

- Learning is disabled on border links—network MAC addresses are never learned.
- Learning is enabled on server links—traffic to server is forwarded based on destination MAC address.

Learned MAC addresses never age unless the server link goes down or is deleted; in that case, server MACs can move (in case of repinning).

The table explains where certain traffic is sent when received.

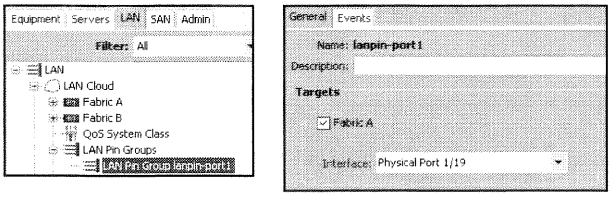
© 2012 Cisco Systems, Inc.
https://t.me/learningnets

Understand Network Connectivity

This topic discusses the Cisco UCS network connectivity.

LAN Pinning

- To identify the specific uplink port or port channel on one or both fabric interconnects
- Administrative pinning—using pin groups
  - Pin traffic from specific blade network adapter to specific uplinks with service profile
- Dynamic pinning (default)—not using pin groups
  - Cisco UCS automatically chooses uplink for adapter
  - Could choose incorrectly if LAN uplinks are on different Layer 2 domains



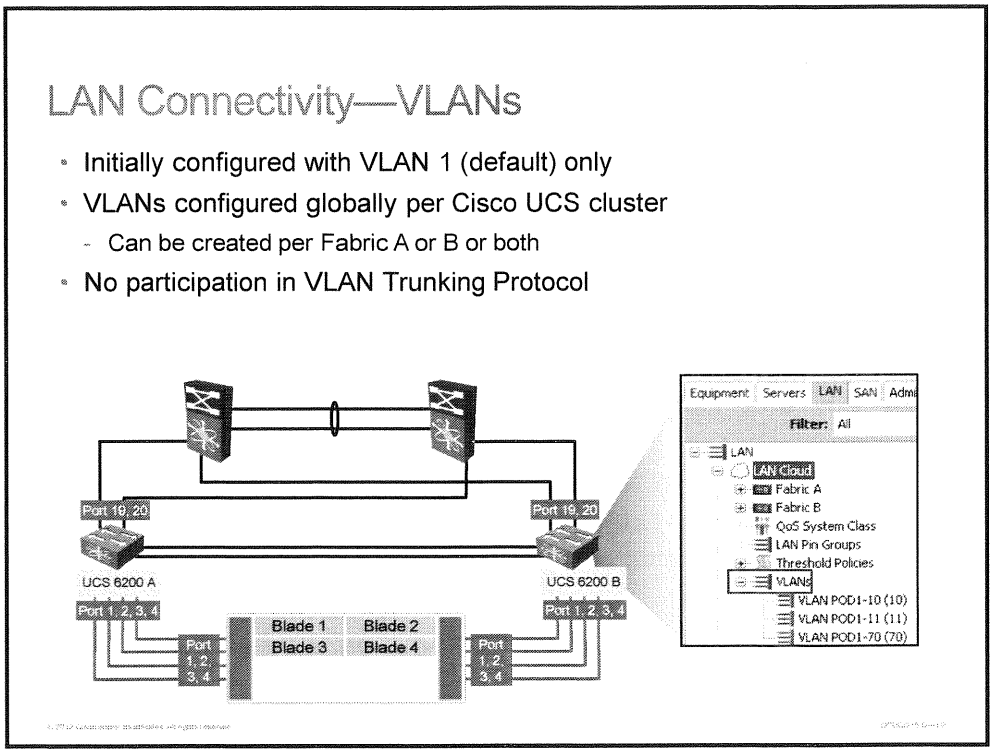
Cisco UCS allows the use of a mechanism called a pin group.

With pin groups, traffic from a specific blade server adapter is tied to a particular uplink port or port channel on each fabric interconnect. This configuration is achieved via the service profile configuration.

If pin groups are not used in service profiles, Cisco UCS automatically chooses an uplink port or port channel for the adapter on the blade associated with each profile.

If you need to force that specific blade to use a certain uplink port, the pin group can be applied via the service profile to achieve the forced use of the link. This might be the case when the uplinks are connected to different Layer 2 domains (VLANs). If left to default, the Cisco UCS Manager does not know what the correct uplink for that specific blade is.

https://t.me/learningnets



Cisco UCS uses VLANs to separate the traffic from different segments.

VLANs

VLANs in a Cisco UCS system are defined to connect the Cisco UCS with the external LAN networks. The VLANs defined in Cisco UCS should match the VLANs defined on the other side of the uplink ports because they are trunked using IEEE v802.1q.

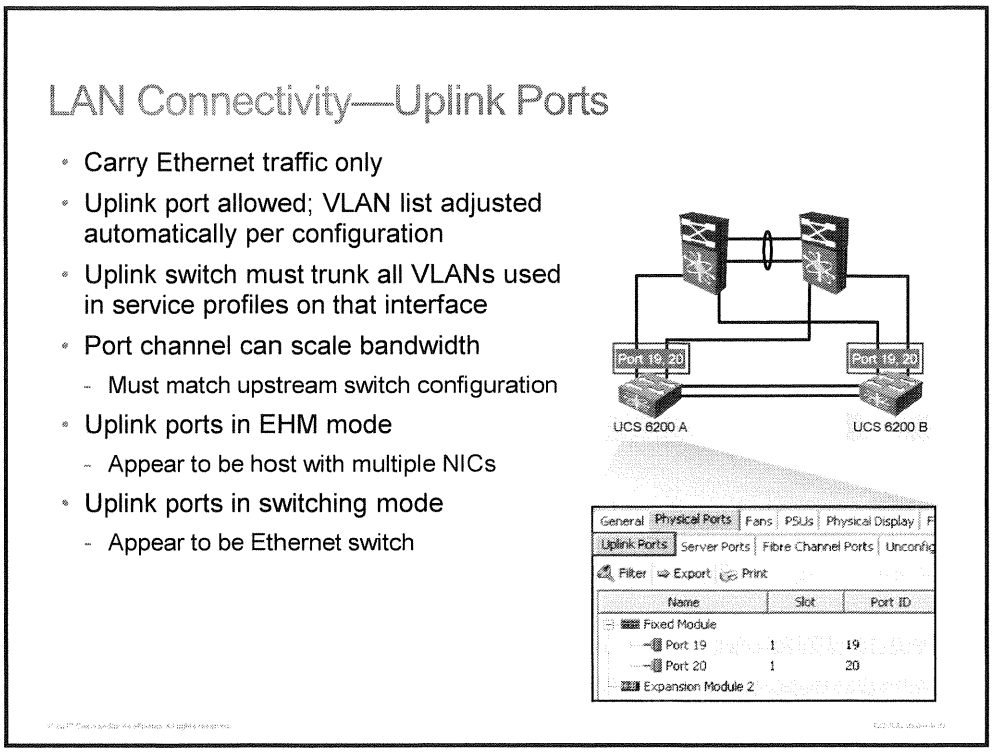
Initially, a new Cisco UCS system is configured with default VLAN 1 only, which is not recommended.

VLAN configuration in Cisco UCS is done globally for the whole system, but it can later be assigned to specific service profiles. Because Cisco UCS in a redundancy setup consists of Fabric A and Fabric B, the VLANs can also be designed per individual fabric or independent of the fabric (which enables the failover option if later used in a service profile).

When created in the Cisco UCS Manager, the VLAN needs the name (which is later used in other configuration parts) and the valid IEEE 802.1Q VLAN ID.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets



Uplink Ports

Uplink ports in a Cisco UCS are those physical ports on the fabric interconnects that are dedicated for connectivity to a LAN device external to Cisco UCS (for example, for the connectivity to Cisco Nexus 7000).

These ports can be either from the fixed port range or from the expansion module, if present.

The ports carry VLAN traffic only and are configured as 802.1Q trunk ports to carry VLAN traffic for all of the VLANs used in the service profiles for a particular fabric. The allowed VLAN list is adjusted automatically based on Cisco UCS VLAN configuration.

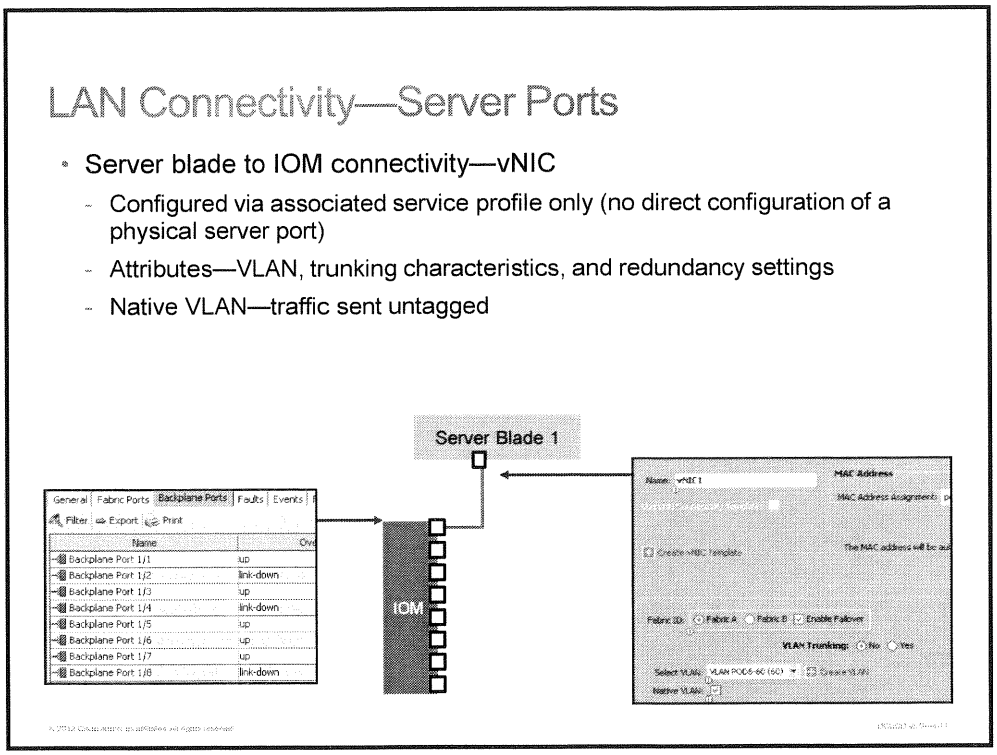
Depending on the Cisco UCS configuration, an uplink port carries VLANs that belong to the fabric the uplink port is part of (A or B). Those VLANs that are not fabric-dependent are defined globally in the LAN cloud.

Port Channels

Uplink ports bandwidth can be scaled using port channels, which use the Link Aggregation Control Protocol (LACP) 802.3ad protocol. The configuration of the port channel must match the other side. (VLANs being trunked must be the same.)

Port channel can be configured using uplink ports from a single Cisco UCS 6200XP Fabric Interconnect in a cluster—that is, from the same fabric.

https://t.me/learningnets



Server Blade to IOM Server Ports—vNIC

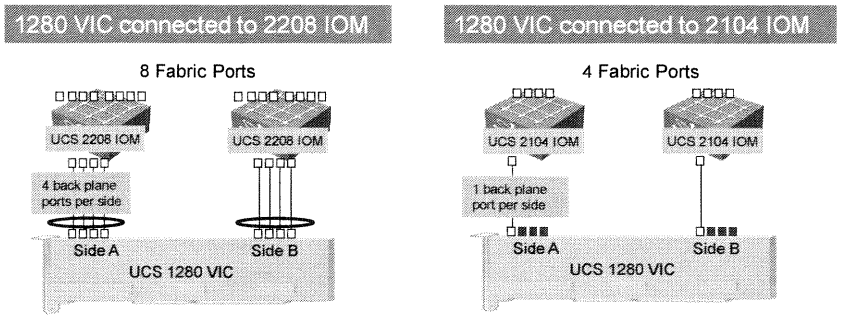
The communication between the server blade and the I/O module (IOM) is governed by the service profile configuration. This configuration includes virtual network interface cards (vNICs), which define how a server connects to the LAN network.

A vNIC has a number of parameters associated with it, the most important of which are VLAN, trunking characteristic (trunk versus nontrunk interface), and redundancy setting, which defines whether, upon primary fabric failure, the communication fails over to the second (for example, from Fabric A to Fabric B, or vice versa).

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

UCS 2.0 Connectivity – Adapter to IOM

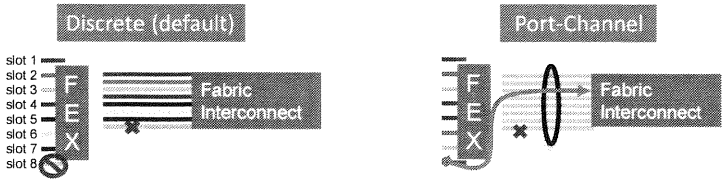
- 40-GB host connectivity per side (i.e., 4 x 10 G links per side).
- 10-GB host connectivity per side (i.e., 1 x 10 G links per side).
- Only one back plane port is brought up—same as current-generation hardware.



With the 2.0 UCSM and the new VIC 1280, the connectivity from the server can be scaled with port channels. The functionality is available in combination with the 2208 IOM.

UCS 2.0 Connectivity – IOM to FI

- Statically pinned to individual fabric links
- Statically pinned to logical link (port channel)



The connectivity from the chassis—effectively, the IOM—can also be scaled with the second generation of the IOMs. Instead of having a discrete mode where server interfaces are statically pinned to the uplinks, the port channel mode distributes the traffic from the server vNIC among the uplinks.

https://t.me/learningnets

UCS 2.0 Connectivity Interoperability

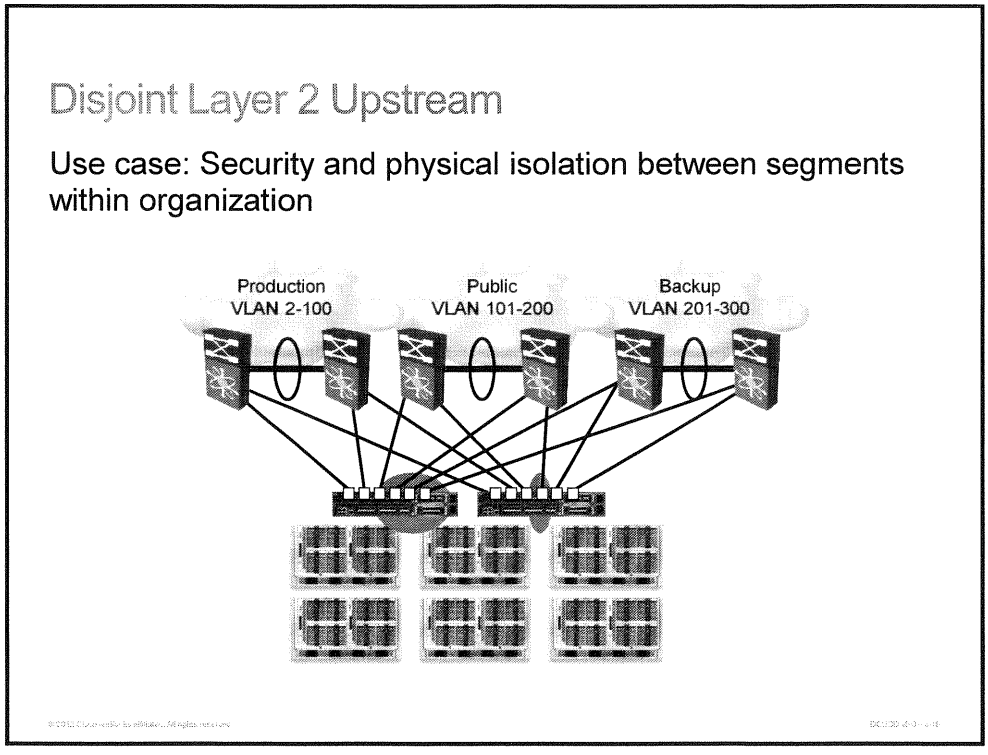
Fabric Interconnect	IOM	Adapter	Connectivity benefits per side
6100	2104	UCS M81 KR	Fabric and host connectivity same as Balboa 1.4(1)
6100	2208/2204	UCS M81 KR	All 8/4 IOM ports in discrete mode 10 GB to the host
6100	2104	UCS1280/1240 VIC	4 IOM ports in discrete mode 5 GB to the host
6100	2208/2204	UCS1280 VIC	All 8/4 IOM ports in discrete mode 40 GB to the host
6200	2104	UCS M81 KR	4 IOM ports in discrete mode 5 GB to the host
6200	2208/2204	UCS M81 KR	All 8/4 IOM ports can be leveraged in discrete mode 10 GB to the host
6200	2104	UCS1280 VIC	All 8 IOM ports in discrete mode 10 GB to the host
6200	2208/2204	UCS1280 VIC	All 8/4 IOM ports can be leveraged in port-channel mode 40 GB to the host

The table lists fabric interconnect, IOM, and virtual interface card (VIC) adapter combinations and provides descriptions of what is available to the server from the aspect of connectivity.

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

Recognize Disjoint Layer 2 Domains

This topic describes Layer 2 disjoint domains concepts and implications.

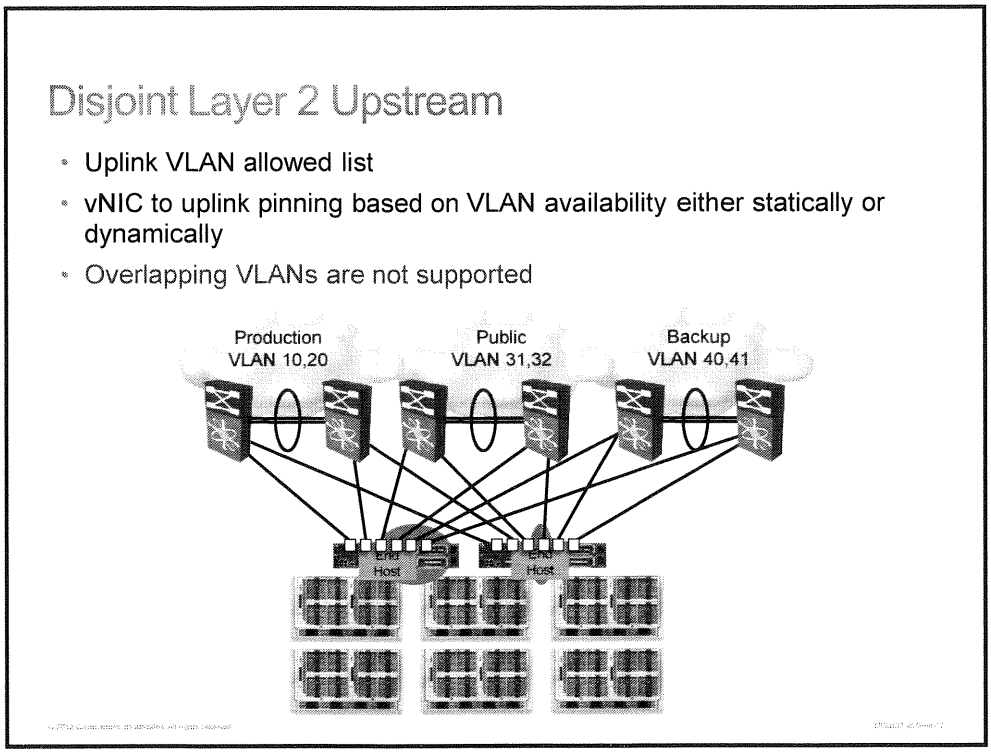


Upstream disjoint Layer 2 networks are required if you have two or more Ethernet "clouds" that never connect, but they must be accessed by servers or virtual machines located in the same Cisco UCS instance. For example, you could configure disjoint Layer 2 networks if you require one of the following:

- Servers or virtual machines to access a public network and a backup network
- In a multitenant system, servers or virtual machines for more than one customer are located in the same Cisco UCS instance and need to access the Layer 2 networks for both customers

By default, data traffic in Cisco UCS works on a principle of mutual inclusion. All traffic for all VLANs and upstream networks travels along all uplink ports and port channels. If what you have set up does not support upstream disjoint Layer 2 networks, you must assign the appropriate uplink interfaces to your VLANs or traffic for those VLANs continues to flow along all uplink ports and port channels.

https://t.me/learningnets



The Layer 2 disjoint networks operate on a principle of selective exclusion. Traffic for a VLAN that is designated as part of a disjoint network can only travel along an uplink Ethernet port or port channel that is specifically assigned to that VLAN and is selectively excluded from all other uplink ports and port channels. However, traffic for VLANs that are not specifically assigned to an uplink Ethernet port or port channel can still travel on all uplink ports or port channels, including those that carry traffic for the disjoint Layer 2 networks. In Cisco UCS, the VLAN represents the upstream disjoint Layer 2 network. When you design your network topology for disjoint Layer 2 networks, you must assign uplink interfaces to VLANs and not the reverse.

Cisco UCS only supports disjoint Layer 2 networks when the Ethernet switching mode of the fabric interconnects is configured for end-host mode.

Cisco UCS does not support overlapping VLANs in disjoint Layer 2 networks. You must ensure that each VLAN only connects to one upstream disjoint Layer 2 domain.

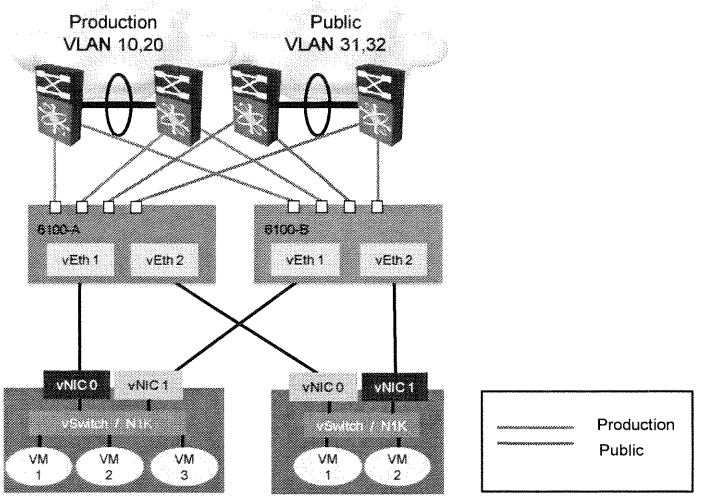
vNIC Requirements

A vNIC can only communicate with one disjoint Layer 2 network. If a server needs to communicate with multiple disjoint Layer 2 networks, you must configure a vNIC for each of those networks. To communicate with more than two disjoint Layer 2 networks, a server must have a Cisco VIC adapter that supports more than two vNICs.

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

Disjoint Layer 2 Upstream (Cont.)

- vNIC considerations for vSwitch and bare metal



The following considerations apply for vSwitch and bare-metal implementations:

- A vNIC can only participate in single Layer 2 network upstream. All selected VLANs need to map to the same Layer 2 network.
- A vNIC pair per Layer 2 network is highly recommended
- Cisco virtualized adapters are recommended to support multiple Layer 2 networks for redundancy and load balancing
- A maximum of 31 disjoint networks is supported

https://t.me/learningnets

Disjoint Layer 2 Maximums

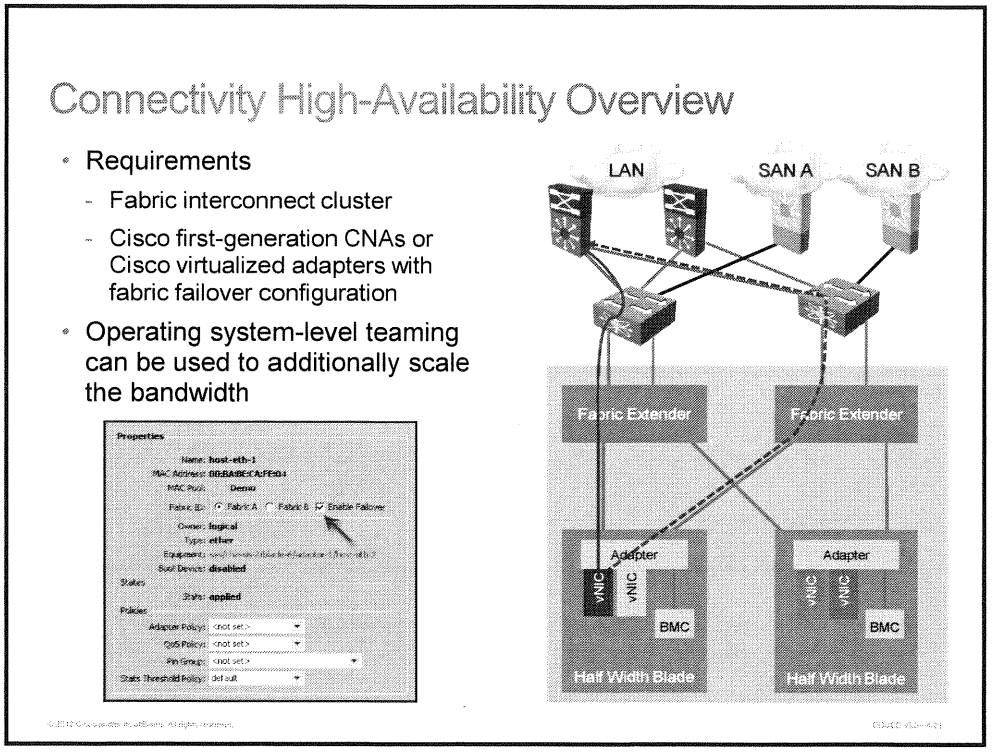
Adapter	VM-FEX Dynamic	Static	Fabric Failover
1280	114	26 with ESX **	Yes
VIC	114 or 56 *	26 with ESX **	Yes
Emulex Gen 1	NA	2	Yes
QLogic Gen 1	NA	2	Yes
Emulex Gen 2	NA	2	No
QLogic Gen 2	NA	2	No

The maximums listed in the table apply when implementing Layer 2 disjointed domains. The number of disjoint Layer 2 networks per blade is dependent on the adapter type.

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

Define the Network High-Availability Aspects

This topic describes the network high-availability mechanisms for Cisco UCS network connectivity.

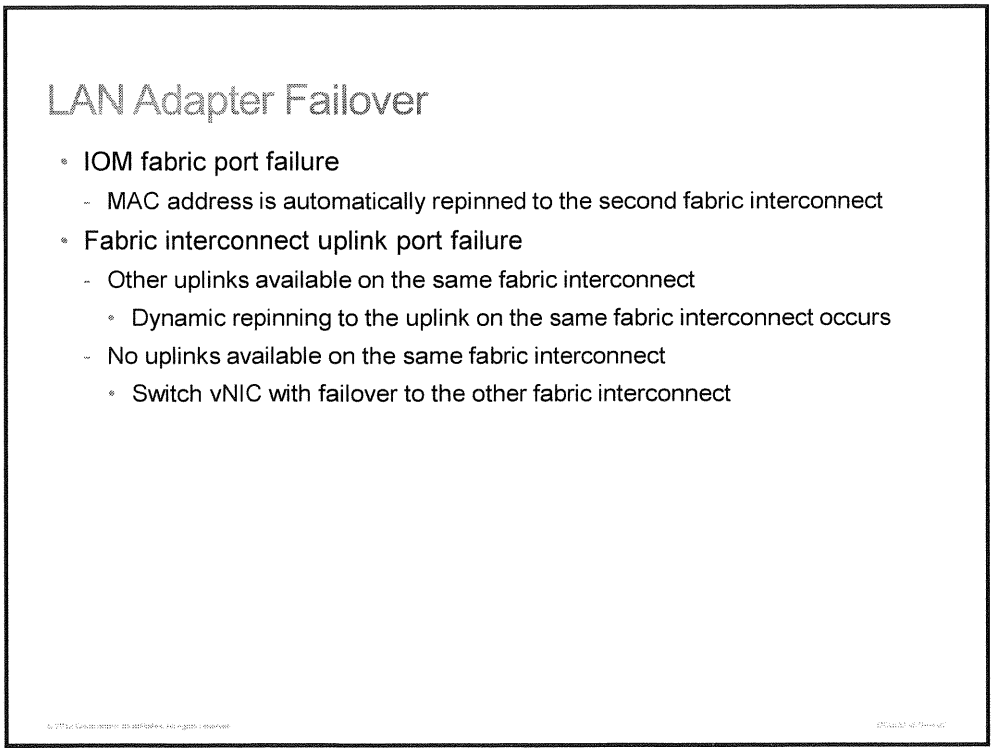


To deploy fabric failover for server connectivity, the following requirements must be met:

- Cisco UCS must be deployed in a cluster.
- The server must be equipped with either first-generation converged network adapters (CNAs) or Cisco virtualized adapters.
- Fabric failover must be configured for LAN adapters.

The LAN-based hardware failover eliminates the need for operating system-level teaming for the failover purpose. Still, the operating system-level teaming can be used to scale the bandwidth when at least two LAN adapters are used.

https://t.me/learningnets



Within Cisco UCS, the LAN adapter can manage either IOM fabric port failure or fabric interconnects uplink port failure.

When IOM fabric port failure occurs, the LAN adapter is automatically repinned to the second fabric.

When fabric interconnect uplink port failure occurs, two scenarios are possible:

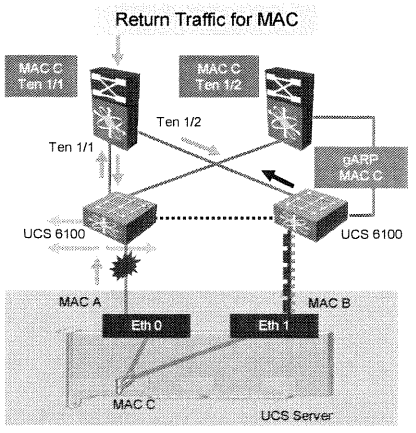
- When other uplinks on the same fabric interconnect are available, dynamic repinning to the uplink of the same fabric interconnect is performed.
- When no other uplinks on the same fabric interconnect are available, fabric failover to the second fabric occurs.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

Fabric Failover and Synchronization

- Fabric Failover
  - When active path fails, failover to standby fabric occurs
  - 6x00 updates path changes to upstream switches via gratuitous ARP and also for vNIC MAC upon active link failure



Fabric failover is the advanced functionality available with Cisco UCS, with the following characteristics:

- Fabric provides redundant path for each vNIC
- Hardware-based active/standby failover mechanism is implemented for Ethernet traffic
- Host-operating system is unaware of failure and recovery
- Unlike operating system NIC teaming, redundancy is provided with a single interface
- Fabric failover load-balances traffic on a per-vNIC basis
- Each vNIC maps to one of the fabrics for active data path

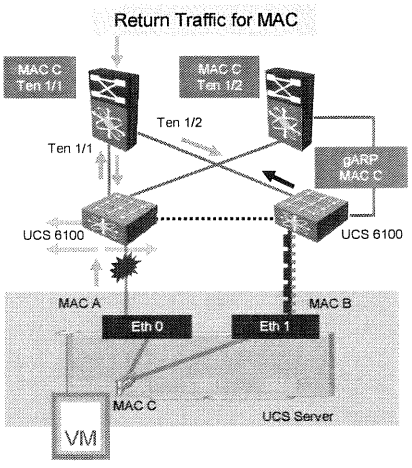
When the active path fails, failover to standby fabric occurs and UCS 6200XP updates path changes to the upstream switches via gratuitous Address Resolution Protocol (ARP) of vNIC MAC upon active link failure. This is performed only for the static vNICs with the following problems:

- Virtual machine (VM) MAC behind a vSwitch is not updated to upstream switches
- For silent VMs (one-way communication), it does not provide a path update

https://t.me/learningnets

Fabric Failover and Synchronization (Cont.)

- Fabric Synchronization
  - Maintains synchronized MAC address tables between 6100XP
  - 6x00 updates path changes to upstream switches via gratuitous ARP and also for VM MACs upon active link failure



With fabric synchronization, the MAC address tables are synchronized between Cisco UCS 6x00 XP fabric interconnects. The UCS 6x00XP updates path changes to upstream switches via gratuitous ARP not only for vNIC MACs, but also for VM MACs behind a vSwitch upon active link failure.

The feature is always enabled from version 1.4 of the software onwards.

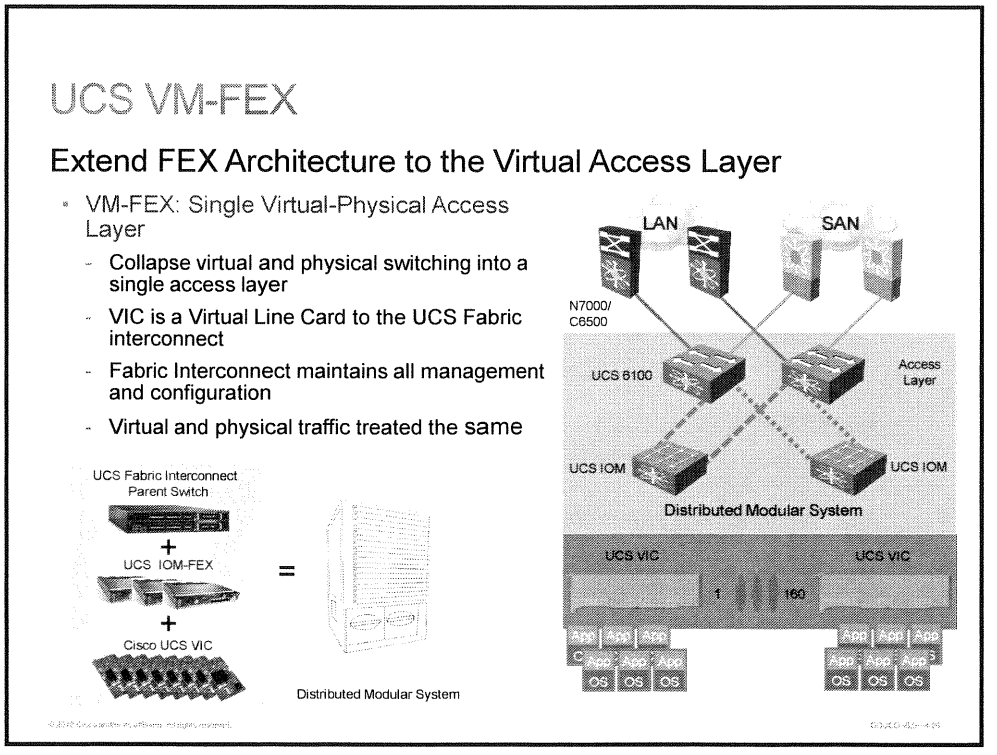
The feature is specifically useful for Hyper-V deployments, because Hyper-V officially does not support NIC teaming or bonding. Therefore, the UCS hardware-based failover provides redundancy. With fabric synchronization, the VMs running on the Hyper-V in Cisco UCS benefit from MAC table synchronization.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

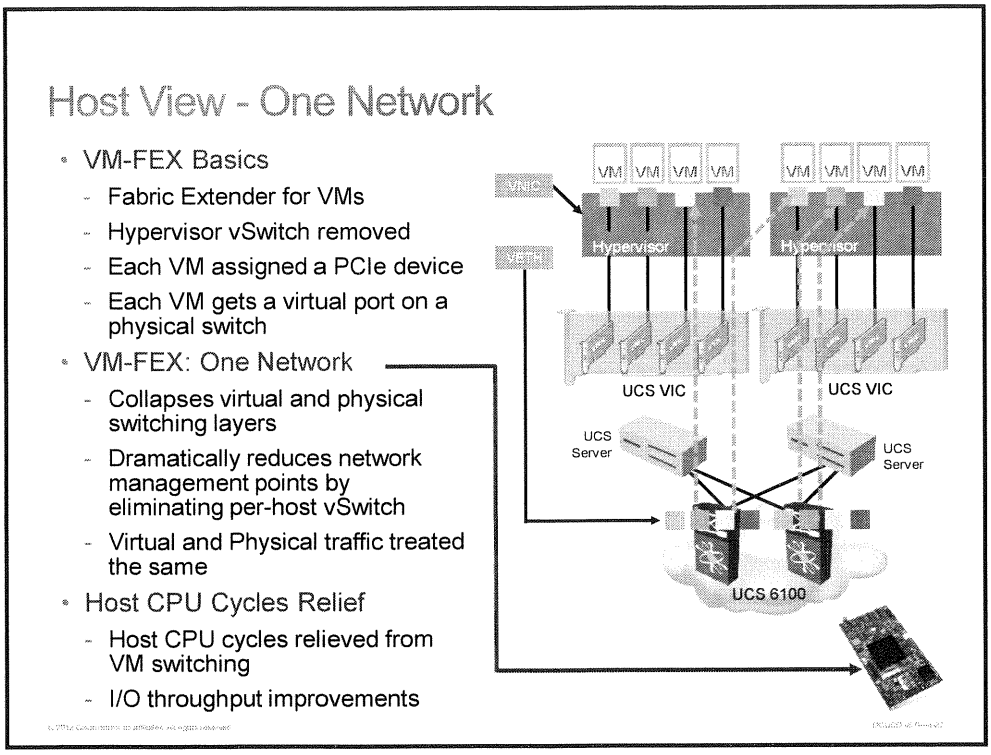
Define VM-FEX Aspects

This topic describes the Virtual Machine Fabric Extender (VM-FEX) requirements for the Cisco UCS B- and C-Series.



Cisco VM-FEX technology extends fabric interconnect ports directly to virtual machines, eliminating software-based switching in the hypervisor. Cisco VM-FEX technology collapses virtual and physical networking infrastructure into a single infrastructure that is fully aware of locations and networking policies of the virtual machines. Cisco FEX-VM technology is implemented by Cisco VICs with a prestandard implementation of IEEE 802.1Qbh Port Extender.

https://t.me/learningnets



A Cisco VIC adapter, such as the Cisco UCS M81KR Virtual Interface Card, is a converged network adapter (CNA) designed for both single-OS and VM-based deployments. The VIC adapter supports static or dynamic virtualized interfaces, including up to 128 vNICs).

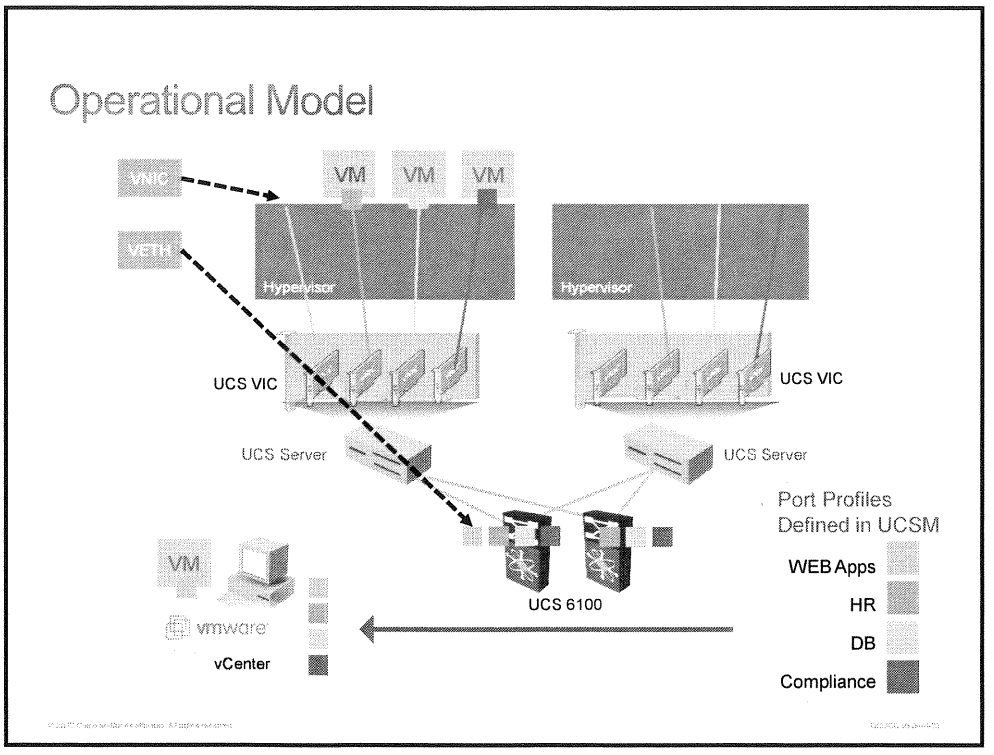
VIC adapters support VM-FEX to provide hardware-based switching of traffic to and from virtual machine interfaces.

In a VMware environment, VM-FEX supports the standard VMware integration with VMware ESX hypervisors installed on the server and all virtual machine management performed through the VMware vCenter.

The Cisco UCS Virtual Interface Card 1280 supports up to 256 dynamic virtual adapters and interfaces.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets



VM-FEX ports can operate in standard mode or high-performance mode.

Standard Mode

In standard mode, traffic to and from a virtual machine passes through the distributed virtual switch (DVS) and the hypervisor.

The figure shows the traffic paths taken by VM traffic on a Cisco UCS server with a VIC adapter that has VM-FEX ports in standard mode.

High-Performance 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:

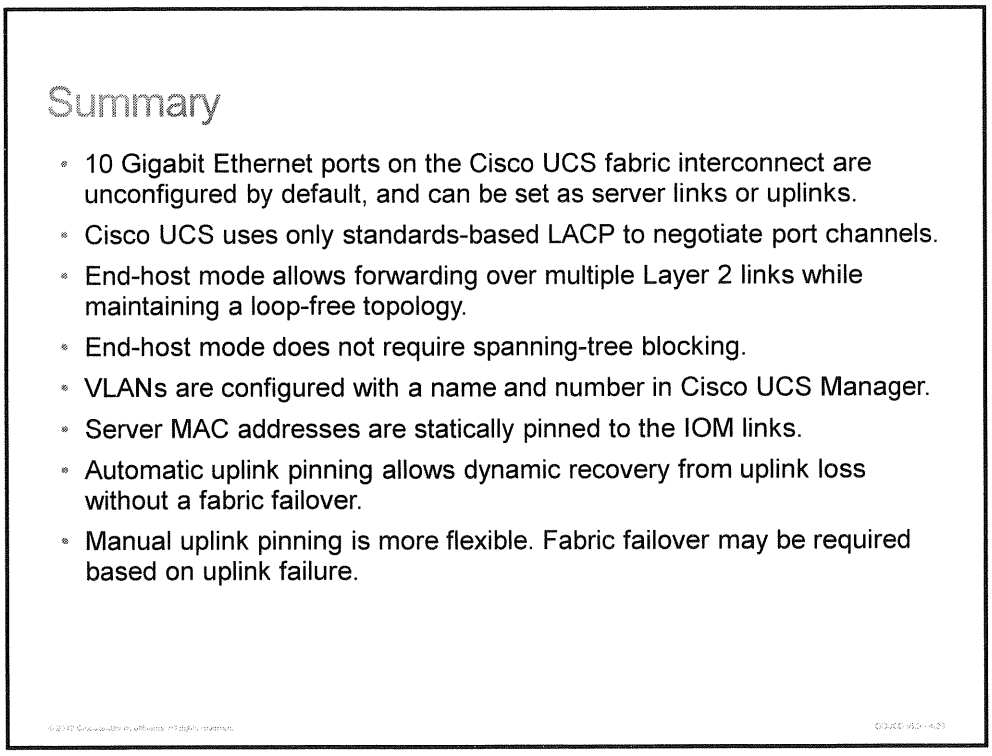
- Increases I/O performance and throughput
- Decreases I/O latency
- Improves CPU utilization for virtualized I/O-intensive applications

With VMware, high-performance mode also supports VMotion. During VMotion, 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.

https://t.me/learningnets

Summary

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



© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Lesson 2

Designing Unified Computing Storage

Overview

This lesson discusses the SAN design aspects of the Cisco data center unified computing solution.

Objectives

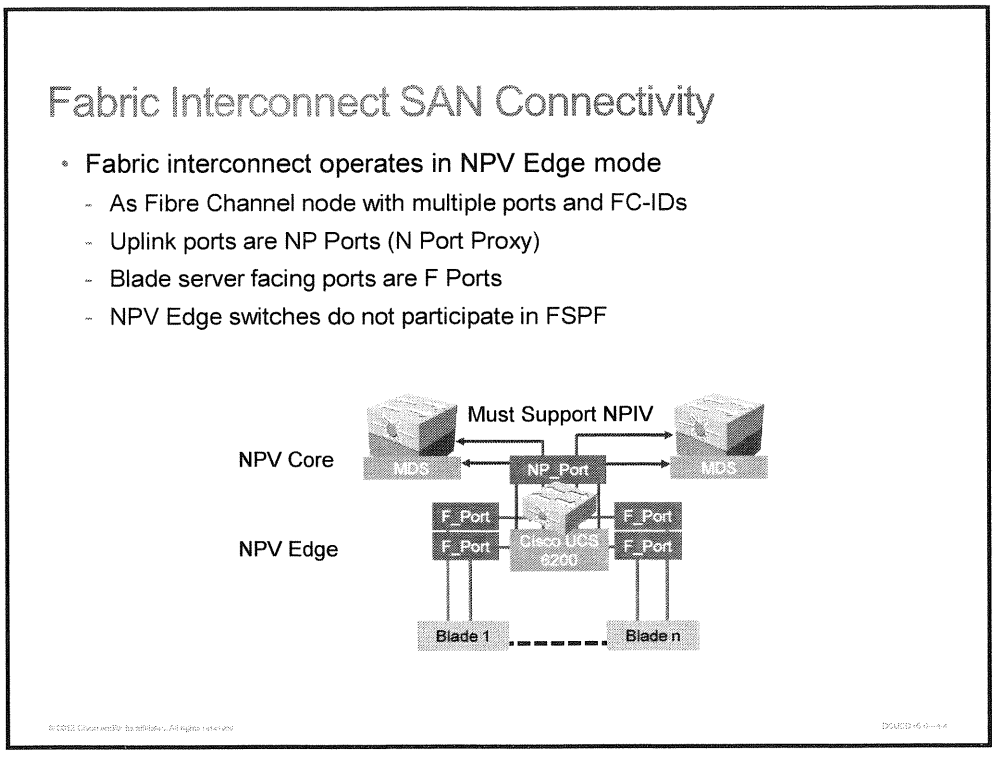
Upon completing this lesson, you will be able to design the SAN of the Cisco data center unified computing solution. You will be able to meet these objectives:

- Recognize SAN operational modes of Cisco UCS 6200 Series Fabric Interconnects
- Understand SAN connectivity design aspects
- Define the SAN high-availability mechanisms for Cisco UCS 6200 Series Fabric Interconnects

https://t.me/learningnets

Define Fabric Interconnect SAN Operational Modes

This topic describes SAN operational modes of Cisco Unified Computing System (UCS) 6200 Series Fabric Interconnects.



The Cisco UCS fabric interconnects operate in the N-Port Virtualizer (NPV) edge mode.

The upstream Fibre Channel switch (for example, MDS) must support and be enabled with the N-Port ID Virtualization (NPIV) feature, which allows multiple Fibre Channel IDs (FCIDs) to be assigned a single node port (N Port).

In the NPIV topology, there are two types of interfaces:

- **Server interface:** The server-facing interface is either physical Fibre Channels or virtual Fibre Channel interfaces operating in a fabric port (F Port) mode.
- **Border interface:** Border interfaces are network-facing and always operate in a proxy N-Port mode.

There is no local switching of the Fibre Channel traffic on the Cisco UCS 6200XP; all packets are forwarded to the NPV core switch.

The fabric login (FLOGI)-related processing is relayed in software (FLOGI, fabric discover [FDISC], and corresponding link service accept [LS_ACC], link service reject [LS_RJT], and so on) to the same uplink interface.

Note	Every uplink can be connected to different Fibre Channel switches and VSANs.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

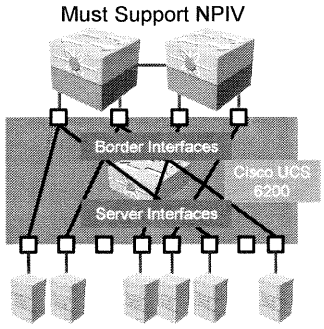
N-Port Virtualizer

NPV edge mode—default

- Each server interface pinned to one border interface
- Pinning logic distributes server interfaces between border interfaces (round-robin)
- All traffic follows the pinned port
- All traffic passed to upstream device for switching
- Supports nested NPIV

Switching mode

- Ability to turn on/off
- Limited Fibre Channel switching functionality
- No zoning configuration
- Zoning may be inherited from external switch

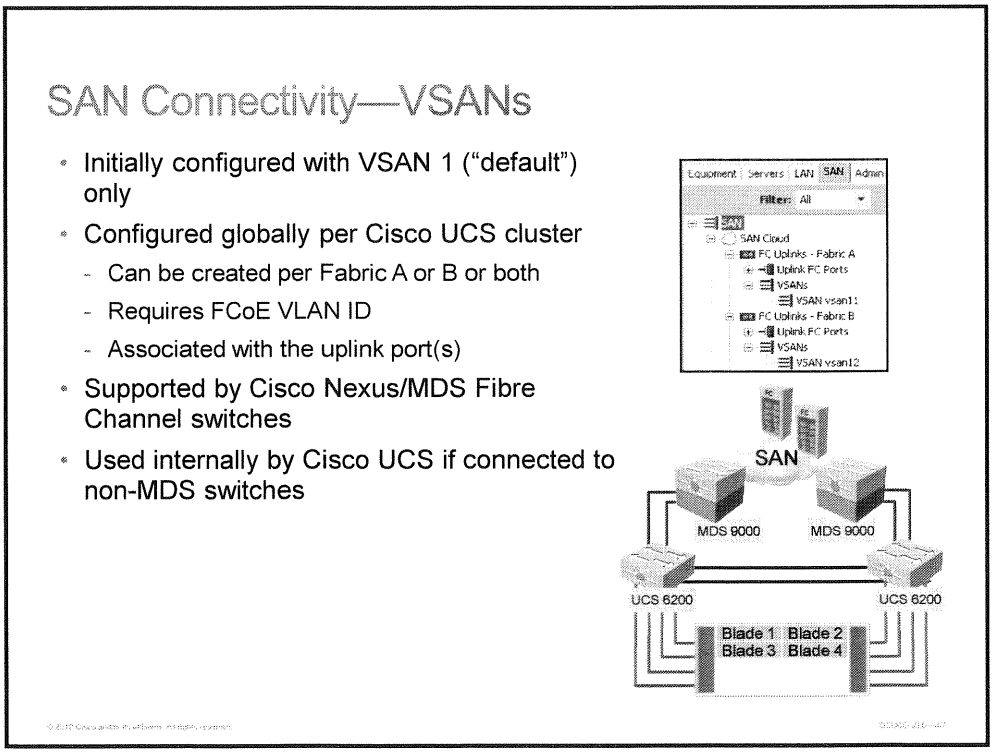


Cisco UCS 6x00XP in NPV edge mode pins each server link to exactly one border link. The pinning logic load-balances server links to various border links, while all traffic is forwarded to the upstream SAN device for switching.

https://t.me/learningnets

Understand SAN Connectivity

This topic discusses SAN connectivity design aspects.



Cisco UCS SAN connectivity consists of these elements:

- Virtual storage area networks (VSANs): Defined to separate traffic for different Fibre Channel fabrics
- Uplink connectivity: Connect the Cisco UCS system to the external SAN network
- Server connectivity: Connect the blade servers via uplink ports to the external SAN network

VSANs

VSANs in a Cisco UCS system are defined to connect the Cisco UCS with the external SAN networks. The VSANs defined in Cisco UCS should match the VSANs defined on the other side of the uplink port. Initially, a new Cisco UCS system is configured with VSAN 1 only as the default, which is not recommended.

VSAN configuration in Cisco UCS is done globally for the entire system, but it can be assigned later to specific service profiles. Because Cisco UCS in a redundancy setup consists of Fabric A and Fabric B, the VSANs can also be designed for an individual fabric or independent of the fabric.

When created in the Cisco UCS, the VSAN needs the following:

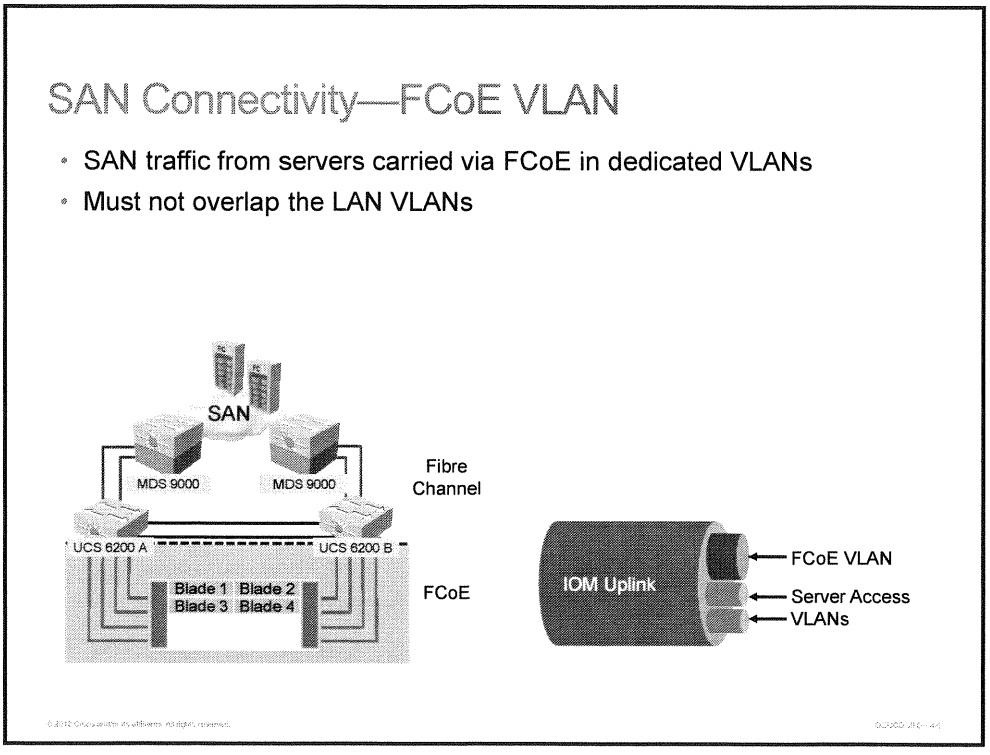
- Name, which is later used in other configuration parts
- VSAN ID
- Fibre Channel over Ethernet (FCoE) VLAN ID, used to carry the VSAN traffic over from the server blade to the I/O module (IOM) to the fabric interconnect

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

If you want to support any VSAN, it needs to be configured globally into Cisco UCS Manager, and it can then be associated with a particular virtual host bus adapter (vHBA). The default VSAN is preconfigured into Cisco UCS Manager and is automatically chosen as the default connectivity for each vHBA.

While Cisco Multilayer Director Switches (MDS) use VSANs, other vendors have not yet implemented VSAN technology within their switching architectures. Cisco UCS continues to use VSANs internally to isolate physical fabrics. Each uplink connects to one physical fabric and is mapped internally to one or more VSANs. Through this method, the same Cisco UCS fabric interconnect could be connected via uplink to multiple, physically separate Fibre Channel fabrics without causing those fabrics to merge. All Fibre Channel services would be kept isolated by VSANs. No inter-VSAN Routing (IVR) is possible.

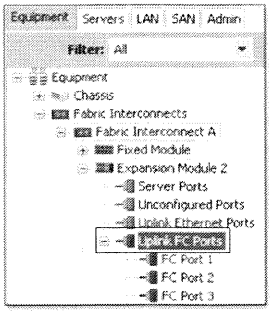


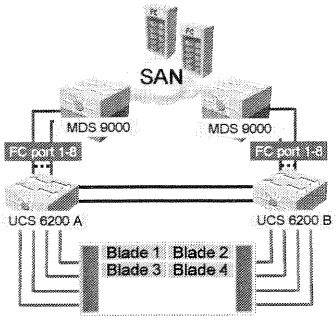
Within Cisco UCS, an FCoE VLAN ID is used for VSAN. The FCoE VLAN ID should not overlap with regular VLANs used for LAN connectivity.

https://t.me/learningnets

SAN Connectivity—Uplink Ports

- Connectivity to external SAN devices
  - Carries Fibre Channel traffic only
- Fibre Channel-enabled ports (fixed or expansion module based)





Uplinks in a Cisco UCS are the fixed physical Fibre Channel ports or those on the expansion modules used for connectivity to a SAN device external to Cisco UCS (for example, for the connectivity to MDS 9000). Note that a single port carries traffic for one VSAN only.

Depending of the Cisco UCS configuration, an uplink port carries VSANs that belong to its fabric (A or B).

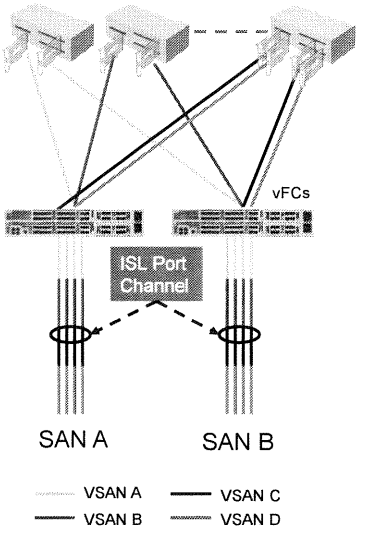
A single uplink port carries traffic for one or more VSANs (trunking), meaning it is connected to one or more physical fabrics and is mapped internally to one or more VSAN numbers. The same Cisco UCS Fabric Interconnect can be connected via uplinks to multiple, physically separate Fibre Channel fabrics without causing those fabrics to merge. All Fibre Channel services are thus kept isolated using VSANs, and no IVR is possible.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Fibre Channel PortChannels and Trunking

- Up to 16 Fibre Channel ports aggregated in a single port channel
  - Different combination of Fibre Channel ports from different expansion modules on the FI can be placed on the same port channel
  - In case of port speed mismatch, port channel forces port speed to highest commonly supported speed
- VSANs can be trunked over the port channel
- VSAN trunking and port channel supported for both NPV and switch mode FI operation



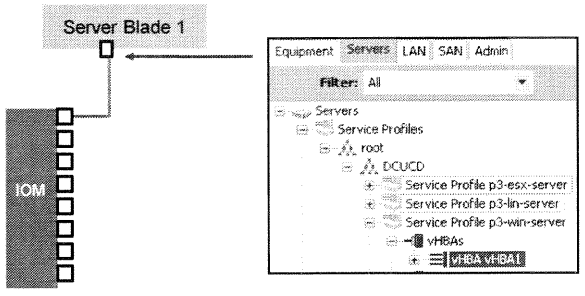
There are two features that can be used to scale the SAN implementation:

- Fibre Channel port channels:
  - Up to 16 Fibre Channel ports can be aggregated together for a single port channel
  - Different combination of Fibre Channel ports on the UCS 6200XP can be placed on the same port channel
  - In case of port speed mismatch, the port channel forces port speed to the highest commonly supported speed
- Fibre Channel port trunking:
  - VSANs can be trunked Fibre Channel ports
  - VSAN trunking and port channel are supported for both NPV and switch mode fabric interconnect operation

https://t.me/learningnets

SAN Connectivity—Server Ports

- Defined in a service profile with vHBA:
  - Assigned to a single VSAN
  - VSAN and properties assigned dynamically via service profile
- VSAN used internally to isolate fabrics even if uplinks are connected to non-MDS switches



A server SAN port is configured as a vHBA that corresponds to VSAN. The concept is similar to the one used in LAN: VLAN/vNIC.

For the server to be connected to the SAN, the service profile must be configured with the vHBA, and a VSAN must be selected. A vHBA configuration is applied to the Fibre Channel interface on the physical blade when the service profile is associated with the blade server.

Before the VSAN is associated with vHBA, it must be configured globally in Cisco UCS Manager.

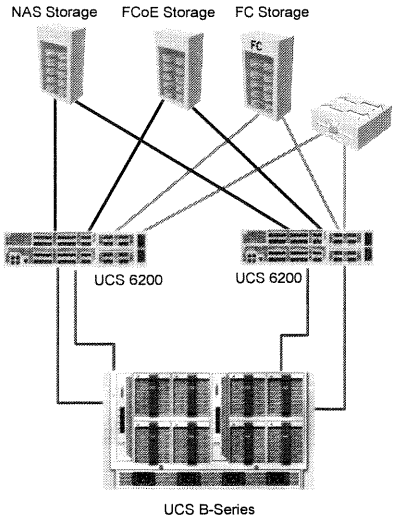
VSANs are supported on Cisco MDS switches but not by other vendors even though Cisco UCS internally still uses VSANs to distinguish between and isolate the fabrics.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Direct Connection of Storage

- Support for NetApp and EMC DAS
- Ethernet and Fibre Channel switching modes are independent
- Allow direct-connected NAS devices in EHM mode
  - New port type allows customers with direct connect storage to leave switch mode off
  - New "appliance" port type for Ethernet appliances and direct connect storage



Cisco UCS supports directly connected storage that can be Fibre Channel, FCoE, or Network File System (NFS).

The NFS type of appliance is denoted as an Ethernet appliance that can be any specialized device for use on an Ethernet network (for example network-attached storage [NAS], Internet Small Computers Interface [iSCSI], security appliances, or Cisco Nexus 1010) that does not run STP. Note that this type of port should not be used for switch connectivity in order to avoid traffic loops.

Even though new types of storage attachment are available, the same rules apply: There is no hardware-assisted fabric failover for storage due to the complexity of the operation. A traditional SAN host multipathing must be used for fabric failover.

The direct attached storage brings several benefits for the customer:

- Cisco UCS 6x00XP switch mode no longer needed for such environments
- External Fibre Channel switch may still be required but no longer in data path (if zoning has to be implemented)

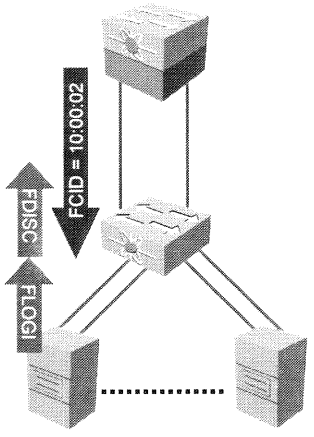
https://t.me/learningnets

Define the SAN High-Availability Aspects

This topic describes the SAN high-availability mechanisms for Cisco UCS 6200 Series Fabric Interconnects.

Pinning and Load Balancing

- Server interfaces pinned to uplink interfaces (not FLOGI)
- Traffic on server interface
  - Sent to the pinned interface
  - No forwarding lookup
    - NPV switch does not participate in FSPF
  - Binding check performed to verify frame SID is on the right server interface
    - Prevent address spoofing
- Traffic on border interface
  - Forwarding lookup is performed per frame destination ID (DID)
  - DID points to the server interface
  - Packets arriving on a border interface with a wrong or nonexistent DID are discarded and are not forwarded to the server interface



With NPV edge mode, each downstream device (server or blade server) is pinned to an uplink port based on a round-robin algorithm.

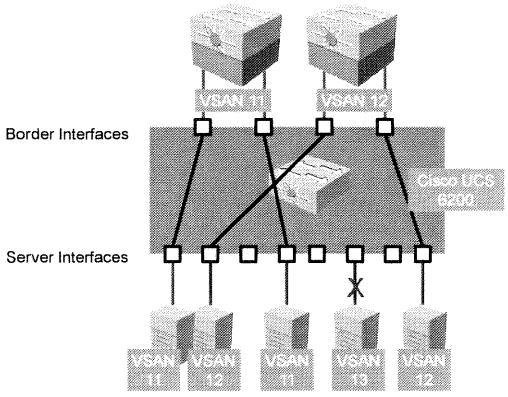
The Cisco UCS 6200XP switch in NPV edge mode no longer processes FLOGI login requests or makes routing decisions using Fabric Shortest Path First (FSPF). Instead, these operations are passed to the upstream switch known as the NPV core switch. The NPV core switch utilizes NPIV in order to interpret multiple logins from the same port.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Pinning with VSANs

- Pinning based on VSANs
  - Server interface pinned to border interface with same VSAN
  - Server interface kept down if no interface with same VSAN available



With VSANs, the NPV edge mode pinning takes into account uplink port VSAN as well. The server is pinned to an uplink port based on the uplink interface VSAN membership (still in a round-robin fashion).

SAN Pinning

- To identify a specific Fibre Channel uplink port on one/both fabric interconnects
- Administrative pinning—using pin groups
  - Pin traffic from specific blade SAN adapter to specific uplink using service profile
- Dynamic pinning (default)—not using pin groups
  - Cisco UCS automatically chooses uplink for SAN adapter based on VSAN

If required, the SAN pinning can be administratively defined with pin groups. The configuration is applied via a service profile.

https://t.me/learningnets

SAN Adapter Failover

- No hardware-assisted fabric failover
  - Traditional SAN host multipathing must be used for fabric failover
- Fibre Channel uplink failure
  - Other uplinks in same VSAN available on the same fabric interconnect:
    - Interface goes down/up to trigger server FLOGI
    - Dynamic repinning to the uplink on same fabric interconnect
    - If failed uplink comes back, no repinning to avoid disruption
  - No uplinks available on the same fabric interconnect
    - Server port brought down

Within the Cisco UCS, SAN-level fabric failover is not available. Thus the traditional SAN host-based multipathing must be used to achieve that level of high availability.

If there are multiple Fibre Channel uplinks on the same fabric interconnect for the same VSAN, the following occurs:

- Interface goes down and up to trigger server FLOGI
- Dynamic repinning to the uplink of the same fabric interconnect in the same VSAN occurs
- Upon the failed uplink restoration, repinning is avoided in order to prevent disruption

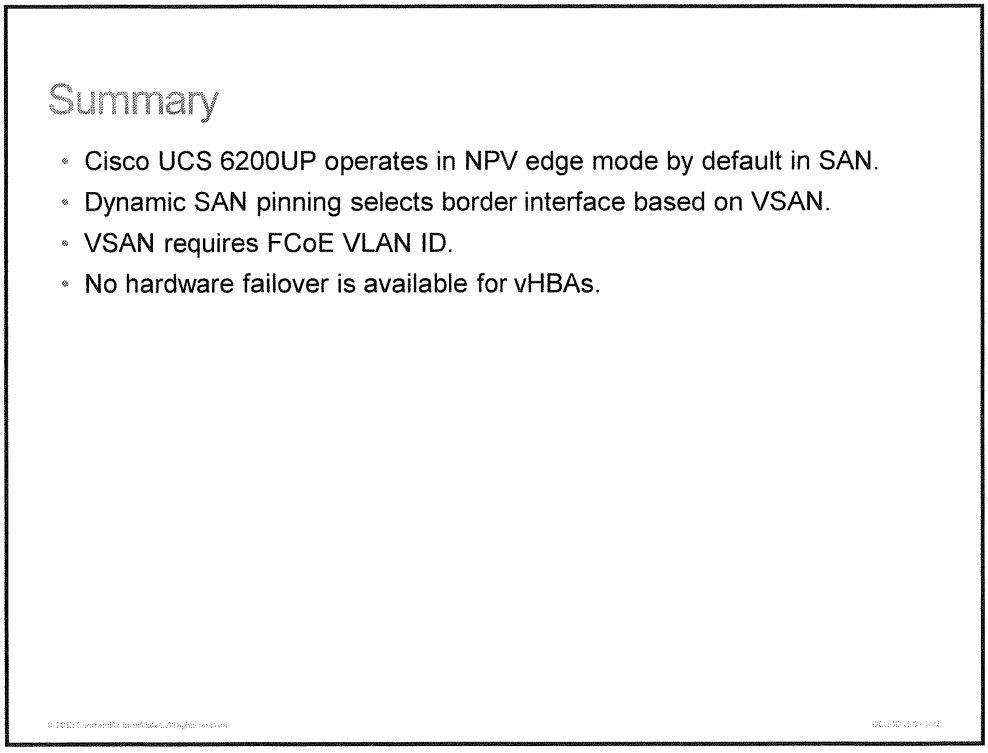
If there are no Fibre Channel uplinks in the same fabric interconnect for the same VSAN, the server port is brought down.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Summary

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



https://t.me/learningnets

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Lesson 3

Designing the Virtual Access Layer

Overview

This lesson discusses the Cisco Nexus 1000V Series Switch and its placement in the virtual network access layer of the Cisco Unified Computing System (UCS) solution.

Objectives

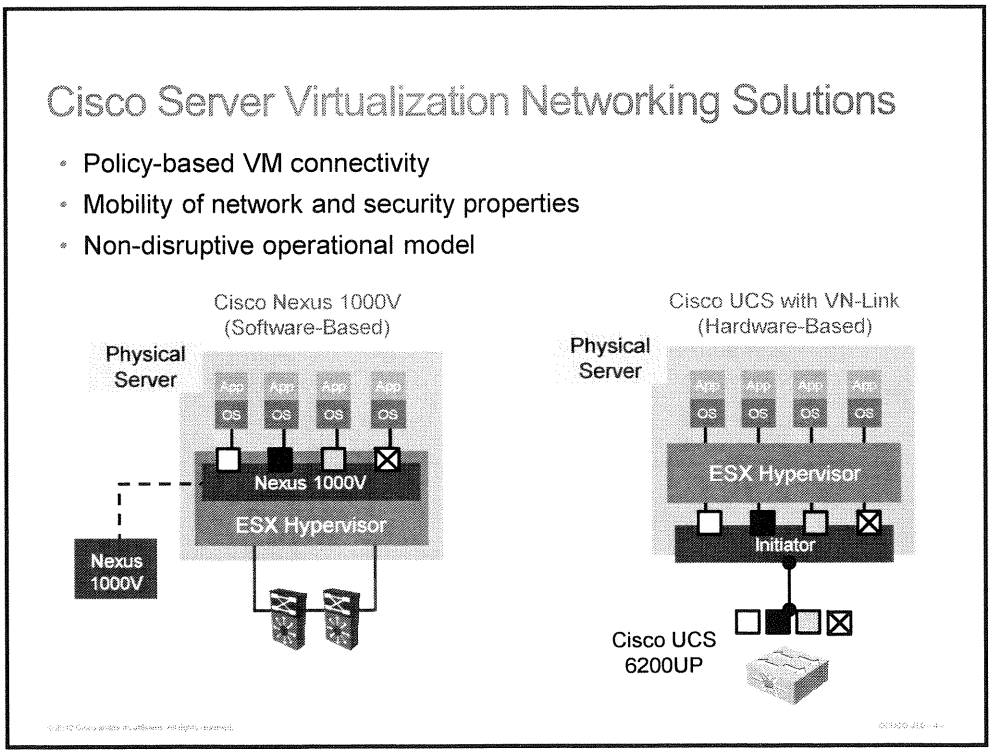
Upon completing this lesson, you will be able to recognize Cisco Nexus 1000V and aspects of Nexus 1000V design. You will be able to meet these objectives:

- Identify and describe the Cisco Nexus 1000V
- Identify and describe Cisco Nexus 1000V integration with VMware vCenter

https://t.me/learningnets

Cisco Nexus 1000V Overview

This topic identifies and describes Cisco Nexus 1000V.



Cisco server virtualization solution uses technology jointly developed by Cisco and VMware. The network access layer is moved into the virtual environment to provide enhanced network functionality at the virtual machine (VM) level.

This can be deployed as a hardware- or software-based solution, depending on the data center design and demands. Both deployment scenarios offer VM visibility, policy-based VM connectivity, policy mobility, and a nondisruptive operational model.

Cisco Nexus 1000V

The Cisco Nexus 1000V is a software-based solution providing VM-level network configurability and management. The Cisco Nexus 1000V works with any upstream switching system to provide standard networking controls to the virtual environment.

VM-FEX

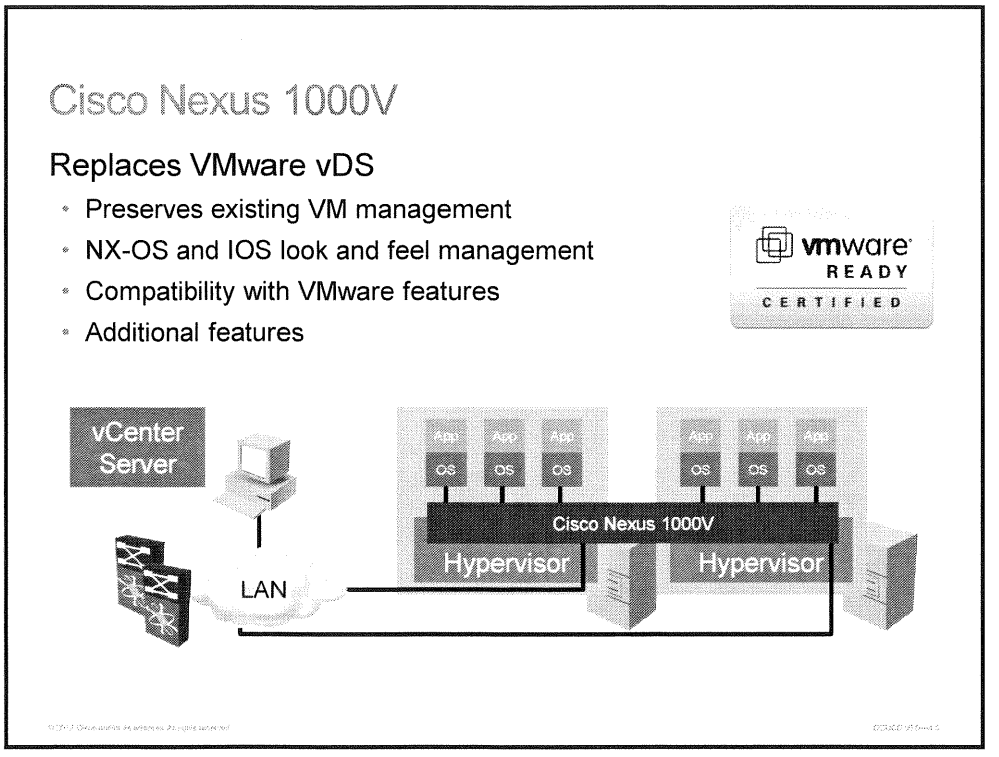
Virtual Machine Fabric Extender (VM-FEX) is designed to move the network access layer into the virtual environment in order to provide enhanced network functionality at the VM level.

Cisco UCS 6200UP

With the Cisco UCS 6200UP, VM-FEX can be deployed as a hardware-based solution offering VM visibility, policy-based VM connectivity, policy mobility, and a nondisruptive operational model.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets



Cisco Nexus 1000V Series Switches are VM access switches that are an intelligent software switch implementation for VMware vSphere environments running the Cisco NX-OS operating system. Operating inside the VMware ESX hypervisor, the Cisco Nexus 1000V Series supports Cisco VM-FEX server virtualization technology to provide the following:

- Policy-based VM connectivity
- Mobile VM security and network policy
- Nondisruptive operational model for your server virtualization and networking teams

The Cisco Nexus 1000V bypasses the VMware vSwitch with a Cisco software switch. This model provides a single point of configuration for the networking environment of multiple ESX hosts.

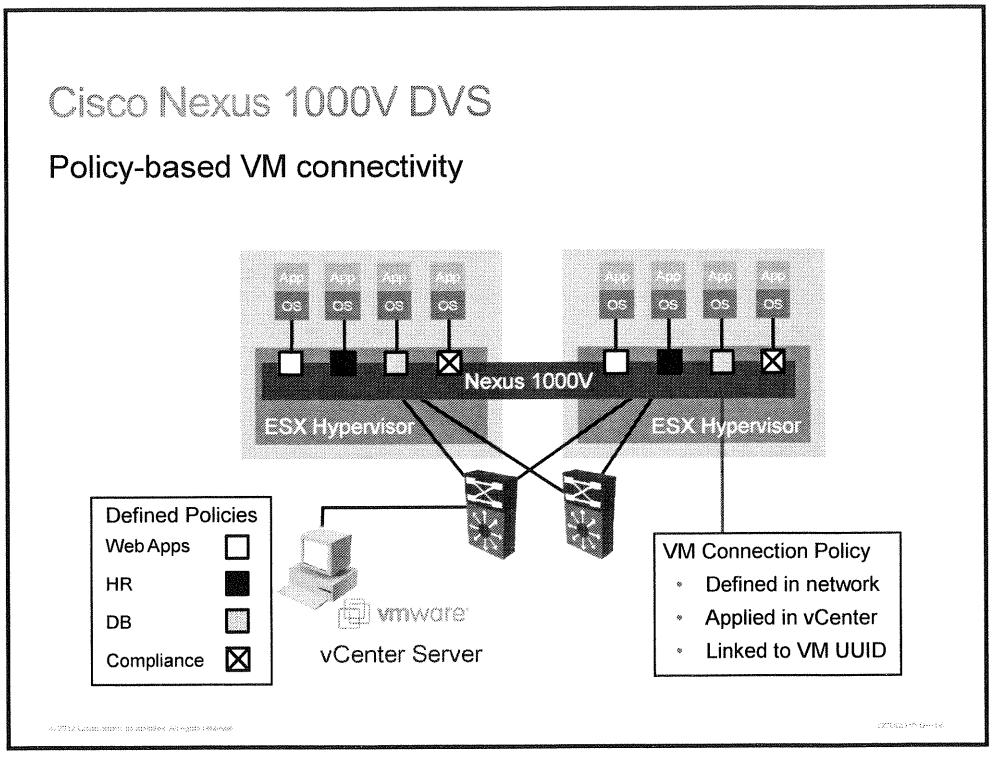
When server virtualization is deployed in the data center, virtual servers typically are not managed the same way as physical servers. Server virtualization is treated as a special deployment, leading to longer deployment time, with a greater degree of coordination among server, network, storage, and security administrators.

With the Cisco Nexus 1000V Series, you can have a consistent networking feature set and provisioning process all the way from the VM access layer to the core of the data center network infrastructure. Virtual servers can now leverage the same network configuration, security policy, diagnostic tools, and operational models as can their physical server counterparts attached to dedicated physical network ports.

Virtualization administrators can access predefined network policy that follows mobile virtual machines to ensure proper connectivity, thereby saving valuable time to focus on VM administration.

This comprehensive set of capabilities helps deploy server virtualization faster and realize its benefits sooner.

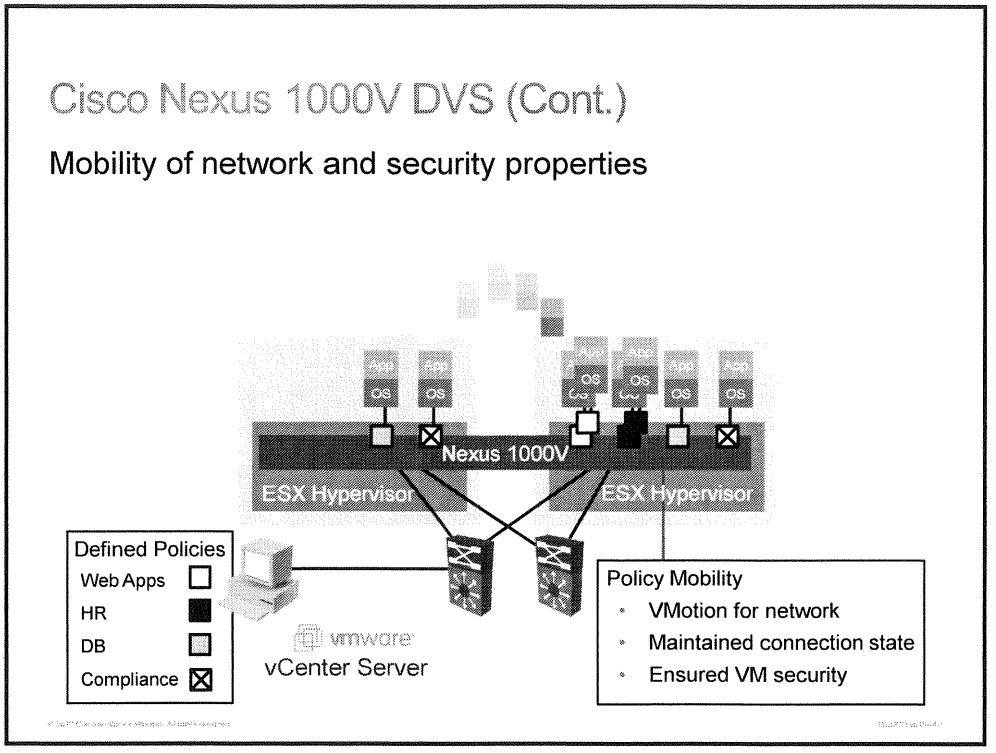
https://t.me/learningnets



VM connection policies are defined in the network and applied to individual VMs from within VMware vCenter. These policies are linked to the Universally Unique Identifier (UUID) of the VM and are not based on physical or virtual ports.

To complement the ease of creating and provisioning VMs, the Cisco Nexus 1000V includes the Port Profile feature to address the dynamic nature of server virtualization from the network perspective. Port Profiles enable you to define VM network policies for different types or classes of VMs from the Cisco Nexus 1000V Virtual Supervisor Module (VSM), and then apply the profiles to individual VM virtual network interface cards (vNICs) through the VMware vCenter GUI for transparent provisioning of network resources. Port Profiles are a scalable mechanism to configure networks with a large numbers of VMs.

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



Through the VMware vCenter application programming interfaces (APIs), the Cisco Nexus 1000V monitors VM migration and ensures policy enforcement as machines transition between physical ports. Security policies are applied and enforced as VMs migrate through automatic or manual processes.

Network and security policies defined in the Port Profile follow the VM throughout its lifecycle, whether it is being migrated from one server to another, suspended, hibernated, or restarted.

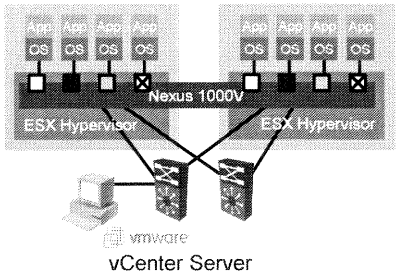
In addition to migrating the policy, the Cisco Nexus 1000V VSM also moves the VM network state, such as the port counters and flow statistics. VMs participating in traffic-monitoring activities, such as Cisco NetFlow or Encapsulated Remote Switched Port Analyzer (ERSPAN), can continue these activities uninterrupted by VMotion operations.

When a specific Port Profile is updated, the Cisco Nexus 1000V automatically provides live updates to all of the virtual ports using that same Port Profile. With the ability to migrate network and security policies through VMotion, regulatory compliance is much easier to enforce with the Cisco Nexus 1000V because the security policy is defined in the same way as physical servers and constantly enforced by the switch.

https://t.me/learningnets

Cisco Nexus 1000V DVS (Cont.)

- Nondisruptive operational model
- Server benefits
  - Existing VM management preserved
  - Reduced deployment time and operational workload
  - Improved scalability
  - VM-level visibility
- Network benefits
  - Unified network management and operations
  - Improved operational security
  - Enhanced VM network features
  - Policy persistence
  - VM-level visibility



When using the Cisco Nexus 1000V, the management model for VMs stays the same and is processed by the VM administrator. Network administrators create security profiles, and these policies are applied to individual VMs by the VM admin. Deployment time is reduced through preconfigured repeatable processes, thereby reducing the operational workload. The benefits include unified network management and operations, enhanced network features at the VM level, and VM-level visibility.

Because of its close integration with VMware vCenter Server, the Cisco Nexus 1000V allows virtualization administrators to continue using VMware tools to provision VMs. At the same time, network administrators can provision and operate the VM network the same way they do the physical network—using Cisco command-line interface (CLI) and Simple Network Managing Protocol (SNMP) along with tools such as ERSPAN and NetFlow.

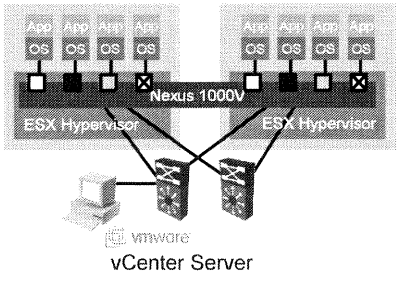
While both teams work independently, using familiar tools, the Cisco Nexus 1000V enforces consistent configuration and policy throughout the server virtualization environment. This level of integration lowers the cost of ownership while supporting various organizational boundaries among server, network, security, and storage teams.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Cisco Nexus 1000V Features

- Layer 2
  - VLAN, PVLAN, 802.1Q
  - LACP
  - vPC host mode
- QoS classification and marking
- Security
  - Layer 2, 3, 4 access lists
  - Port security
- SPAN and ERSPAN
- Compatibility with VMware
  - VMotion, Storage VMotion
  - DRS, HA, FT



Cisco Nexus 1000V supports the same features as a physical Cisco Catalyst or Nexus switches while maintaining compatibility with VMware advanced services like VMotion, Distributed Resource Scheduler (DRS), Fault Tolerance, High Availability, Storage VMotion, Update Manager, and vShield Zones.

vPC Host Mode

Virtual PortChannel host mode (vPC-HM) allows member ports in a port channel to connect to two different upstream switches. With vPC-HM, ports are grouped into two subgroups for traffic separation. If Cisco Discovery Protocol is enabled on the upstream switch, then the subgroups are automatically created using Cisco Discovery Protocol information. If Cisco Discovery Protocol is not enabled on the upstream switch, then the subgroup on the interface must be manually configured.

Layer 2 Features

The following Layer 2 features are supported by the Cisco Nexus 1000V:

- Layer 2 switch ports and VLAN trunks
- IEEE 802.1Q VLAN encapsulation
- Link Aggregation Control Protocol (LACP): IEEE 802.3ad
- Advanced port channel hashing based on Layer 2, 3, and 4 information
- Virtual PortChannel host mode
- Private VLANs with promiscuous, isolated, and community ports
- Private VLAN on trunks
- Internet Group Management Protocol (IGMP) snooping versions 1, 2, and 3
- Jumbo frame support of up to 9216 bytes
- Integrated loop prevention with bridge protocol data unit (BDPU) filter without running Spanning Tree Protocol (STP)

https://t.me/learningnets

QoS Features

The following quality of service (QoS) features are supported by the Cisco Nexus 1000V:

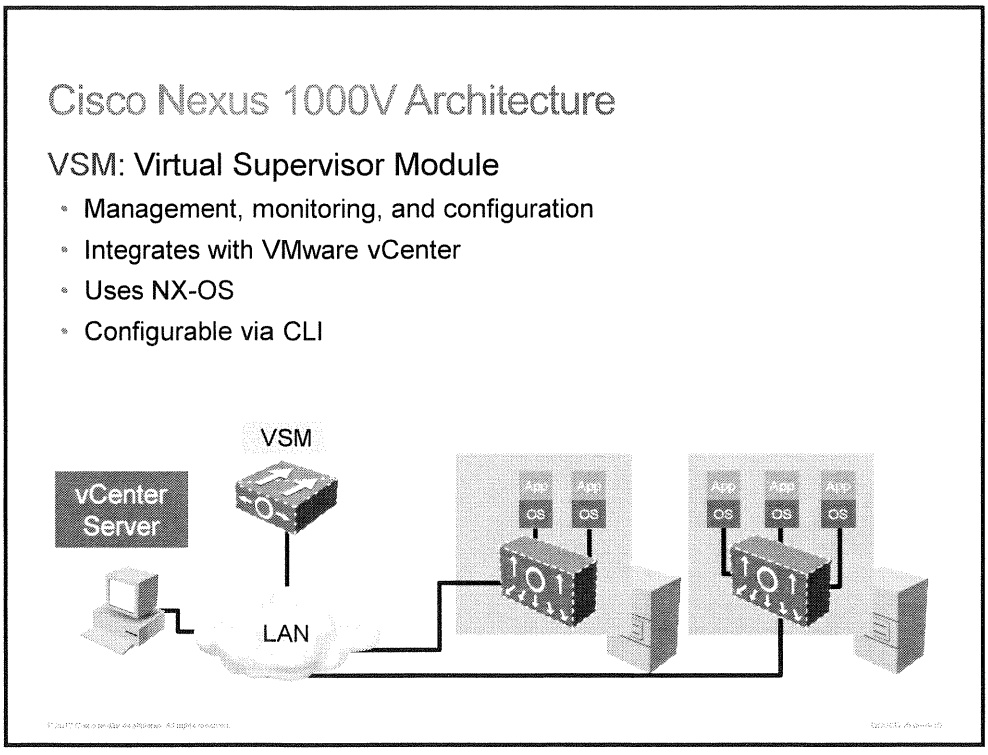
- Classification per access group (access control list [ACL]), IEEE 802.1p class of service (CoS), IP Type of Service: IP precedence or differentiated services code point (DSCP) (RFC 2474), User Datagram Protocol (UDP) ports, packet length
- Marking per two-rate, three-color marker (RFC 2698), IEEE 802.1p CoS marking, IP Type of Service: IP precedence or DSCP (RFC 2474)
- Traffic policing (transmit- and receive-rate limiting)
- Modular QoS CLI Modular QoS CLI (MQC) compliance

Security Features

The following security features are supported by the Cisco Nexus 1000V:

- Ingress and egress access control lists (ACLs) on Ethernet and virtual Ethernet ports
- Standard and extended Layer 2 ACLs
- Standard and extended Layer 3 and 4 ACLs
- Port ACLs (PACLs)
- Named ACLs
- ACL statistics
- Cisco integrated security features
- Virtual service domain for Layer 4 through 7 service VM

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



Cisco Nexus 1000V is licensed per each server CPU regardless of the number of cores. It comprises the following:

- **Virtual Supervisor Module (VSM):** Performs management, monitoring, and configuration tasks for the Cisco Nexus 1000V and is tightly integrated with the VMware vCenter—the connectivity definitions are pushed from Cisco Nexus 1000V to the vCenter.
- **Virtual Ethernet Module (VEM):** Enables advanced networking capability on the ESX hypervisor and provides each VM with a virtual dedicated switch port.

A Cisco Nexus 1000V deployment consists of VSM (one or two for redundancy) and multiple VEMs installed in the ESX hosts—a vNetwork Distributed Switch.

A VSM is a supervisor module similar to regular physical modular switches, while VEMs are remote Ethernet line cards that are connected to a VSM.

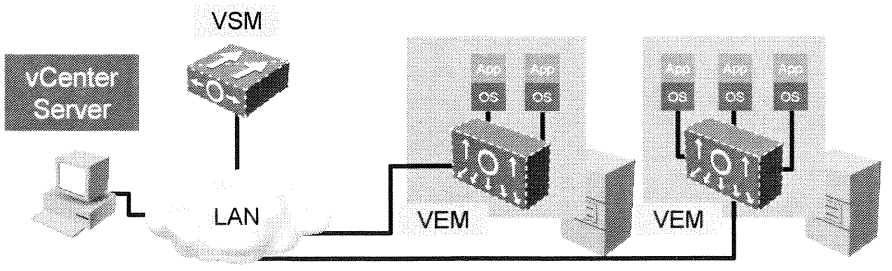
In Cisco Nexus 1000V deployments, VMware provides the vNIC and drivers, while the Cisco Nexus 1000V provides the switching and management of switching.

https://t.me/learningnets

Cisco Nexus 1000V Architecture (Cont.)

VEM—Virtual Ethernet Module

- Replaces ESX virtual switch
- Enables advanced networking on ESX hypervisor
- Provides each VM with dedicated port



The VEM is a software replacement for the VMware vSwitch on a VMware ESX Host. All traffic-forwarding decisions are made by the VEM.

The VEM leverages the VMware vNetwork Distributed Switch (vDS) API, which was developed jointly by Cisco and VMware to provide advanced networking capability to virtual machines. This level of integration ensures that the Cisco Nexus 1000V is fully aware of all server virtualization events, such as VMware VMotion and DRS. The VEM takes configuration information from the VSM and performs Layer 2 switching and advanced networking functions:

- **PortChannel**
- **QoS**
- **Security:** Private VLAN, access control lists, port security
- **Monitoring:** NetFlow, Switched Port Analyzer (SPAN), ERSPAN

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

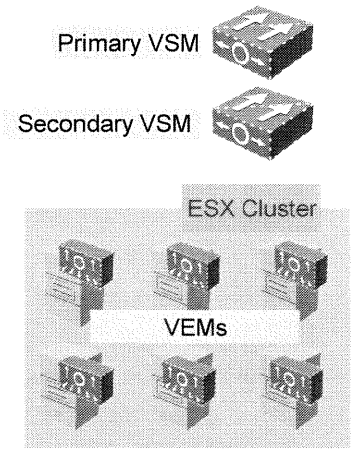
Cisco Nexus 1000V Architecture (Cont.)

ESX cluster = single Cisco Nexus 1000V switch

- Two VSM for high availability
- One VEM per ESX host

License per CPU socket for each VEM

- Installed on the VSM
- One or more license files can be installed
- Installation is nondisruptive to operation
- Dual supervisor environment runs licensed software on both



The Cisco Nexus 1000V uses distributed architecture. This architecture separates control and data plane functionality. The control plane functionality is represented by VSM, which manages multiple distributed data planes (VEM in each ESX host). Thus, a VSM acts as a supervisor module for remote VEMs.

All configuration and supervisor functions are processed by the VSM. Using Console, Telnet, or Secure Shell (SSH), an administrator makes all configuration changes on the VSM. When a change is made on the VSM, the configuration is passed to vCenter, and the changes are then made on the distributed virtual switch (DVS). DVS changes are then passed down to the corresponding VEM.

Each VEM will act as a module on the VSM and the VSM or VEM appears as a single switch to Cisco Discovery Protocol neighbors. The VSM does not reside in the data path and therefore cannot directly receive or respond to Cisco Discovery Protocol messages. Cisco Discovery Protocol and other network management packets are transferred between the VEM to the VSM on one of three required VLANs, known as the Packet VLANs.

The Cisco Nexus 1000V Series is licensed based on the number of physical CPUs on the server on which the VEM is running.

https://t.me/learningnets

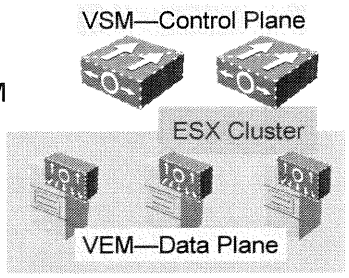
VSM: Control Plane

Centralized control plane—VSM

- Manages multiple data planes
- Software appliance on physical server or VM
- Two VSMs for high availability

Maximum configuration:

- 64 VMware hosts per VSM
- 512 active VLANs
- 2048 vEth ports per vDS
- 216 vEth per host
- 32 physical NICs per host
- 256 port channels per vDS
- 8 port channels per host



VSM VM	Minimum Req.
OS	Other Linux (64-bit)
Memory	2 GB
vCPU	1
Network	3x e1000
Disk	3 GB

VSM can be deployed as a software appliance on a physical server (Control Plane Physical Appliance [CPPA]) or on a VM (Control Plane Virtual Appliance [CPVA]). A redundant Cisco Nexus 1000V deployment would incorporate two VSM appliances.

An individual Cisco Nexus 1000V supports the following:

- One data plane per ESX host
- Up to 64 ESX hosts
- 512 active VLANs
- 32 physical NICs per ESX host
- 2048 virtual Ethernet ports
- 216 virtual Ethernet ports per ESX host

The VSM has the following minimum system requirements when run as a VM:

- 2-GB memory
- Single virtual CPU (vCPU)
- Three e1000 network adapters named Management, Control, and Packet in the right sequence
- 4-GB disk

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

VEM: Data Plane

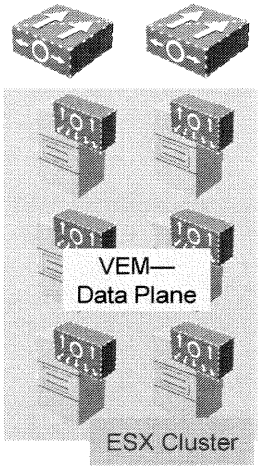
Distributed data plane—VEMs

- Each operates independently
- No address learning across VEM
- No backplane between VEM
- No forwarding from ingress line card to egress line card
- No EtherChannel across VEMs

Internal switching—VSM not required

- VSM failure does not fail data path
- Traffic continues to flow by VEM

No STP



Each VEM acts as a separate switching line card, with no concept of a fabric between VEMs. This means there are no port channels or connections between VEMs, and they do not rely on one another for operation. Switching decisions and frame forwarding all happen on the VEM and are not reliant on the VSM.

Only the uplinks in a host can be bundled in a port channel for load balancing and high availability. The Cisco Nexus 1000V does not support EtherChannels across different VEMs.

The Cisco Nexus 1000V does not run STP because it will deactivate all but one uplink to an upstream switch, thereby preventing full utilization of uplink bandwidth. Instead, each VEM is designed to prevent loops in the network topology.

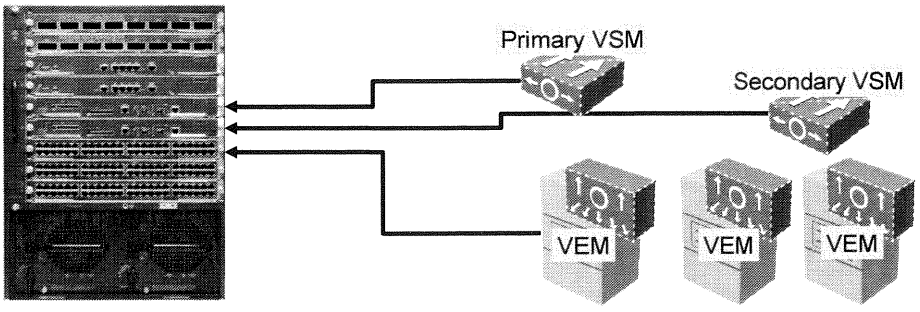
In the event of loss of communication with the VSM, the VEM has Nonstop Forwarding capability to continue to switch traffic based on the last known configuration. In short, the VEM provides advanced switching with data-center reliability for the server virtualization environment.

Unlike with the VSM, VEM resources are unmanaged and dynamic. Although the storage footprint of the VEM is fixed (approximately 6.4 MB of disk space), RAM utilization on the VMware ESX host is variable based on the configuration and scale of the Cisco Nexus 1000V Series deployment. In a typical configuration, each VEM can be expected to require 10 to 50 MB of RAM, with an upper hard limit of 150 MB for a fully scaled solution with all features turned on and utilized to their design limits.

https://t.me/learningnets

Cisco Nexus 1000V Virtual Chassis

```
n1kv# show module
Mod  Ports  Module-Type                      Model              Status
---  -----  -------------------------------- ------------------ ------------
1    0      Virtual Supervisor Module        Nexus1000V         active *
2    0      Virtual Supervisor Module        Nexus1000V         ha-standby
3    248    Virtual Ethernet Module          NA                 ok
4    248    Virtual Ethernet Module          NA                 ok
5    248    Virtual Ethernet Module          NA                 ok
... output omitted ...
```



The Cisco Nexus 1000V Series uses a virtual chassis model to represent a pair of VSMs and their associated VEMs.

Like any Cisco chassis base platform, the Cisco Nexus 1000V Series virtual chassis has slots and modules, or line cards, associated with it. The VSMs are always associated with slot numbers 1 and 2 in the virtual chassis.

The VEMs are sequentially assigned to slots 3 through 66 based on the order in which their respective hosts were added to the Cisco Nexus 1000V Series Switch. With the 64 VEMs and the redundant supervisors, the Cisco Nexus 1000V can be viewed as a 66-slot modular switch.

When a VEM comes online for the first time, the VSM assigns the module number and tracks that module using the UUID of the VMware ESX server, helping ensure that if the VMware ESX host loses connectivity or is powered down for any reason, the VEM will retain its module number when the host comes back online.

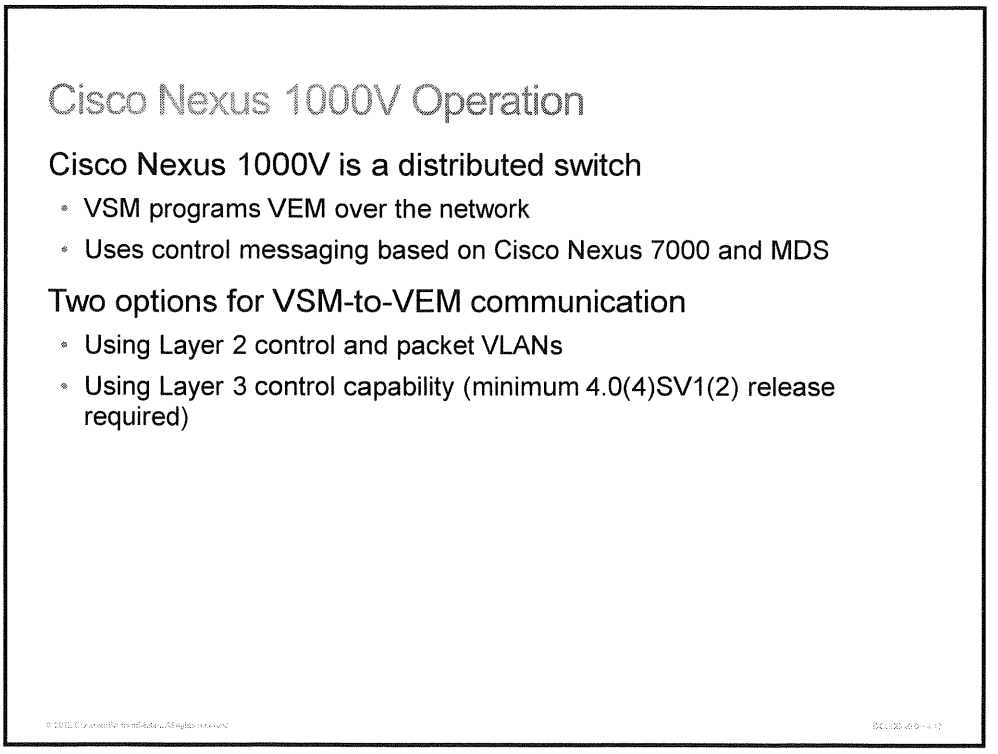
The VSM maintains a "heartbeat" with its associated VEMs transmitted at 2-second intervals. If the VSM does not receive a response within 8 seconds, the VSM considers the VEM removed from the virtual chassis. If the VEM is not responding because of a connectivity problem, the VEM will continue to switch packets in its last known good state. When communication is restored between a running VEM and the VSM, the VEM is reprogrammed, causing a slight (1 to 15 seconds) pause in network traffic.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Cisco Nexus 1000V VMware Integration

This topic identifies and describes Cisco Nexus 1000V integration with VMware vCenter.



In many ways, the Cisco Nexus 1000V Series Switches are similar to physical Ethernet switches. For packet forwarding, the Cisco Nexus 1000V Series uses the same techniques that other Ethernet switches apply, with a MAC address-to-port mapping table used to determine where packets should be forwarded.

The Cisco Nexus 1000V Series maintains forwarding tables in a slightly different manner than other modular switches. Unlike physical switches, which have a centralized forwarding engine, each VEM maintains a separate forwarding table. There is no synchronization between forwarding tables on different VEMs. In addition, there is no concept of forwarding from a port on one VEM to a port on another VEM. Packets destined for a device not local to a VEM are forwarded to the external network, which in turn may forward the packets to a different VEM.

MAC Address Learning

This distributed forwarding model within a centrally managed switch is demonstrated by the way the Cisco Nexus 1000V Series manages MAC address learning. A MAC address can be learned multiple times within a single Cisco Nexus 1000V Series Switch in either of two ways: statically or dynamically. Static entries are automatically generated for virtual machines running on the VEM—these entries do not time out. For devices not running on the VEM, the VEM can learn a MAC address dynamically through the physical NICs in the server.

Each VEM maintains a separate MAC address table. Thus, a single Cisco Nexus 1000V Series Switch may learn a given MAC address multiple times, possibly as often as once per VEM. For example, one VEM may be hosting a virtual machine, and the MAC address of the virtual machine will be statically learned on the VEM. A second VEM, in the same Cisco Nexus 1000V Series Switch, may learn the MAC address of the virtual machine dynamically. Thus, within the Cisco NX-OS CLI, you may see the MAC address of the virtual machine twice: a dynamic entry and a static entry.

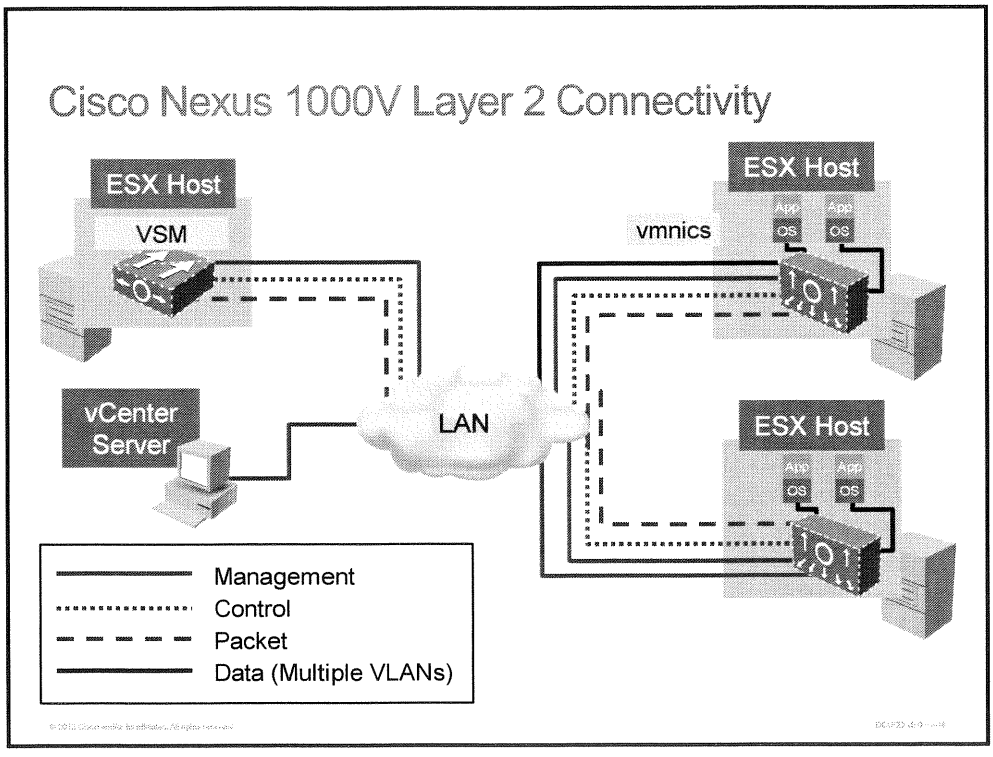
https://t.me/learningnets

Loop Prevention

Another differentiating characteristic of the Cisco Nexus 1000V Series is that it does not run STP. Although this might seem to be a significant departure from other Ethernet switches, potentially causing catastrophic network loops, the Cisco Nexus 1000V Series implements a simple and effective loop-prevention strategy that does not require STP.

Because the Cisco Nexus 1000V Series does not participate in STP, it does not respond to bridge protocol data unit (BPDU) packets, nor does it generate them. BPDU packets that are received by Cisco Nexus 1000V Series Switches are dropped.

The Cisco Nexus 1000V Series uses a simple technique to prevent loops. Similar to a physical Ethernet switch, the Cisco Nexus 1000V Series performs source and destination MAC address lookups in order to make forwarding decisions. The VEM applies loop-prevention logic to every incoming packet on Ethernet interfaces. This logic is used to identify potential loops. Every ingress packet on a physical Ethernet interface is inspected to help ensure that the destination MAC address is internal to the VEM. If the destination MAC address is external, the Cisco Nexus 1000V Series will drop the packet, thus preventing a loopback to the physical network.



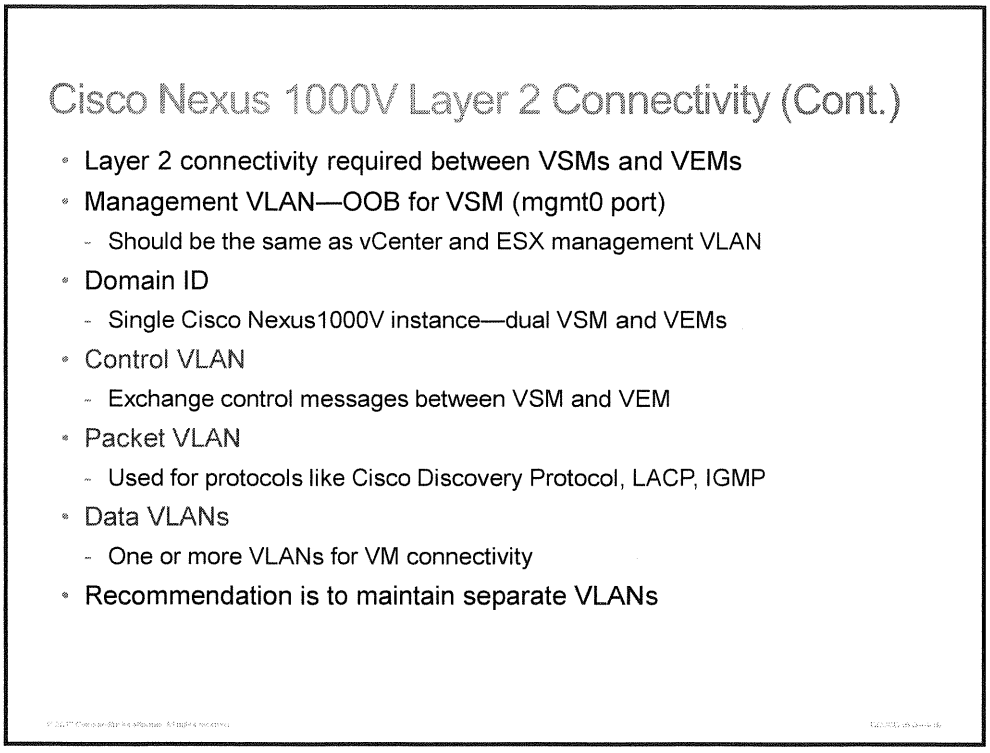
The VSM, VEM, vCenter, and VM connectivity use dedicated VLANs—management, control, packet, and one or more data networks.

Like the VSM, each VEM has a control and a packet interface. These interfaces are unmanaged and are not directly configurable by the end user. The VEM uses the opaque data provided by VMware vCenter Server to configure the control and packet interfaces with the correct VLANs. The VEM then applies the correct uplink port profile to the control and packet interfaces in order to establish communication with the VSM. There are two ways of connecting the VSM and the VEM.

If the VSM and the VEM are in the same Layer 2 domain, the best way to connect them is to use the Layer 2 connectivity mode.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets



The Cisco Nexus 1000V VLANs consist of a management, control, packet, and one or more data networks.

Management VLAN

Each VMware ESX Host, the VSM, and the vCenter Server must all reside in the same management network and be part of the same Layer 2 domain.

Domain ID

A single Cisco Nexus 1000V instance, including dual redundant VSMs and managed VEMs, forms a switch domain. Each Cisco Nexus 1000V domain within a VMware vCenter Server needs to be distinguished by a unique integer called the domain identifier.

Control and Packet VLANs

Control and packet VLANs from the VSM must be accessible by uplink profiles on each VEM. The control VLAN and the packet VLAN are used for communication between the VSM and the VEMs within a switch domain.

The packet VLAN is used by protocols such as Cisco Discovery Protocol, Link Aggregation Control Protocol (LACP), and Internet Group Management Protocol (IGMP).

The control VLAN is used for the following:

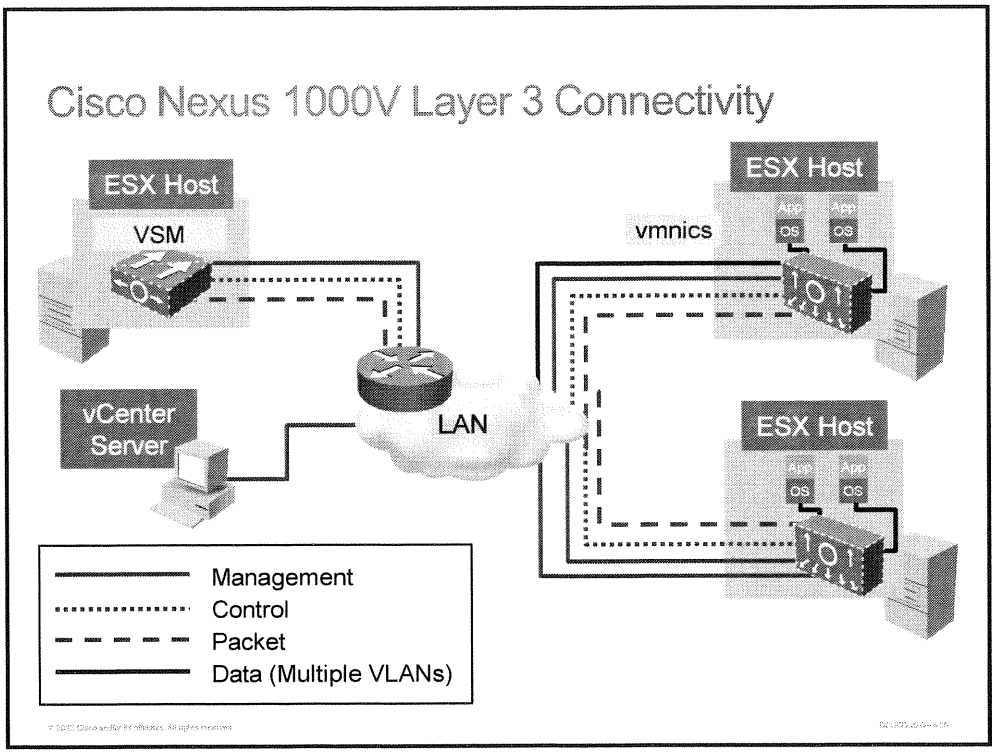
- VSM configuration commands to each VEM, and their responses
- VEM notifications to the VSM—for example, a VEM notifies the VSM of the attachment or detachment of ports to the DVS
- VEM NetFlow exports are sent to the VSM, from where they are then forwarded to a NetFlow Collector

https://t.me/learningnets

Data VLANs

The data networks carry VM packet traffic (server data). One or more data VLANs are defined for this purpose. Data traffic from the VM is not sent to the VSM, and the VSM does not require access to the data VLANs. All VSM management is out-of-band, and switching decisions do not rely on the VSM.

Note It is recommended that the control VLAN and packet VLAN should be separate VLANs, and they should be on separate VLANs from those that carry data.



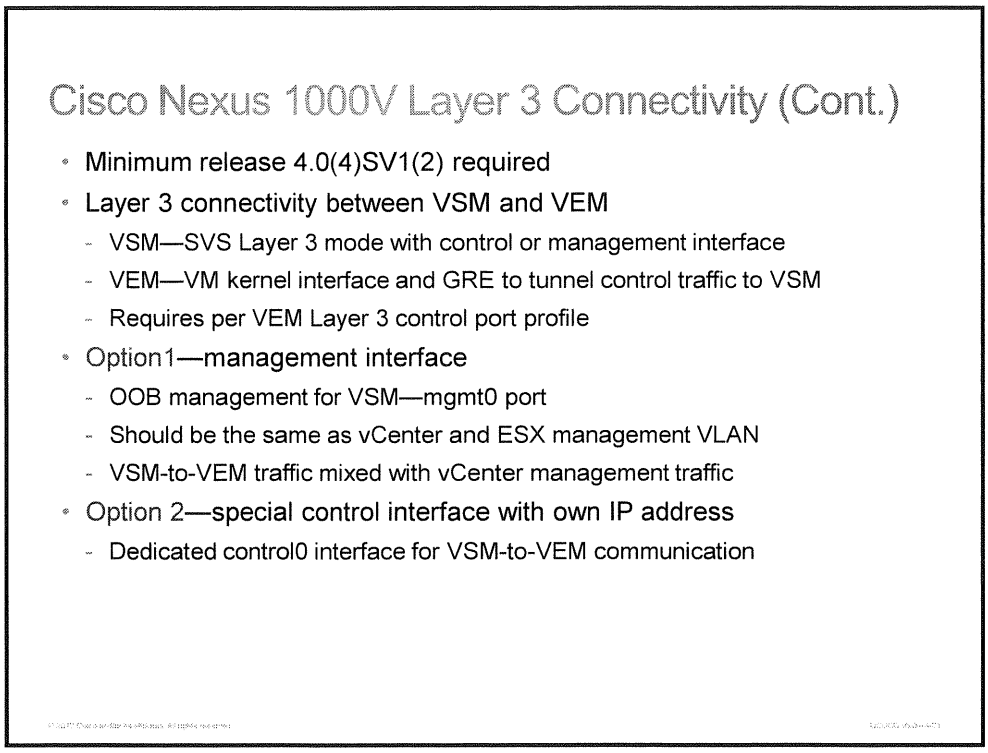
If the VSM and the VEM are in different Layer 2 domains, the Layer 3 connectivity mode should be used.

The Layer 3 mode encapsulates the packet of the Layer 2 mode using Generic Routing Encapsulation (GRE).

This process requires configuration of a VMware VMkernel interface on each VMware ESX host. This VMware VMkernel interface will need to be configured and attached to a port profile with the l3control option.

The l3control configuration tells the VEM that it can then use this interface to send Layer 3 packets. So even if the Cisco Nexus 1000V Series is a Layer 2 switch, it can still send IP packets.

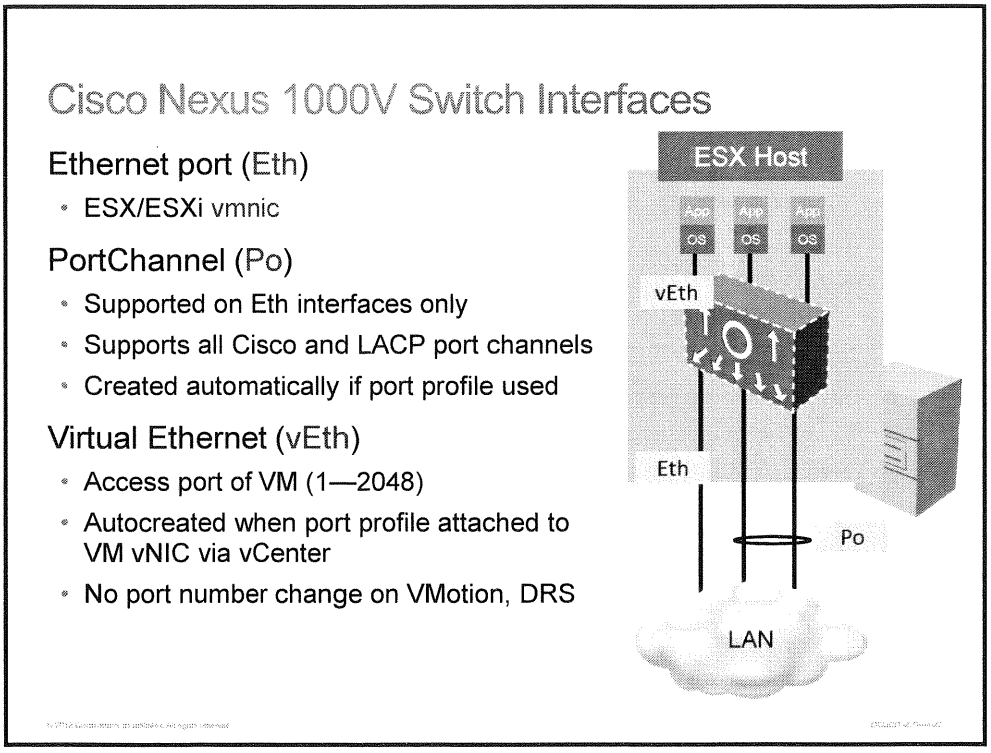
© 2012 Cisco Systems, Inc.
https://t.me/learningnets



For Layer 3 connectivity between VSM and VEM, the following options are available:

- Out-of-band (OOB) management interface (mgmt0 port), which must be in the same VLAN as vCenter Server and ESX management interfaces. With this option, the VSM-to-VEM traffic is mixed with vCenter management traffic.
- Dedicated control interface (Control0) with its own IP address for VSM-to-VEM communication

https://t.me/learningnets



The Cisco Nexus 1000V Series supports multiple switch-port types for internal and external connectivity: virtual Ethernet (vEth), Ethernet (Eth), and PortChannel. The most common port type within a Cisco Nexus 1000V Series environment is a new concept called a virtual Ethernet (vEth) interface. This interface type represents either the switch port connected to a virtual machine vNIC or connectivity to specialized interface types such as the vswif or vmknic interface.

vEth Interface

A vEth interface has several characteristics that differentiate it from other interface types. In addition to the fact that vEth interfaces are virtual and therefore have no associated physical component, the interface naming convention is unique.

Unlike a traditional Cisco interface, a vEth interface name does not indicate the module with which the port is associated. Whereas a traditional physical switch port may be notated as GigX/Y—where X is the module number and Y is the port number on the module—a vEth interface is notated as such: vEthY. This unique notation is designed to work transparently with VMware VMotion, keeping the interface name the same regardless of the location of the associated VM.

The second characteristic that makes a vEth interface unique is its transient nature. A given vEth interface appears or disappears based on the status of the VM connected to it. The mapping of a virtual machine vNIC to a vEth interface is static. When a new VM is created, a vEth interface is also created for each vNIC of the VM. The vEth interfaces will persist as long as the VM exists. If the VM is temporarily down (the guest operating system is shut down), the vEth interfaces will remain inactive but still bound to that specific VM. If the VM is deleted, the vEth interfaces will become available for connection to newly provisioned VMs.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

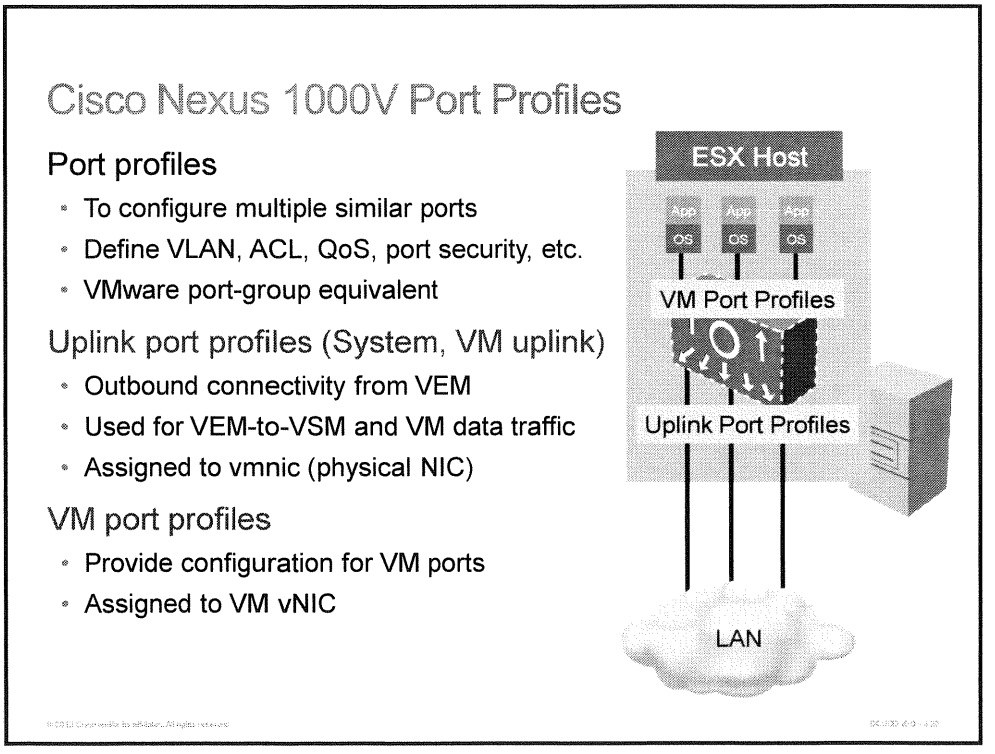
Eth Interface

The Cisco Nexus 1000V Series contains two interface types related to the VMNICs (physical NICs) within a VMware ESX host.

An Ethernet, or Eth, interface is the Cisco Nexus 1000V Series representation of a VMNIC. An Eth interface is represented in standard Cisco interface notation (EthX/Y) using the Cisco NX-OS naming convention "Eth" rather than a speed such as "Gig" or "Fast," as is the custom in Cisco IOS Software. These Eth interfaces are module-specific and are fairly static within the environment.

PortChannel Interface

PortChannel interfaces are the third interface type supported by the Cisco Nexus 1000V Series. A PortChannel interface is an aggregation of multiple Eth interfaces on the same VEM.



In the Cisco Nexus 1000V, port profiles are used to configure interfaces. A port profile can be assigned to multiple interfaces, giving them all the same configuration. Changes to the port profile can be propagated automatically to the configuration of any interface assigned to it.

In the VMware vCenter Server, a port profile is represented as a port group. The vEthernet or Ethernet interfaces are assigned in vCenter Server to a port profile for the following:

- Defining port configuration by policy
- Applying a single policy across many ports
- Supporting both vEthernet and Ethernet ports

https://t.me/learningnets

Port Profile Configuration

A port profile is a set of interface configuration commands that can be dynamically applied to either the physical (uplink) or virtual interfaces. A port profile can define a set of attributes including the following:

- VLAN
- Port channels
- Private VLAN (PVLAN)
- Access control list (ACL)
- Port security
- NetFlow
- Rate limiting
- QoS marking

The network administrator defines port profiles in the VSM. When the VSM connects to vCenter Server, it creates a DVS, and each port profile is published as a port group on the DVS. The server administrator can then apply those port groups to specific uplinks, VM vNICs, or management ports, such as Virtual Switch Interfaces (VSI) or VM kernel NICs.

A change to a VSM port profile is propagated to all ports associated with the port profile. The network administrator uses the Cisco NX-OS CLI to change a specific interface configuration from the port profile configuration applied to it. For example, a specific uplink can be shut down, or a specific virtual port can have ERSPAN applied to it without affecting other interfaces using the same port profile.

Note Although the configuration can be applied to the individual virtual port on Cisco Nexus 1000V, it is recommended to apply the entire configuration via port profiles.

By using port profiles, hundreds or even thousands of VMs can be provisioned rapidly with detailed network configurations such as port state, QoS tagging, and ACL controls. Additionally, port profiles reduce the risk of misconfiguration on groups of similar ports by maintaining the configuration in a central location. While individual port configuration is still possible using the Cisco Nexus 1000V, it is recommended to use port profiles rather than configuring ports directly. This method reduces administration time and simplifies network troubleshooting.

Uplink Port Profiles

The server administrator can assign port profiles that are configured as uplinks to physical NICs.

System Uplink Port Profiles

When a server administrator adds a host to the DVS, the VEM in that host needs to be able to configure the VSM. Because the ports and VLANs for this communication are not yet in place, system port profiles and system VLANs are configured to meet this need. VSM sends minimal early configuration to the vCenter Server, which then propagates it to the VEM when the host is added to the DVS.

A system port profile is designed to establish and protect vCenter Server connectivity. It can carry the following VLANs:

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

- System VLANs or vNICs used when bringing up the ports before communication is established between the VSM and VEM
- The uplink that carries the control VLAN
- Management uplinks used for VMware vCenter Server connectivity or SSH or Telnet connections. There can be more than one management port or VLAN—for example, one dedicated for vCenter Server connectivity, one for SSH, one for SNMP, a switch interface, and so on.
- VMware kernel NIC for accessing VMFS storage over Internet Small Computer Systems Interface (iSCSI) or Network File System (NFS).

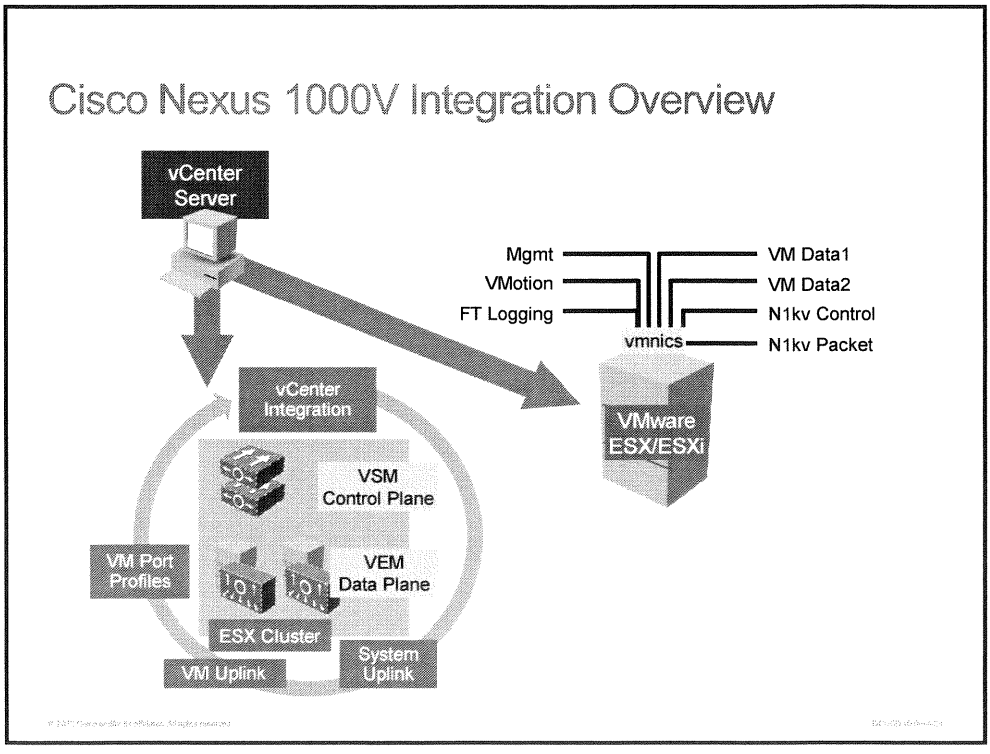
VM Uplink Port Profiles

VM uplink port profiles are used to define VM uplink characteristics that are separate from the control and packet VLANs.

VM Port Profiles

VM profiles are used to provide configuration for VMs and typically require an uplink profile to access the physical network.

When a VM profile is configured without a corresponding uplink profile, it creates internal VM networks. If a VM profile is created by accessing a VLAN that is not trunked to the physical network, then the assigned VMs will be able to communicate only with other VMs assigned to the profile in the same host. This configuration is similar to creating internal-only vSwitches or port groups within a standard VMware networking environment.



The Cisco Nexus 1000V Series is tightly integrated with VMware vCenter. This integration enables the network administrator and the server administrator to collaborate efficiently without each having to learn a different management tool. The network administrator uses the Cisco NX-OS CLI on the VSM, and the server administrator continues to use VMware vCenter. Because of the tight relationship between the VSM and VMware vCenter, the communication between the two needs to be reliable and secure.

https://t.me/learningnets

Cisco Nexus 1000V and VMware vCenter Server Communication

The VSM maintains a link to VMware vCenter Server that is used to maintain the definition of the Cisco Nexus 1000V Series within VMware vCenter Server as well as to propagate port profiles. The server and network administrators both have roles in establishing the link between the Cisco Nexus 1000V Series and VMware vCenter Server.

After installation is finished, the communication between the VSM and VMware vCenter Server is enabled under the svs configuration. The installer application will register the VSM plug-in with VMware vCenter Server, which will then establish the link and create the instance of the Cisco Nexus 1000V Series within VMware vCenter Server.

Each VSM contains a unique extension key used to bind that specific VSM to the VMware vCenter Server.

In creating the Cisco Nexus 1000V Series within VMware vCenter Server, the VSM propagates any port profiles that are already defined, as well as important information required for VEM installation called opaque data. The opaque data provides limited configuration details to the VEM so that it can communicate with the VSM after installation. The VSM is considered the authoritative container for all configuration information. If the connection between the VSM and VMware vCenter Server is disrupted, the VSM helps ensure that any configuration changes that have been made during this period of disrupted communication are propagated to VMware vCenter Server when the link is restored.

After the connection between the VSM and VMware vCenter Server is established, the link is primarily used to propagate new port profiles and any changes to existing port profiles.

Cisco Nexus 1000V Series VMware vCenter Server Extension

VMware vCenter Server is an extensible application that allows third-party management plug-ins, thus enabling external applications to extend the capabilities of VMware vCenter Server and its companion GUI, VMware vSphere Client.

The Cisco Nexus 1000V Series uses a VMware vCenter Server extension to properly display a representation of the Cisco Nexus 1000V Series and its main components within VMware vSphere Client.

The Cisco Nexus 1000V Series extension is a small XML file (cisco_nexus_1000v_extension.xml) that is directly installed and registered using the installation GUI. The plug-in can also be downloaded from the management IP address of the VSM by using a web browser. This plug-in must be installed before the VSM can establish a link to VMware vCenter Server.

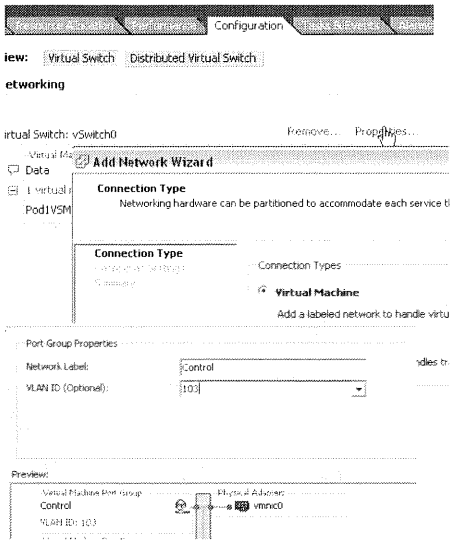
© 2012 Cisco Systems, Inc.
https://t.me/learningnets

Prepare VMware Hosts for Nexus 1000V

Define management, control, packet VLANs

- Required when VSM is running as VM

Create VM port groups for each VLAN on VSM hosts



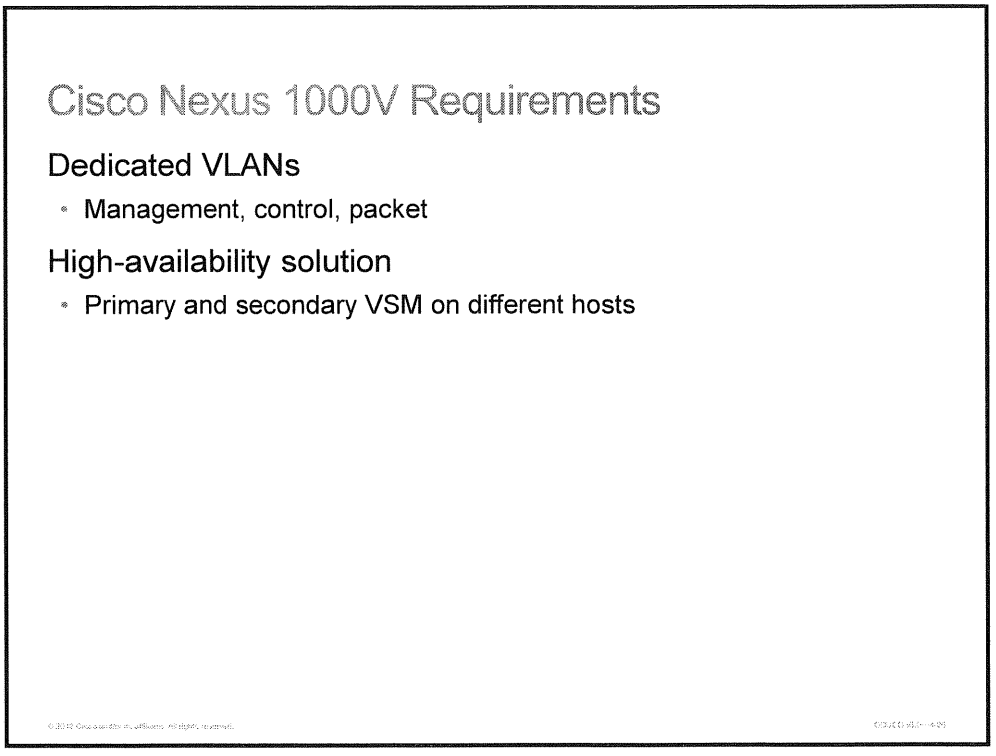
VLAN	ID	Port Group
Mgmt	1	Management
Control	999	Control
Packet	998	Packet

To deploy and integrate Cisco Nexus 1000V with VMware vCenter and hosts, the following prerequisites have to be met:

- vCenter Server must be installed and prepared.
- The VMware Enterprise Plus license must already be installed on the hosts.
- All hosts for the VEM must be running ESX or ESXi 4.0.
- There is at least one ESX or ESXi 4.0 host. If you plan to use VMotion, you need two ESX or ESXi 4.0 hosts.
- An ESX host is available to run the VSM VM.
- The VSM VM can be hosted on the VEM in an ESX host that it is managing, or it can be hosted on a separate ESX or ESXi host (3.5 or 4.0) running the regular VMware vSwitch.
- The ESX host requires a minimum of 4 GB of physical RAM to host a VSM VM as the ESX server alone requires a minimum of 2 GB of physical RAM. Additional memory may be required to run the vCenter Server VM on the same host.
- Each host has a minimum of the following physical NICs (PNICs): one PNIC for a Service Console or management, one PNIC for the traffic between VSM and VEM and for VM data traffic.
- All ESX hosts within a Cisco Nexus 1000V must have Layer 2 connectivity to one another.
- If you are using a set of switches, make sure that the interswitch trunk links carry all relevant VLANs, including control and packet VLANs. The uplink should be a trunk port carrying all VLANs configured on the ESX host.
- On the host running the VSM VM, the control and packet VLANs are configured through the VMware switch and the VMNIC.
- The optional VMware Update Manager (VUM) manages the Cisco Nexus 1000V software installation for ESX hosts in a data center.

https://t.me/learningnets

- The VSM VM control interface must be on the same Layer 2 VLAN as the ESX 4.0 host that it manages. If you configure Layer 3, then you do not have this restriction. In each case, the two VSMs must run in the same IP subnet.
- You have now completed configuring the valid network mapping for control, management, and packet VLAN access.
- The host for the VSM VM runs on 64-bit server hardware and can run ESX 4.0 or ESX 3.5.



Planning the Cisco Nexus 1000V deployment must take into consideration Cisco Nexus 1000V, VMware, and uplink switch aspects.

From the Cisco Nexus 1000V perspective, the following must be addressed:

- Licensing: Cisco Nexus 1000V is licensed per each server CPU. The designer must thus know how many ESX hosts will be initially used and how many will be used in the future in order to plan and select the proper licensing pack.
- VLAN scheme: For the Cisco Nexus 1000V deployment, a minimum of three VLANs is required: management, control, and packet. The designer must reserve and assign these VLAN IDs from the free VLAN ID pool. Although the VLAN IDs could be any of those already used, it is recommended to separate VLAN IDs.
- Design VSM deployment.

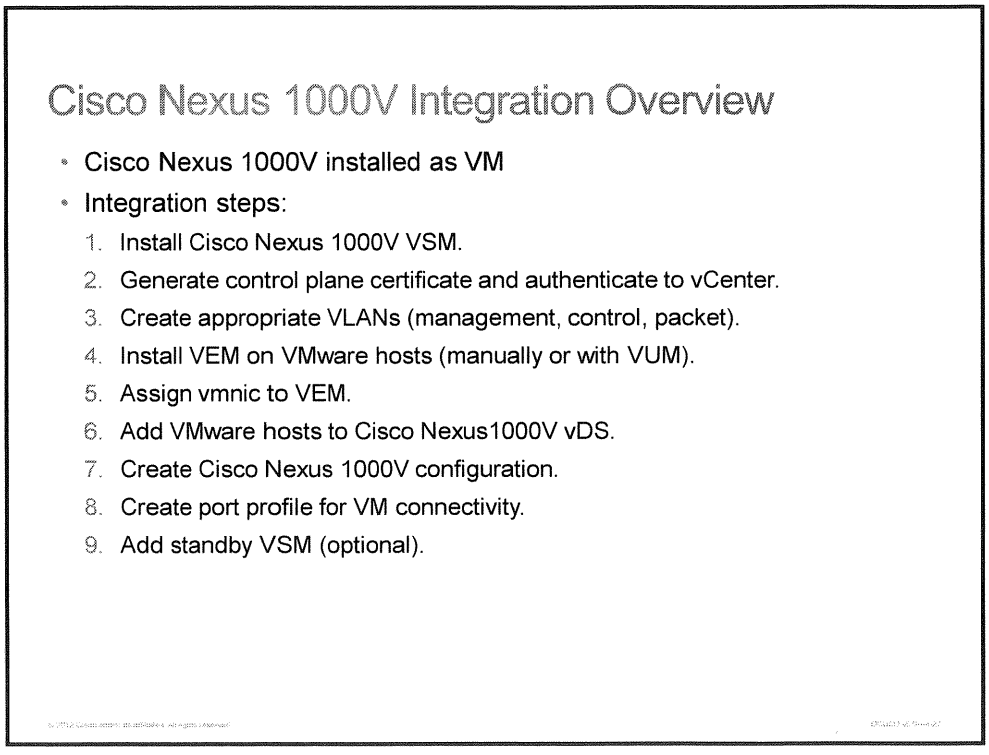
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

VSM Deployment Design

VSM deployment design must address the following Cisco Nexus 1000V aspects:

- Select the ESX host or hosts where the VSM appliance will be running—the host should have sufficient resources for the VSM VM.
- Define the management parameters, such as management IP address and login credentials.
- Define the domain ID parameters for multiple Cisco Nexus 1000V deployments. A unique domain ID per Cisco Nexus 1000V domain is recommended.
- If you must use the same control and packet VLAN pair for multiple domains, you must ensure that their domain identifiers are different.
- Define VSM restrictions: prohibited VMotion, disabled DRS, and fault tolerance for the VSM appliance.
- Define system uplink port profile for control and packet VLANs.
- Define the redundancy scheme for the VSM—standalone or active-standby (primary-secondary). Select the ESX host for the secondary VSM appliance. The selected host should be different from the one selected for the primary VSM.



The process of deploying and integrating the Cisco Nexus 1000V can be divided into nine steps, as defined in the figure.

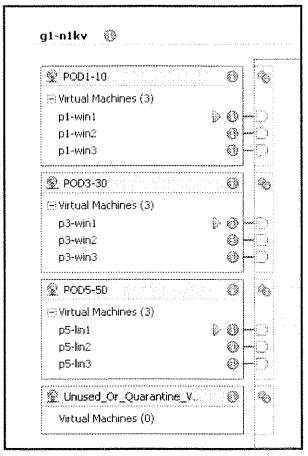
On the VMware side, the physical NICs have to be selected for the system uplink port profiles in order to enable VSM-to-VEM communication.

On the uplink switch side, the interface where ESX is attached has to be configured for 802.1Q trunking and must allow the expected VLANs—control, packet, and management.

https://t.me/learningnets

VM Deployment Considerations

- Cisco Nexus 1000V considerations
  - Define data VLANs
  - Define VM uplink port profile—trunk data VLANs
  - Define VM port profile—for VM connectivity
  - Per VM policies—security, QoS, etc.
- VMware vCenter considerations
  - Select vmnic for VM uplink port profile for VM connectivity
- Upstream switch considerations
  - ESX attached interface for trunk with allowed VLANs



Inside VMware vCenter Server, VMs are configured as before. Instead of defining network configuration in vCenter Server, port profiles are defined on the Cisco Nexus 1000V Virtual Supervisor Module and are displayed by vCenter as port groups.

Virtualization administrators can take advantage of preconfigured port groups and focus on VM management, while network administrators can use port profiles to apply policy for many ports at the same time. Together, both teams can deploy server virtualization more efficiently and with lower operational cost.

Once the Cisco Nexus 1000V deployment is designed, the VM deployment can be planned. This includes Cisco Nexus 1000V, VMware, and upstream switch settings.

Cisco Nexus 1000V Design for VM Deployment

The design should define the following:

- A VLAN scheme for data traffic from different VMs. The scheme would typically define multiple VLANs for VM connectivity.
- VM port profiles for VM connectivity. A common VM profile for applications with the same connectivity requirements should be defined.
- A per-VM port profile policy, which includes QoS, security, and other settings.

VMware Design for VM Deployment

The design should specify the port groups derived from port profiles for different VMs. The port group description should specify the connectivity policy, which is derived from the port profiles.

Upstream Switch Design

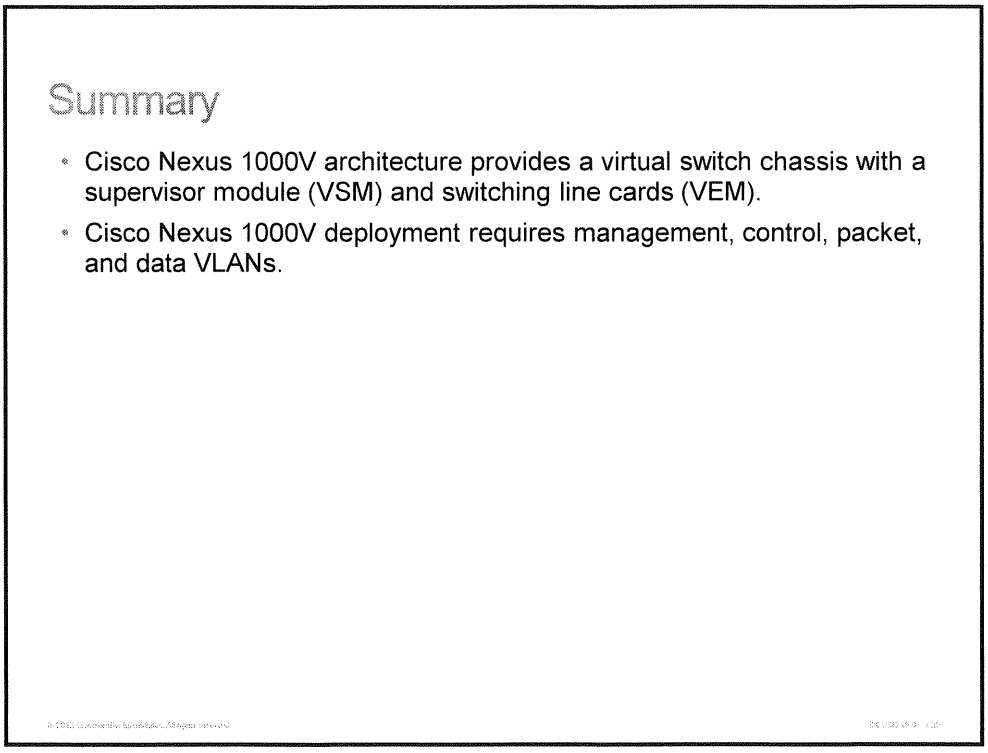
The design should specify the interface settings of the upstream switch where the ESX server is connected for the VMNIC that is carrying data VLANs.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Summary

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

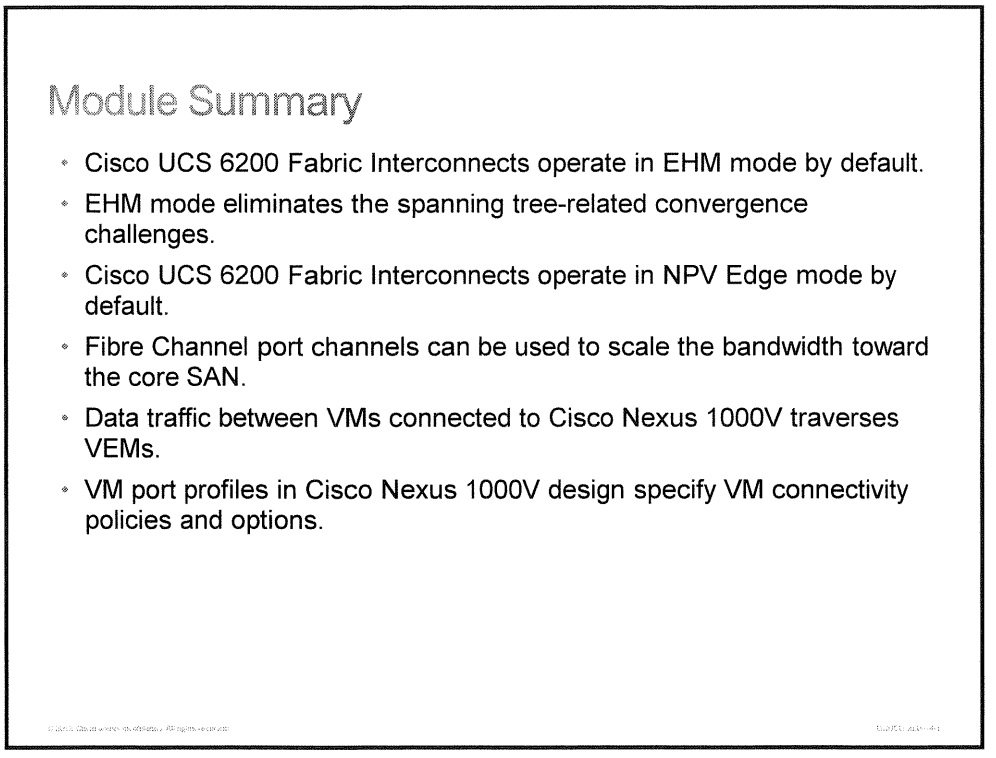


https://t.me/learningnets

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

Module Summary

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



https://t.me/learningnets

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

Module Self-Check

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

Q1) What happens by default when one of the LAN uplinks fails on fabric interconnect A? (Source: Designing the Unified Computing Network)

A) All vNICs that are pinned to that link fail over to fabric B.
B) All vNICs that are pinned to that link get repined to the remaining LAN uplinks.
C) All vNICs that are pinned to that link fail.
D) The operating system teaming driver disables the failed vNIC.

Q2) Which option can be used to connect DMZ-hosted servers to a DMZ network that is not connected to the internal LAN? (Source: Designing the Unified Computing Network)

A) switching mode
B) fabric failover
C) disjoint Layer 2 domain
D) VM-FEX

Q3) Which VLAN IDs can be used for FCoE VLAN? (Source: Designing Unified Computing Storage)

A) 3000 and higher
B) 4090 to 4096
C) any VLAN ID
D) any VLAN ID that is not already in use

Q4) What happens if the Cisco Nexus 1000V VSM fails and there is no secondary VSM? (Source: Designing the Virtual Access Layer)

A) All VM communication between hosts fails.
B) ESXi hosts part of the Cisco Nexus 1000V restart.
C) One of the VEMs takes over the VSM role.
D) VM communication is still operational.

https://t.me/learningnets

Module Self-Check Answer Key

Q1) B

Q2) C

Q3) D

Q4) D

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Module 5

Design Cisco Unified Computing Solutions Server Deployment

Overview

The final part of the Cisco Unified Computing System (UCS) solution design is the server deployment plan, which includes the naming convention, proper addressing, and service profile definition.

This module recognizes the server deployment aspects and provides examples of the naming convention and common policies with server deployment.

Module Objectives

Upon completing this module, you will be able to create the Cisco UCS server deployment design for the given solution examples. This ability includes being able to meet these objectives:

- Design the properties and characteristics of the Cisco UCS deployment that are in common for various application solutions
- Design management aspects of the given Cisco data center unified computing solution

https://t.me/learningnets

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

Lesson 1

Designing Cisco UCS Server Deployment

Overview

This lesson discusses the design of the properties and characteristics of the Cisco Unified Computing System (UCS) deployment that are common for various application solutions.

Objectives

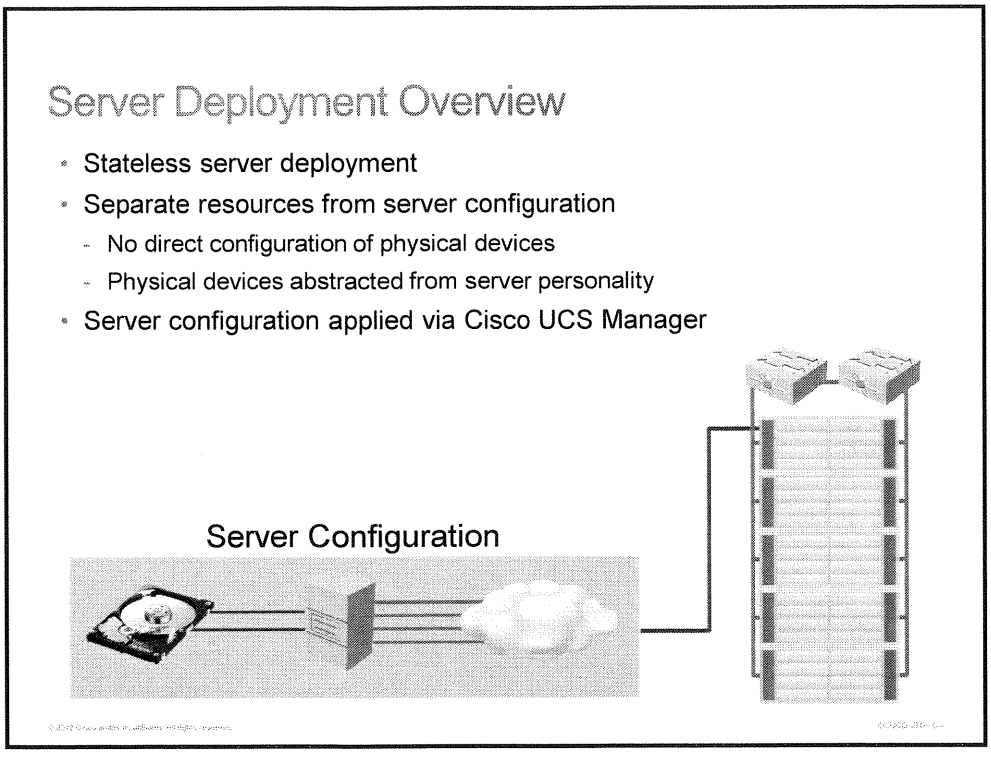
Upon completing this lesson, you will be able to design the properties and characteristics of the Cisco UCS deployment that are common for various application solutions. You will be able to meet these objectives:

- Identify the aspects of the server deployment
- Define the common naming convention for given solutions
- Define the UUID addressing for given solutions
- Define the MAC addressing for given solutions
- Define the WWN addressing for given solutions
- Define the common policies for hosts of a given solution

https://t.me/learningnets

Identify Server Deployment Aspects

This topic describes the aspects of the server deployment.



The Cisco UCS server deployment model separates the physical hardware of the server from the server configuration and personality by decoupling the logical domain from the physical resources. With this deployment model, no configuration is performed on the physical components. The configuration is applied implicitly on the server personality that is assigned to a physical server.

With this type of separation, stateless computing can be deployed where the applied attributes are not tied to the physical hardware. The server or computing hardware is simply a generic computing capacity that is flexible and can be dynamically controlled. The flexibility of the computing capacity enables the dynamic provisioning of servers where the entire server-related configuration is applied via the Cisco UCS Manager, with server elements also being controlled from that single management point.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Using Service Profiles

Every server is stateless and must be associated with a service profile in order to gain its identities and personality. A service profile is associated with a stateless server via manual association or automatically via a blade pool.

After you associate a service profile with a server, the server is then ready to have an operating system and applications installed. You can use the service profile to review the configuration of the server. When a service profile is disassociated from a server, the identity and connectivity information for the server is reset to factory defaults, and the server is returned to the pool (if it originated from a pool).

Servers and service profiles have a one-to-one relationship; each server can be associated with only one service profile, and each service profile can be associated with only one server. A service profile can be modified, cloned, or used to instantiate a template.

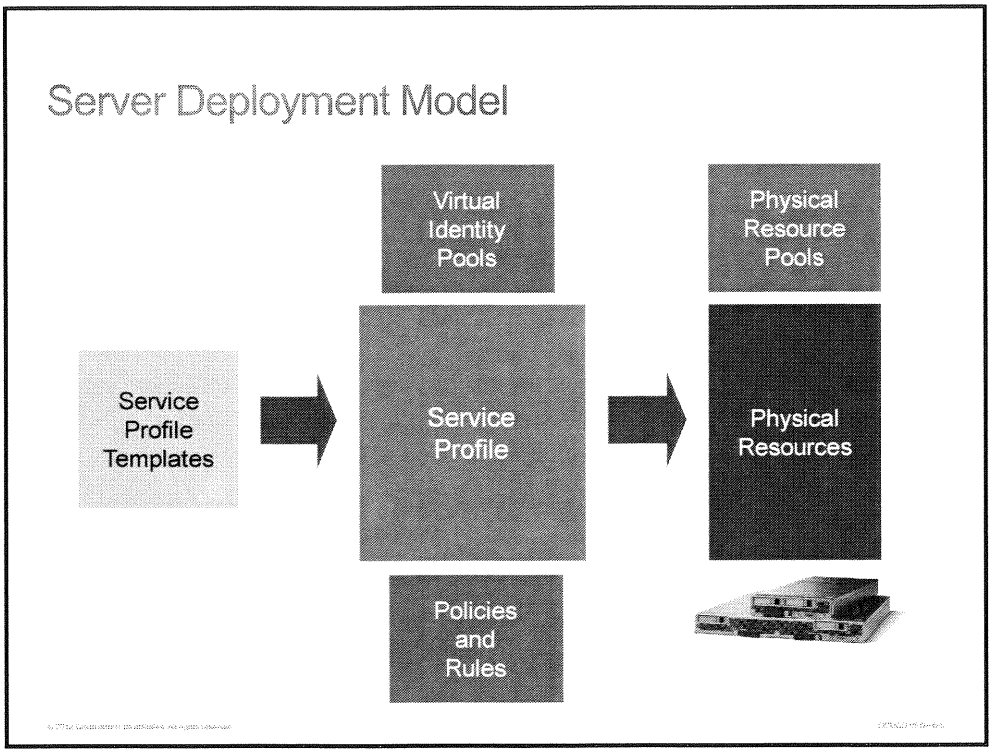
Note that within the Cisco UCS, the service profiles must be used for any server deployment, even if the blades will be managed as traditional individual rack servers. In such cases, the service profile provides LAN and SAN connectivity configuration.

A server that is associated with a service profile is similar to a traditional bare metal server in that it is ready to be used for production and run business services or application software.

Server Mobility

Server mobility enables the transfer of server identity seamlessly between the server blades and the service profiles. When a server must be moved from one blade to another, a simple operation of associating a service profile to another available server in a Cisco UCS system automatically includes complete migration of identities, firmware, and connectivity to the LAN and the SAN.

https://t.me/learningnets



The Cisco UCS server deployment model is built around the following elements, which will be described in more detail later in this lesson.

Server Personality

A server personality is composed of a service profile, which is a description of an individual server and all its related characteristics. A service profile can use various identity pools (for example, a MAC pool, a world wide name [WWN] pool, and so on) from which identity parameters are taken. Apart from the identity, the service profile is also assigned various policies that define rules, which in turn govern how the server is initialized and managed.

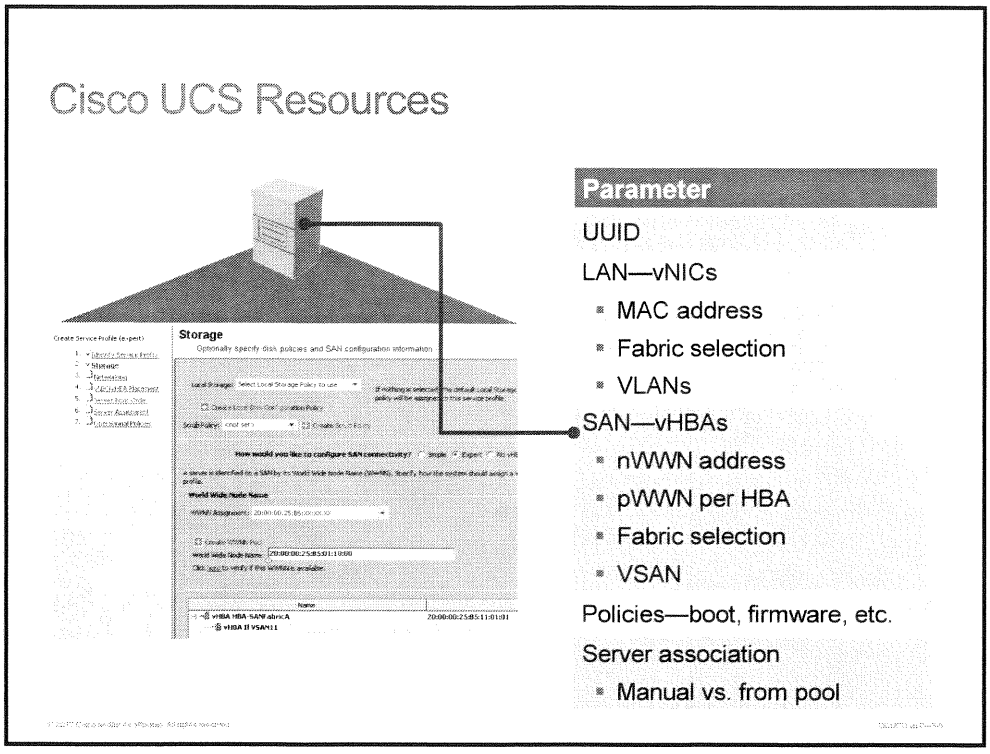
Service profile templates offer an approach to server deployment that eases large-scale server deployments.

Server Physical Resources

Server physical resources (for example, server blades) can be used individually or be pooled into common resource pools upon which service profiles are then applied. A blade must be associated with a service profile in order to be used for the operating system and application deployment.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets



As mentioned previously, in order to provide a flexible server deployment model, Cisco UCS separates resources and identities (personality) from the physical server (resources). The abstraction provides a flexible environment in which resources can be provisioned and migrated between physical devices rapidly.

Cisco UCS categorizes resources separately from the configurations that may reside on them. This is to allow rapid, stateless provisioning of the computing cloud.

LAN Resources

The Cisco UCS management system that runs on the Cisco UCS fabric interconnect controls various LAN resources. These resources include MAC addresses, VLANs, and LAN ports.

SAN Resources

The Cisco UCS management system is responsible for VSANs, node world wide name (nWWN), and port world wide name (pWWN) address assignment, as well as SAN ports. The central assignment of WWNs is critical to the portability of service profiles.

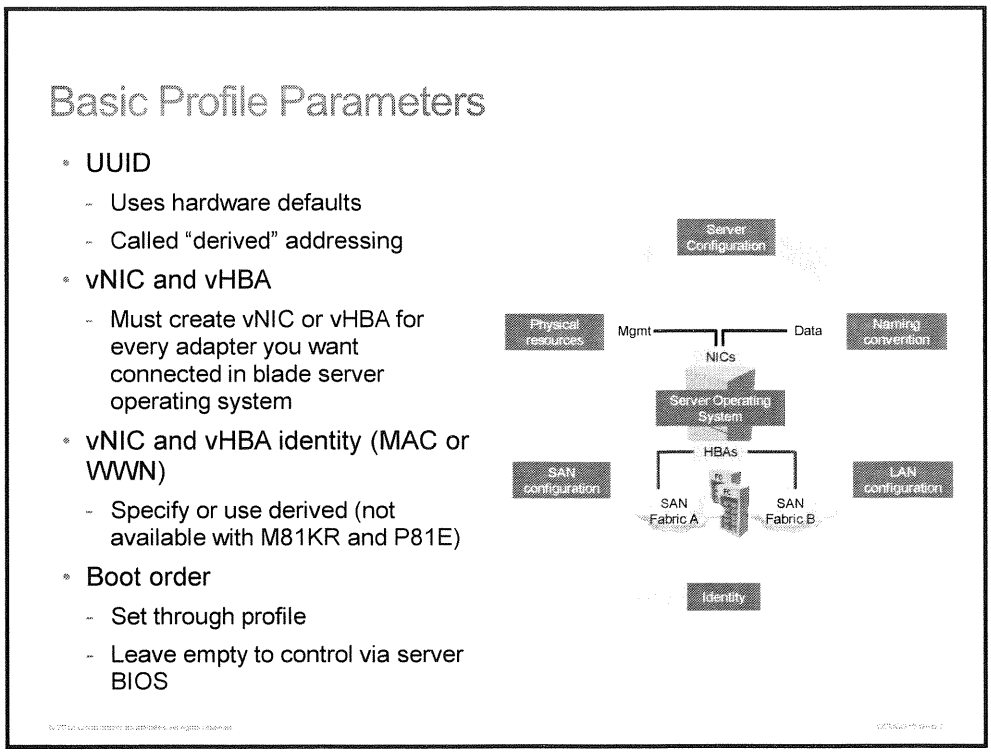
Compute Resources

The Cisco UCS management system manages compute nodes—blade servers, firmware, and Universally Unique Identifiers (UUIDs) within the Cisco UCS environment.

Management Resources

All system management is processed through the Cisco UCS Manager that runs on the fabric interconnect. The management VLANs, server keyboard, video, mouse (KVM) attachments, and low-level control are also Cisco UCS resources.

https://t.me/learningnets



Service Profile

A service profile is a self-contained logical representation (object) of a desired physical server, including its connectivity, configuration, and identity.

Each service profile serves a specific purpose so as to ensure that the associated server hardware has the necessary configuration, identities, and connectivity to LAN and SAN based on the requirements from the applications the server will host.

The service profile maintains configuration information about the server hardware, interfaces, fabric connectivity, and server and network identity—the server personality. This includes all of the unique information for the server, including MAC, pWWN, UUID, boot order, and so on.

The personality information is stored in a format that can be managed through the Cisco UCS Manager. All service profiles are centrally managed and stored in a database on the fabric interconnects.

Cisco UCS uses pools, policies, and service profiles to abstract state and configuration information from all the components that define a server, its service quality, and its connectivity. This makes the Cisco UCS system a stateless computing system. Cisco UCS enables an operational state and services to be separated from the server hardware and physical connectivity.

Network Connectivity

Each service profile specifies the LAN and SAN network connections for the server within the Cisco UCS infrastructure as well as out to the external network. An administrator is not required to manually configure the network connections for Cisco UCS servers and other components. All network configuration is performed through the service profile. When a service profile is associated with a server, the Cisco UCS internal fabric is configured with the information in the service profile. If the profile was previously associated with a different server, the Cisco UCS network infrastructure is reconfigured to support identical network connectivity to the new server.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

Hardware Component Configuration

When a service profile is associated with a server, the following components are configured according to the data in the profile:

- Server, including BIOS and Baseboard Management Controller (BMC)
- Adapters
- Fabric interconnect

The administrator is not required to configure these hardware components directly.

Server Identity Management

The administrator can use the network and device identities that are burned into the server hardware at the time of manufacture or use the identities that are specified in the associated service profile either directly or through identity pools, such as MAC, WWN, and UUID.

The following are examples of configuration information that the administrator can include in a service profile:

- Profile name and description
- Unique server identity (the UUID)
- LAN connectivity attributes, such as the MAC address
- SAN connectivity attributes, such as the WWN

Operational Aspects Configuration

The administrator can configure some of the operational functions for a server in a service profile, such as:

- Firmware packages and versions
- Operating system boot order and configuration
- Intelligent Platform Management Interface (IPMI) and KVM access

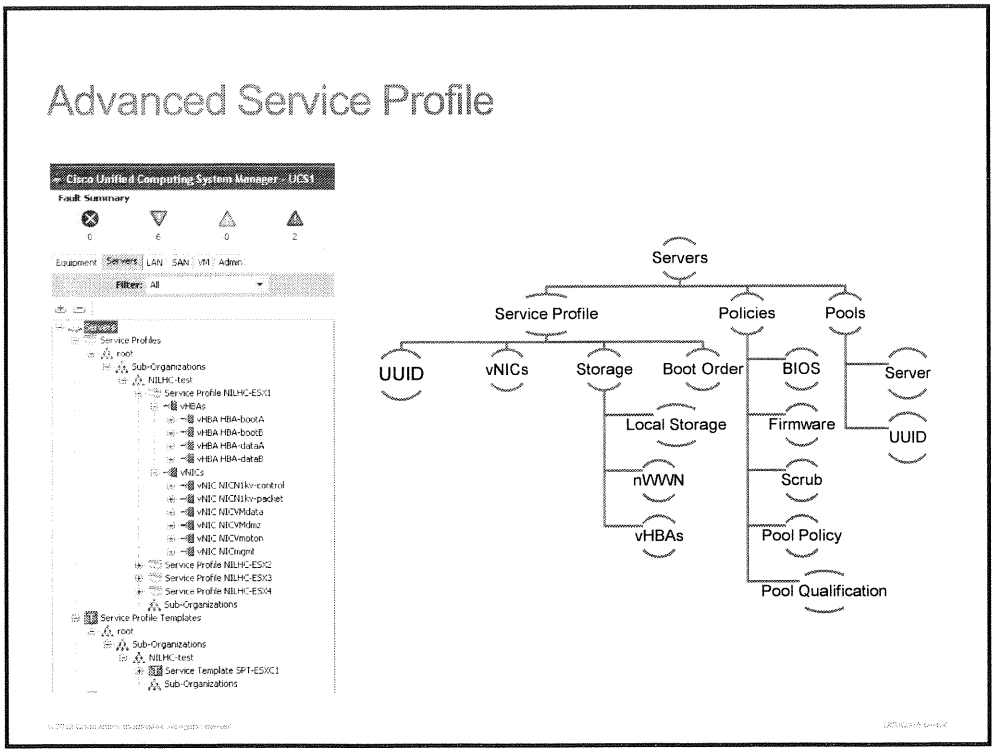
vNIC Configuration

A virtualized network interface card (vNIC) is configured on a physical network adapter and appears to be a physical network interface card (NIC) to the operating system of the server. The type of adapter in the system determines how many vNICs the administrator can create. For example, a Cisco UCS CNA M71KR adapter has two NICs, which means a maximum of two vNICs for each of those adapters can be created. A vNIC communicates over Ethernet and manages LAN traffic. At a minimum, each vNIC must be configured with a name and with fabric and network connectivity.

vHBA Configuration

A virtual host bus adapter (vHBA) is configured on a physical network adapter and appears to be a physical host bus adapter (HBA) to the operating system of the server. The type of adapter in the system determines how many vHBAs the administrator can create. For example, a Cisco UCS CNA M71KR has two HBAs, which means a maximum of two vHBAs for each of those adapters can be created. By contrast, a Cisco UCS 82598KR-CI does not have any HBAs, and so no vHBAs can be created. A vHBA communicates over Fibre Channel over Ethernet (FCoE) and manages SAN traffic. At a minimum, each vHBA must be configured with a name and fabric connectivity.

https://t.me/learningnets

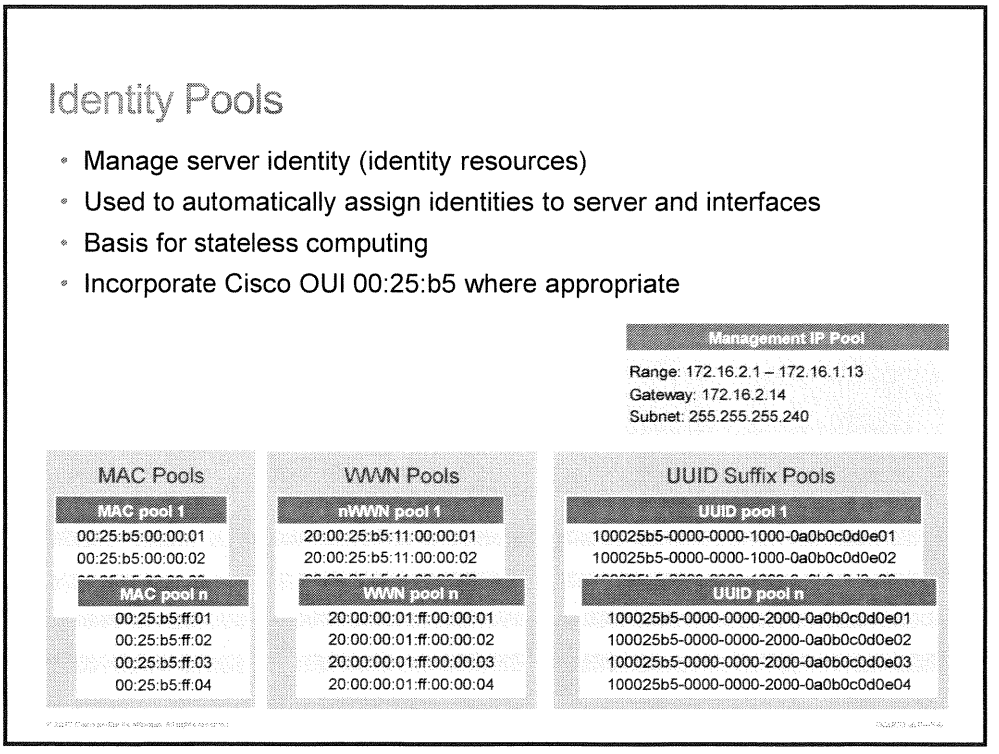


Cisco UCS resource pools are collections of identities—or physical or logical resources—that are available in the system. Pools are containers for resources and identity definitions. A pool contains only one type of resource. For instance, a MAC pool must contain only MAC addresses and an nWWN pool must contain only nWWN addresses.

All pools increase the flexibility of server deployments and allow an administrator to centrally manage the Cisco UCS system resources, which reduces the need for administrators to actively manage the use of resources.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets



Pools

There are two pool types—identity pools and blade pools.

If a pool is used to group identity information, such as MAC addresses, it can be preassigned with identity ranges for servers that will host specific information. For example, database servers could be configured to use the same range of MAC addresses, UUIDs, and WWN addresses.

Blade pools can be used to segment unconfigured servers with similar characteristics. Later, they can be associated with service profiles by means of a policy.

Multiple pools with different names can contain the same types of resources and even the same resources as well.

For example, a blade can be a member of multiple pools. Once a blade is consumed and assigned by the system, that blade is no longer available to any resource pools.

Default vs. User-Created Pools

Cisco UCS also has default pools that are predefined but not prepopulated. Default pools are used as a last resource when a regular pool has been drained of its resources in the local organization and there are no other resources available in either the parent or grandparent organization.

Identity Pools

Identity resource pools are used for server identity management and assignment. They allow the Cisco UCS solution to automatically assign identities to servers and their interfaces.

Using identity pools minimizes the administrative tasks that are needed for network, storage, and server administration. Identity pools are also key components for the stateless computing model, because the physical servers are unaware of the configuration and any server can be assigned to any identity and workload at any time.

There are various identity resource pools in existence, such as network MAC pools, storage WWN pools, UUID pools, and management IP pools.

https://t.me/learningnets

IEEE OUI

The Organizationally Unique Identifier (OUI) is assigned by the IEEE to allow hardware manufacturers guaranteed universal uniqueness of the addressing schemes for MAC and WWN addresses.

The OUI that is assigned to Cisco is 00:25:b5.

MAC Pools

A MAC address is a hardware address that uniquely identifies each node on a LAN. A MAC address—a 48-bit structure that is defined by the EUI-48 standard, which is a 48-bit extended universal identifier—has the form of MM:MM:MM:nn:nn:nn, where:

- MM:MM:MM (the first 24 bits) represents the OUI of the device manufacturer
- nn:nn:nn (the last 24 bits) is assigned by the organization

MAC addresses should be globally unique, or at least remain within an Ethernet broadcast domain.

Note A Cisco UCS MAC pool is restricted to MAC addresses that remain within the 00:25:b5:nn:nn:nn range.

When defining MAC addresses that are part of a certain pool, an administrator can encode meaning into the organizationally assigned field (the last 24 bits). Such a meaning could be the Cisco UCS cluster ID and data center, which would help preserve uniqueness within a Cisco UCS management organization.

The MAC address resource pools allow dynamic assignment of unique MAC addresses, thereby allowing servers to be stateless.

A MAC pool is a collection of network MAC addresses. The pool MAC addresses are unique in the Layer 2 environment and are available to be assigned to the vNICs on the servers. An administrator who creates one or more blocks of MAC addresses in a pool populates the MAC pool.

A MAC address block is a sequence with a first MAC address and a last MAC address. The block is created by specifying a starting MAC address and number of addresses.

An example of a MAC address block is presented here:

- First MAC: 00:25:b5:00:00:01
- Last MAC: 00:25:b5:00:01:00

The MAC address resource pools allow dynamic assignment of unique MAC addresses and thus allow servers to be stateless.

When a MAC address resource pool is defined, the size of the address block should remain within reasonable boundaries in order to omit unnecessarily large blocks that add size and overhead to the Cisco UCS Manager database.

Using MAC Resource Pools

Using MAC pools with service profiles releases an administrator from having to manually configure the MAC addresses that are used by the server that is associated with the service profile.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

MAC pools can be implemented in multitenancy environments. In such environments, you should use an organizational hierarchy in order to ensure that MAC pools can be used only by specific applications or business services.

For Cisco UCS to assign the MAC address from the pool to a server, the MAC pool must be included in a vNIC policy, which must be included in the service profile that represents the server.

The MAC address pools typically should be defined in the initial phase of Cisco UCS deployment in order to allow all stakeholders to agree upon the addressing protocols that will be used later on in the service profiles.

WWN Pools

A WWN is an address that uniquely identifies a server in a SAN environment. WWN addresses are used for two purposes:

- An nWWN is used to address the server.
- A pWWN is used to address each port on the vHBA.

The nWWN and pWWN address resource pools remove the server dependency of SAN resources so that hardware changes remain transparent to the storage network and devices. The zoning configuration, which couples initiators and targets, does not need to change upon hardware changes.

Note A WWN pool can include only nWWNs or pWWNs in the 20:nn:nn:nn:nn:nn:nn:nn, 20:00:00:25:b5:nn:nn:nn, or 5n:nn:nn:nn:nn:nn:nn:nn range. All other WWN ranges are reserved.

Defining WWN Addresses

First, WWN addresses should be globally unique. To achieve this within Cisco UCS and maintain standards, it is recommended that the WWN address incorporate an OUI. This way, the WWN address will not conflict with any other WWN addresses used in the SAN of an organization.

Second, the WWN address can also encode a Cisco cluster ID and data center, which would help preserve uniqueness within an organization.

Third, to distinguish between the nWWN and pWWN addresses, the second octet can encode meaning for the nWWN or pWWN—for example, 00 for nWWN and 01 for pWWN.

WWN Resource Pool

A WWN pool is a collection of WWNs for use by the Fibre Channel vHBAs in a Cisco UCS solution. The WWN pool contains blocks of WWN addresses similar to the MAC pool, and it is created and populated by an administrator who defines and configures one or more blocks of WWN addresses within the pool. Each WWN block or individual WWN can be assigned a boot target.

A WWN address block is a sequence that starts with the first WWN address and ends with the last WWN address. The block is created by specifying a starting WWN address and number of addresses.

An example of a WWN address block is presented here:

- First WWN: 20:00:00:01:11:00:00:01
- Last WWN: 20:00:00:01:11:00:01:00

https://t.me/learningnets

When a WWN address resource pool is defined, the size of the address block should remain within reasonable boundaries in order to omit unnecessarily large blocks that add size and overhead to the Cisco UCS Manager database.

Within the Cisco UCS, separate nWWN and pWWN pools are created.

nWWN Pools

An nWWN pool is a WWN pool that contains only nWWNs. If included in a service profile, the associated server will be assigned an nWWN from that pool.

pWWN Pools

A pWWN pool is a WWN pool that contains only pWWNs. If included in a service profile, the port on each vHBA of the associated server will be assigned a pWWN from that pool.

Using WWN Resource Pools

Using WWN pools in service profiles relieves an administrator from manually configuring the WWNs that will be used by the server that is associated with the service profile.

In multitenancy environments, a WWN pool can be used to control the WWNs used by each organization.

Similar to the MAC address pools, the WWN address pools should typically be defined in the initial phase of Cisco UCS deployment to allow all stakeholders to agree upon the addressing that will later be used in the service profiles.

UUID Pools

A UUID is an identifier that uniquely identifies each computing resource server.

A UUID consists of a prefix and a suffix. Within Cisco UCS, a suffix UUID is 8 bytes long, consists of 16 hexadecimal characters, and uses the format HHHH-HHHHHHHHHHHH. The pool is prefixed with 8 bytes that are unique to Cisco UCS in the format FFFFFFFF-FFFF-FFFF.

The UUID assigned to the server is a "prefix-suffix" number in the format FFFFFFFF-FFFF-FFFF-HHHH-HHHHHHHHHHHH.

Like the MAC and WWN addresses, the UUID can also encode the OUI to maintain global uniqueness. The administrator can also encode the UCS cluster ID and data center into the UUID so as to further maintain uniqueness in the organization.

A Cisco UCS UUID suffix pool is a collection of BIOS UUID suffixes that is available to be assigned to servers. The UUID suffix pool ensures that these variable values are unique for each server associated with a service profile, which then uses that particular pool to avoid conflicts.

The suffix UUIDs are populated by an administrator, who creates one or more blocks of UUID numbers in the pool. The UUID suffix pool is created by specifying the UUID prefix and the block of UUID suffixes. A UUID suffix block is created by specifying the first UUID suffix and the number of UUID suffixes.

An example of UUID suffix block definition is presented here:

- UUID prefix: 10000000-0000-0000
- First UUID: 1000-000000000001
- Last UUID: 1000-000000000255

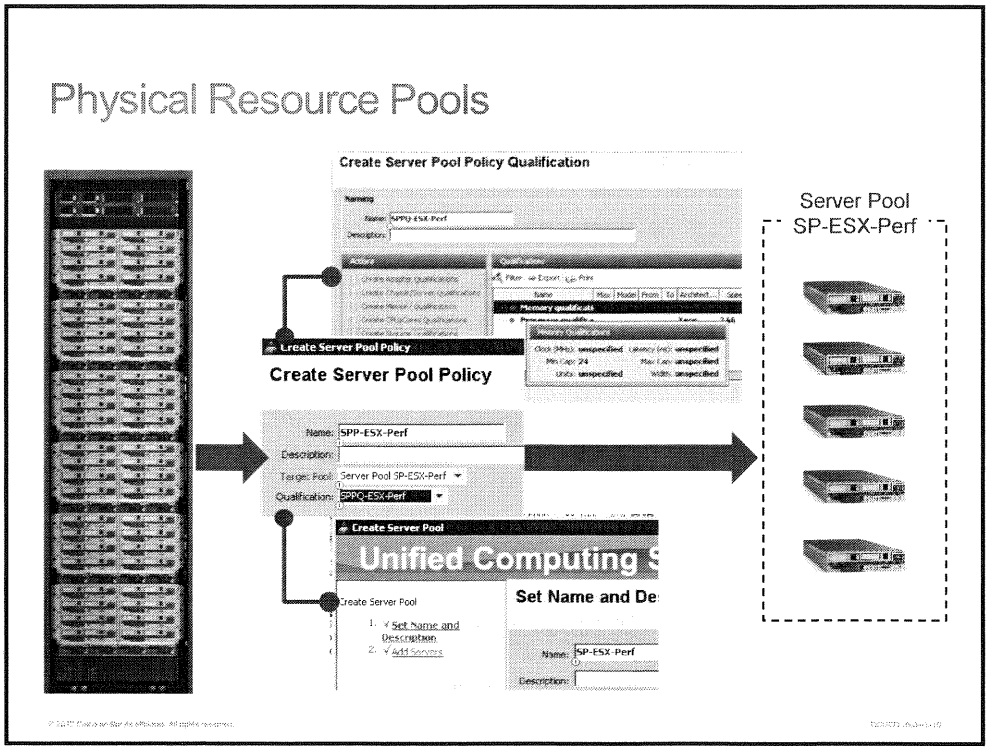
© 2012 Cisco Systems, Inc.

https://t.me/learningnets

When a UUID suffix resource pool is defined, the size of the address block should remain within reasonable boundaries in order to omit unnecessary large blocks that add size and overhead to the Cisco UCS Manager database.

Using UUID Suffix Pools

Using UUID suffix pools in service profiles relieves the administrator from having to manually configure the UUID of the server associated with the service profile.



Server blades can also be pooled. A server pool contains a set of available stateless servers. These servers typically share the same characteristics. These shared characteristics may be the location of each server in the chassis or an attribute such as server type, amount of memory, local storage, type of CPU, or local drive configuration.

A server pool can include servers from any chassis in the system. A given server can belong to multiple server pools.

Server pools can be used in multitenancy deployments, in which organizations designate one or more server pools to be used by a specific organization. For example, a pool that includes all servers with two CPUs could be assigned to the marketing organization, whereas all servers with 64 GB of memory could be assigned to the finance organization.

The Cisco UCS administrator has two options to populate the server pool:

- Manually assign a server to a server pool
- Use server pool policies and server pool policy qualifications to automate the assignment

Whether the server pools will be manually or automatically populated, each pool must first be defined with a name.

https://t.me/learningnets

Manual Server Pool Population

Manual population is a simple, though not very efficient, way to assign pool membership for a server. The administrator manually maintains and manages the server membership and must manually evaluate the inventory of each server in order to assign it to one or more server pools.

Automatic Server Pool Population

A server pool can be automatically populated via policies. Two different policies are used together to determine which pool a server should be a member of:

- Server pool qualification policy
- Server pool policy

Note Server pool automatic population occurs when an individual server blade is discovered. If the administrator wants to apply automatic population to existing server blades, they must be reacknowledged.

Server Pool Qualification Policy

A server pool qualification policy describes the hardware qualification criteria for servers, such as the number of CPUs, memory capacity, type of adapters, and so on. A server that fulfills all of the defined criteria in the policy is considered a qualified server.

Server Pool Policy

A server pool policy describes which server pools the server becomes a member of if it has in fact qualified for a certain server pool qualification. The same server can meet multiple qualifications; thus, multiple pool policies can point to the same pools.

Using Server Pools

An administrator can use a server pool instead of a physical server blade to quickly deploy Cisco UCS service profiles that are based on predefined requirements. An individual server blade must be associated with only one service profile at a time.

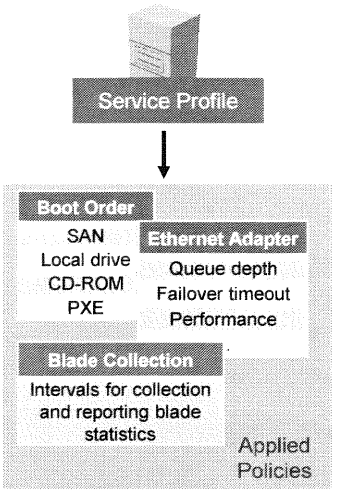
When a server pool is associated with a service profile, Cisco UCS Manager automatically selects an available server blade—one that is currently discovered, is not associated with any other service profile, and is not in the process of being associated or disassociated with a service profile.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Policies Overview

- Policy-driven management of Cisco UCS components
  - Ensures consistent computing environment
- Policy
  - Determines how Cisco UCS components act in given circumstances
  - Separates functions within a system
  - Different polices defined for network, storage, server
- Used within service profiles



Cisco UCS Manager is a policy-driven management device manager. Policy-driven management is one of the key features of Cisco UCS Manager, and it helps IT organizations to better define and implement their own best practices. Because data centers are becoming increasingly more dynamic, the definition, consumption, and resolution of policies is a key enabling technology for making infrastructure devices portable and able to be reconfigured dynamically.

Policies help to ensure that consistent, tested, and compliant systems are used, thereby reducing the risk of issues caused by repetitive manual tasks. Rules are defined in the form of policies inside the Cisco UCS Manager. Policies are centrally defined and enforced at the endpoint. Policies can be used to enforce behavior of Cisco UCS components, such as network, storage, and servers. Such an approach removes the stateful information from a device (blade) and improves mobility and scalability by making it stateless. An administrator can import and export policies into Cisco UCS.

Policies determine how Cisco UCS components will act in specific circumstances. An administrator can create multiple instances of most policies. For example, an administrator might want different boot policies so that some servers can Preboot Execution Environment (PXE) boot, some can SAN boot, and others can boot from local storage.

Policies also allow separation of functions within the system. A subject matter expert can define policies that are used in a service profile, which can then be created by someone who does not have that subject matter expertise. For example, a LAN administrator can create adapter policies and quality of service (QoS) policies for the system. Someone who has limited or no subject matter expertise with LAN administration can then use these policies to create a service profile.

https://t.me/learningnets

Policy Types

An administrator can create and use two types of policies in the Cisco UCS Manager:

- Configuration policies, which configure the servers and other components
- Operational policies, which control certain management, monitoring, and access control functions

Configuration policies make up most the policies in the Cisco UCS and are used to describe configurations of different components of the system. Operational policies determine how the system behaves under specific circumstances.

Policy Scope

Policies also have a scope—they can be either global or local. Global policies are defined for the entire Cisco UCS solution, whereas local policies are defined at the organization level. Examples of global policies include fault collection policies, Call Home policies, and the various statistics collection policies.

Default versus User Policies

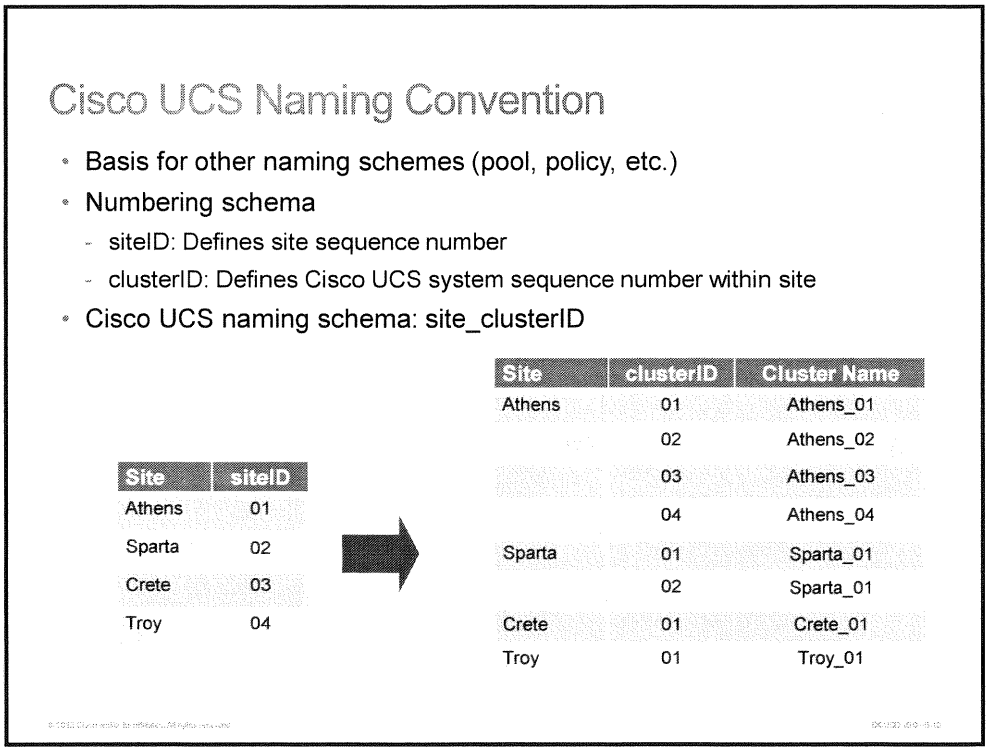
Cisco UCS has default policies that are used in the absence of other user-specific created policies. The predefined policies typically default to normal system operation behavior.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Design Naming Convention

This topic describes how to define the common naming convention for given solutions.



The server deployment design must define the server personalities for the servers the company will use. The personality—which is defined by the service profile—is then applied to the physical server blade that constitutes a server.

The service profile needs information about server characteristics, such as MAC address or addresses, the WWN address or addresses that the server should use, the UUID, which physical blades the logical server should use (which physical blade has the desired processor), the amount of memory, and the proper network adapter.

Thus, the server deployment design can be divided into two major design steps, as follows:

- Defining the resources and policies that are used by service profiles and templates
- Defining the service profile templates and profiles, which utilize the previously defined resources

The resource policy definition consists of:

- Defining the naming convention
- Defining the physical resources pools (the server pools in which servers with similar physical characteristics are placed)
- Defining the virtual identity pools (MAC, WWN, UUID, and management IP address pools)
- Defining the policies and rules used by service profiles and templates (boot policies)

The service profile and template design consists of defining the naming policy for the servers and defining the elements that constitute the template or profile.

https://t.me/learningnets

The first thing to define for the server deployment is the naming convention for all of the elements that will be created during the Cisco UCS configuration. Although this is not necessary, it is recommended to do so in order to ease not only the deployment but also the management and possible troubleshooting.

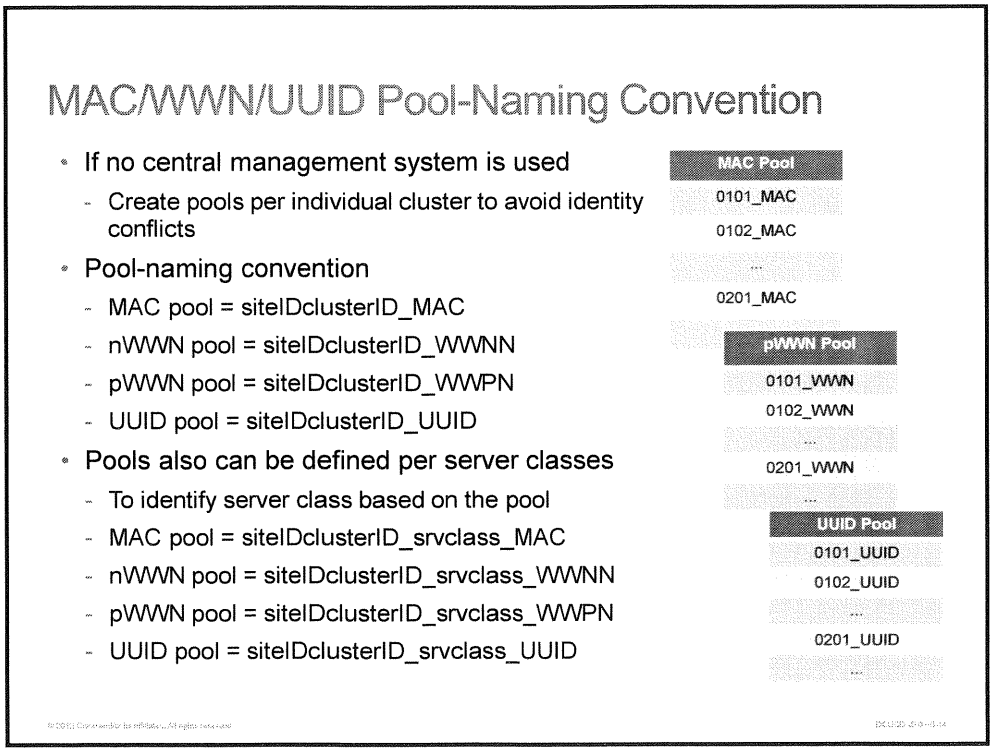
The naming convention must take care of how the various pools, polices, service profiles, templates, and so forth will be named.

When deploying multiple Cisco UCS systems on multiple locations, each system—such as a Cisco UCS cluster—should have a name.

For example, a naming convention can be built using a numbering convention, as follows:

- Each site is given a siteID—a numeric identifier.
- Each cluster (Cisco UCS system) is given the clusterID—a numeric identifier.

The Cisco UCS cluster could be named by using the location name and the Cluster ID.



Second, the naming convention for the MAC, WWN, and UUID pools that will be created must be defined. This naming convention is similar to the naming of Cisco UCS clusters. When multiple Cisco UCS clusters are deployed on multiple locations, the pool name can be formed from the siteID, clusterID, and pool type (MAC, WWN, and UUID).

In a case where a single large pool-per-resource type is required per cluster, the following naming convention can be used:

- siteIDclusterID_MAC for the MAC pool
- siteIDclusterID_WWN for the WWN pool
- siteIDclusterID_UUID for the UUID pool

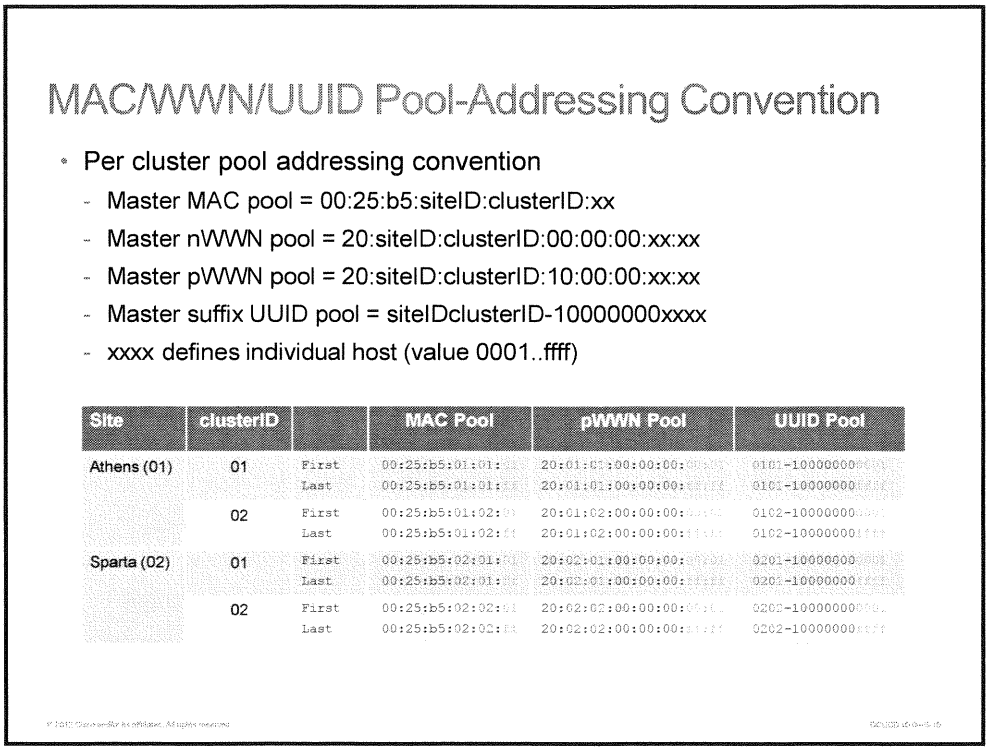
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

If there is a requirement that different server classes should use MAC/WWN/UUID addresses from different pools, the naming convention can take these into consideration by adding the server class into the name:

- siteIDclusterID_srvclass_MAC for the MAC pool
- siteIDclusterID_ srvclass_WWN for the WWN pool
- siteIDclusterID_ srvclass_UUID for the UUID pool

The granularity of pool naming depends on your requirements—in other words, how many of a certain pool type must be defined.



Once the naming convention has been defined, the addressing convention has to be defined. The addressing convection is used to define the list of addresses (MAC, WWN, and UUID) that will be part of a certain pool.

It is recommended that different MAC, WWN, and UUID pools be used on different Cisco UCS clusters, especially if they are on the same physical location. Thus, the clusterID should be part of the individual address.

If you need pools with such granularity to determine the exact Cisco UCS cluster and server class just by looking at the ID value or address in the pool, the addressing convention could be designed in the following manner:

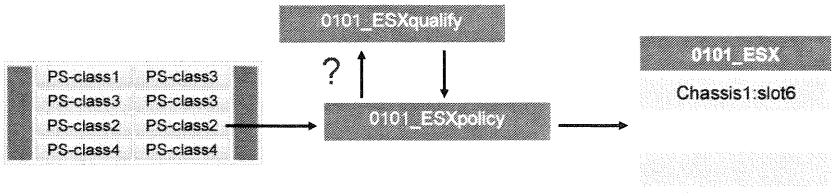
- MAC pool—00:25:b5:siteID:clusterID:xx
- nWWN pool - 20:siteID:clusterID:00:00:yy:xx:xx
- pWWN pool - 20:siteID:clusterID:10:00:yy:xx:xx
- UUID suffix pool—siteIDclusterID-100000yyxxxx

The xxxx stands for the individual host in a range of 0001 to ffff.

https://t.me/learningnets

Server Pool Convention

- Design server pools to group similar physical servers
  - Prefer automatic pool population with policy
- Naming convention
  - Server pool = siteIDclusterID_srvpool
  - Server pool policy qualification = siteIDclusterID_srvqualification
  - Server pool policy = siteIDclusterID_srvpolicy



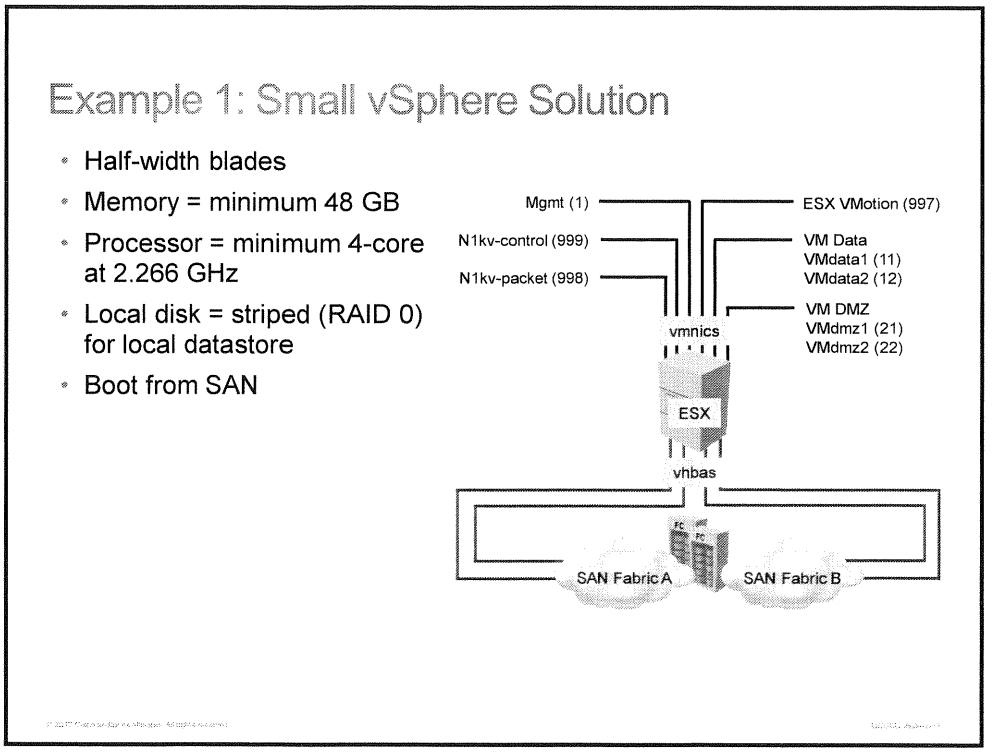
Server pools are used to group server blades that have similar characteristics. The pools can be populated manually, but this is not a recommended technique because the membership must also be maintained manually. A better option is to use automatic pool population—server pool policy qualification and server pool policy are used for this purpose.

The pool and policies must also use names and a naming convention, which can include the siteID and clusterID in a multisite, multicluster Cisco UCS deployment.

- The server pool name can comprise the siteIDclusterID_srvpool, where srvpool describes the server types.
- Server pool policy qualification can comprise the siteIDclusterID_srvqualification, where srvqualification stands for the qualification criteria name.
- The server pool policy can comprise the siteIDclusterID_srvpolicy, where srvpolicy stands for the srvpolicy name.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets



In this example, the VMware vSphere deployment is performed with four ESXi hosts.

LAN Connectivity Requirements

The physical connectivity requirements dictate that more than two vSphere hosts should be connected to two separate fabrics for each LAN segment that requires redundancy.

Thus, there will be two separate physical LAN fabrics on the Cisco UCS cluster level, which converge in the LAN core on the same set of devices. It has been determined that two 10 Gigabit Ethernet uplinks have enough throughput for all the traffic.

SAN Connectivity Requirements

Likewise, the SAN connectivity should be redundant for each segment or HBA that connects vSphere hosts to the SAN fabric. Furthermore, you must use separate segments for accessing the boot and data partitions of the vSphere hosts and environment.

Thus, from the Cisco UCS cluster perspective, the two separate physical SAN fabrics will be used and connected to upstream Fibre Channel switches with the N-Port ID Virtualization (NPIV) feature enabled. Boot and data traffic will be separated through zoning in the SAN and logical unit number (LUN) masking on the drive array side.

https://t.me/learningnets

Host Characteristics

Each VMware ESX host should be deployed using hardware with these minimum characteristics:

- **Blade form factor:** Half-width blade
- **Memory:** At least 48 GB at 1067 MHz
- **CPU:** At least 4-core at 2.266 GHz
- **Boot policy:** ESX should be booted from SAN to support fast blade replacement and complete stateless configuration
- **Local disk:** Used by the local datastore to store ISO images and nonmovable virtual machines (VMs) (such as the Cisco Visual Switch Manager [VSM] for the Cisco Nexus 1000V)
- **Adapter firmware:** Cisco UCS Manager version 2.0(1w) should be used because it is certified to be used with the selected drive array

Host Connectivity Requirements

The individual vSphere host has the following connectivity requirements:

- **Management segment:** This segment connects the host to a management segment that is used for management, configuration, and state maintenance, as well as for communication with the vSphere vCenter server. For this example, the default VLAN 1 will be used.
- **ESX VMotion segment:** This segment connects the vSphere hosts in a cluster to enable migration of VMs from one host to another, as well as to enable advanced VMware vSphere high-availability and resource management functionalities like high availability , fault tolerance, Distributed Resource Scheduler (DRS), and Distributed Power Management (DPM). For this example, VLAN 997 will be used.
- **VM data segment:** This segment is used for communication between VMs that reside in the private sector of the data center. VLANs 11 and 12 have been currently identified as required, but others can be added later. The segment must thusly be configured with trunking capability.
- **VM DMZ segment:** This segment is used for communication between VMs that reside in less trusted networks. VLANs 21 and 22 have been currently identified as required, but others can be added later. As with the private VM data segment, this segment must be configured with trunking capability.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Example 1: Naming Convention and Segments

Identifier	Description
ESX	Denotes ESX hypervisor
SubsegID	Denotes subsegment ID (for example, VLAN on physical segment)
LANseg	LAN segment
SANseg	SAN segment

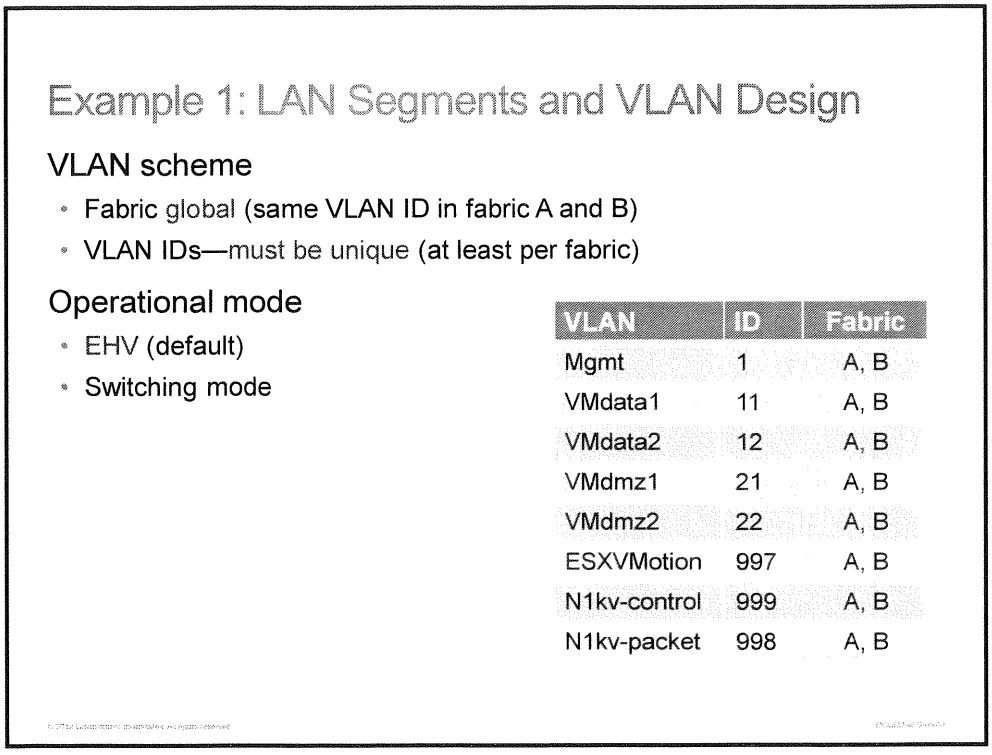
LAN Segment	ID	SAN Segment
Mgmt	001	SAN-FabricA
VMdata	010	SAN-FabricB
VMdmz	020	
ESXVMotion	997	
N1kv-packet	998	
N1kv-control	999	

Type	Naming
MAC	MP-ESX*LANseg*
nWWN	WNP-ESX*poolID*
pWWN	WPP-ESX*SANseg*
UUID	USP-ESX*poolNum*
VLAN	*LANsegSubsegID*
VSAN	*SANsegSubsegID*
vNICs	NIC-*LANseg*
vHBAs	HBA-*SANseg*
Server pool	SRVP-ESX*poolID*
Server pool policy	SRVPP-ESX*poolID*
Server qualification	SPPQ-ESX*poolID*
Host firewall package	HFP-ESX*poolID*

To ease the deployment effort, and address any troubleshooting that might be needed later, the naming convention will incorporate segment information into the addresses and identifiers. This is indicated in the tables in the figure.

Note that this is only a suggestion of a good design practice and is not enforced by Cisco UCS or by the Cisco UCS Manager.

https://t.me/learningnets



The VMware vSphere deployment will be physically connected to a single physical LAN core—that is, the LAN core is not disjointed (formed from two or more disconnected domains).

For this reason—as well as the Cisco UCS level redundancy used to achieve failover between the fabric and manual pinning for the uplink selection—the Cisco UCS cluster will operate in Ethernet host virtualizer (EHV) mode for Ethernet. This also eliminates Spanning Tree Protocol (STP) in the access layer.

VLAN Design

The VLANs that are used will have the same IDs in both fabrics. This means they will be defined globally in Cisco UCS with the VLAN IDs that are specified in the table in the figure.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Example 1: SAN Configuration

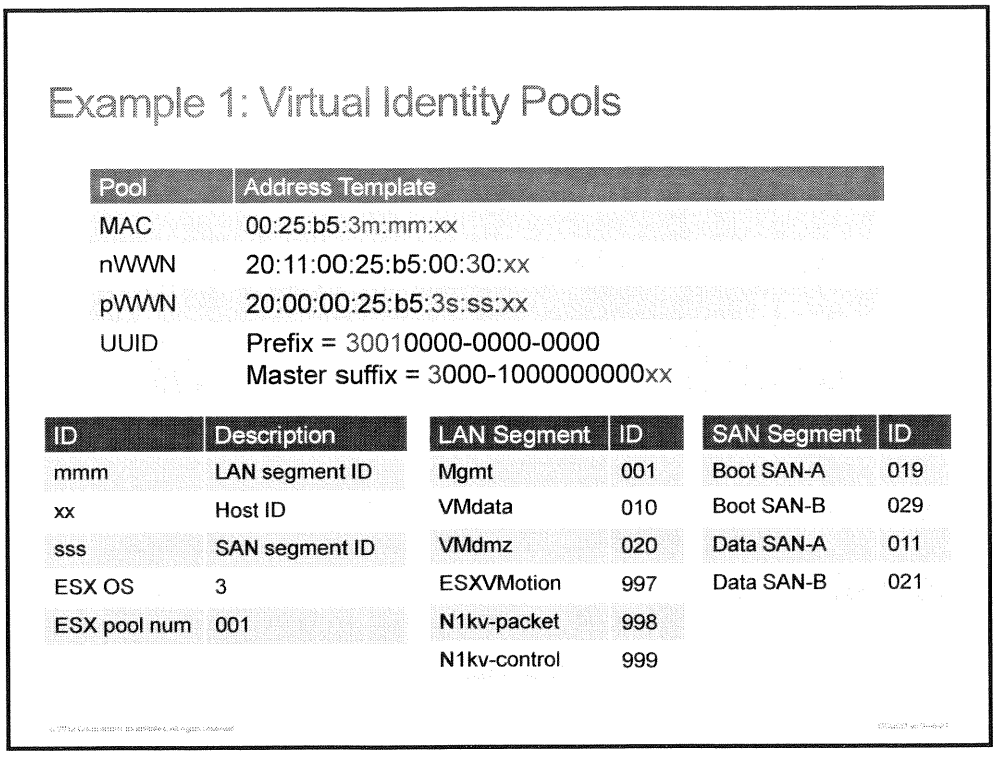
- Must be defined before Fibre Channel uplink configuration
- VSAN scheme
  - Fabric A or B
  - VSAN IDs—must be unique (at least per fabric)
  - FCoE VLAN ID—must be unique (at least per fabric)
- Operational mode is NPV Edge
  - Upstream Fibre Channel switches must be enabled with NPIV

VSAN	VSAN ID	Fabric	FCoE VLAN ID
SAN-FabricA	10	A	1010
SAN-FabricB	20	B	1020

The SAN design is simple: two separate physical fabrics using the VSAN IDs, as specified in the table. Because FCoE is used, VLAN FCoE IDs have to be specified, which is also indicated in the table. A good design practice is to add a prefix to the VSAN ID in order to get an FCoE VLAN ID if the same VLAN ID is already in use (for example, adding 10 in front of VSAN to get the FCoE VLAN ID).

In the core SAN fabric, the core Fibre Channel switches must be enabled with the NPIV feature. Also, for the Cisco Multilayer Director Switch (MDS), the Cisco UCS attached interfaces must be part of the same VSAN (that is, in access mode). Thus, from a design perspective, the Cisco UCS interfaces that are attached to the MDS must have the same configuration as the Cisco UCS Fibre Channel interfaces.

https://t.me/learningnets



To ease the deployment effort and troubleshooting in the future, and to have better control of the environment when it scales, the virtual identity pools with addresses and identifiers will be defined. This includes MAC, UUID, nWWN, and pWWN pools and addresses.

The addresses and identifiers will have segments and IDs encoded to further ease deployment and troubleshooting using the following fields:

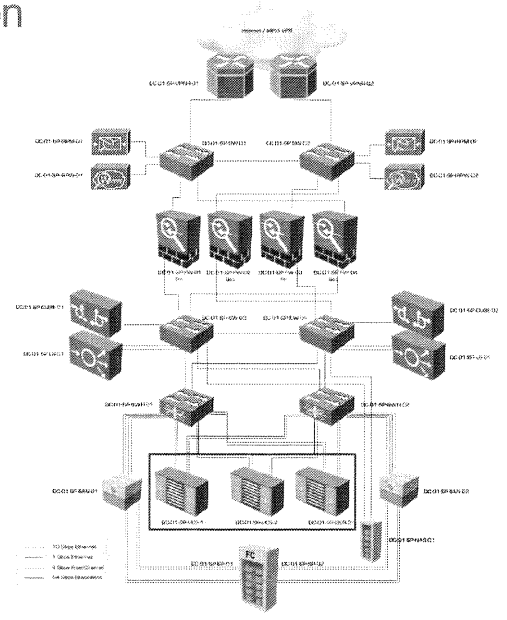
- **mmm:** Denotes the LAN segment ID and encodes the VLAN number within the MAC address. For the VM data segment, because there will be a single NIC used, the identifier will be the least common denominator. In other words, the last part of the VLAN ID will be omitted. The same is true for the VM DMZ segment.
- **xx:** Denotes the host ID, and is picked and determined by Cisco UCS itself.
- **sss:** Denotes the SAN segment ID and encodes the VSAN number that is added by the suffix, where 1 indicates data segment and 9 indicates the boot segment.
- **ESX pool number 001:** To anticipate future addition of vSphere pools, this first ESX pool will be denoted with 001, which will be encoded into the addresses.
- **ESX OS ID 3:** Because Cisco UCS enables deployment of multiple different operating systems and hypervisors, information about the operating system and hypervisor type will also be encoded into the addresses and identifiers. In this case, the vSphere hypervisor will have an ID of 3.

The table at the top of the figure lists the pool types and also specifies a template for creating addresses and identifiers in each pool that have these fields encoded.

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

Example 2: VDI Solution

- VPN aggregating routers
- External LAN switches
- Network, web, and email firewalls
- Internal LAN switches
- Load balancers
- Servers
- Storage systems
- SAN switches

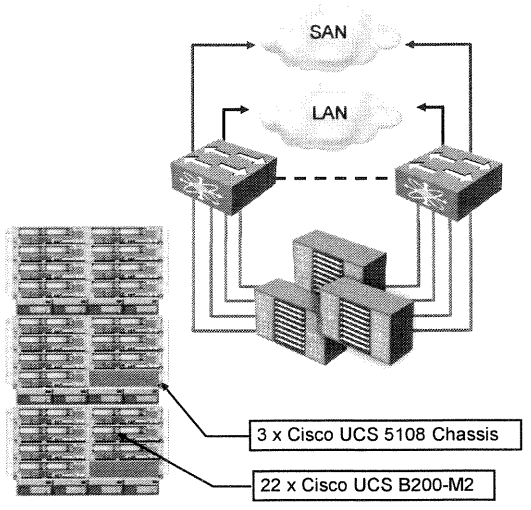


The second example depicts the virtual desktop infrastructure (VDI) data center infrastructure consisting of 22 physical servers in three Cisco UCS chassis. Two servers are used for management purposes and use VMware ESXi virtualization to host all of the supporting servers needed by the VDI solution. The remaining 20 servers also use VMware ESXi virtualization in combination with VMware VDI in order to host up to 2500 virtual desktops.

https://t.me/learningnets

Example 2: VDI Solution (Cont.)

- Three Cisco UCS Blade Chassis
- A total of 22 Cisco UCS B200 Blade Servers
  - 2 servers for infrastructure
  - 20 servers for VDI
- Scalability within the block:
  - Room for two more blade servers in existing chassis
  - Room for another blade chassis in the cabinet



The server blades in the UCS are divided into two separate VMware vSphere clusters:

- Infra cluster with two server blades—namely server blade 1/1 and 2/1 to host the ESXi hosts running VMs that provide the infrastructure services required for running the VDI block (AD, management, vCenter servers, and so on)
- VDI cluster with 20 server blades hosting ESXi hosts running VMs for customer virtual desktops and dedicated infrastructure VMs

The design and configuration of the Cisco UCS connect all of the components not only in a redundant but also highly available manner—the failure of a single component does *not* result in no server connectivity (if the configuration of the server operating systems is also done properly).

- Each server blade in the system is connected via two 10-Gb channels to the hosting chassis backplane—one 10-Gb channel to the left and one 10-Gb channel to the right I/O module (IOM). The two 10-Gb channels are Ethernet Data Center Bridging (DCB)-capable and carry all traffic destined and originated from a server blade (LAN, SAN, and management).
- Each chassis is connected via two 10 Gigabit Ethernet DCB links to individual fabric interconnect switches to achieve better distribution of server traffic between the IOM uplinks and increase bandwidth availability.
- Each fabric interconnect switch is connected with:
  - two 10-Gb/s Ethernet uplinks in a port channel to the LAN (Catalyst 3750 switches)
  - four 4-Gb/s Fibre Channel uplinks in a port channel to the SAN (MDS 9124 switches)

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

Example 2: Object Type Naming Convention

Name structure for nondefault objects: PPPPSSSS

- PPPP: Up to four-character object prefix indicating object type (provided in the table below)
- SSSS: Suffix part of the object name

Object Type	Object Prefix
Bios policy	bios
Boot policy	bp
Customer internal VLAN	CustIn
Enterprise customer internal VLAN	DCust
Host firmware package	hfp
LAN pin group	lpg
Local disk config policy	ldp
MAC pool	mp
Maintenance policy	mtnp
Management firmware package	mfp
Network control policy	ncp
Organization	*Organization name is created from the block ID and purpose and does not have general prefix* *e.g.* DC-01-SP-Infra, DC-01-SP-OiC
Power control policy	pcp
SAN fabric	SANFabric

The Cisco UCS configuration consists of several object types. To make the configuration easier for understanding and to designate the objects by their names, the object names will consist of object type prefix and object suffix. Note that the maximum object name is 16 characters.

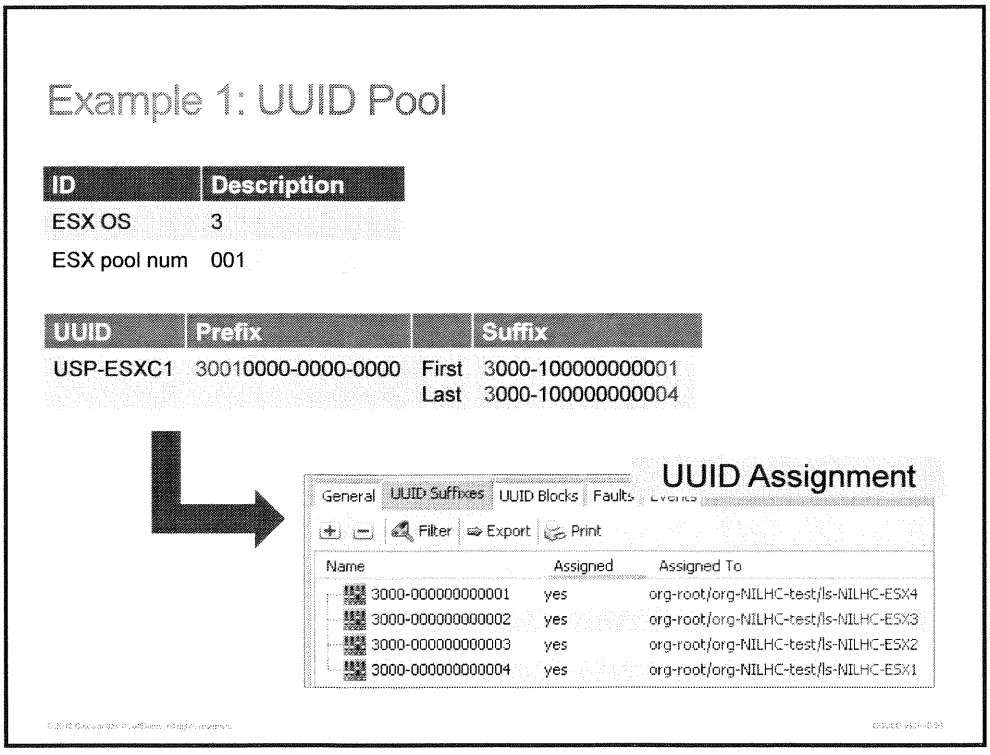
The nondefault objects have the following general name structure: **PPPPSSSS**:

- **PPPP:** Up to four-character object prefix indicating object type
- **SSSS:** Suffix part of the object name

https://t.me/learningnets

Design UUID Addressing

This topic describes how to define the UUID addressing for given solutions.

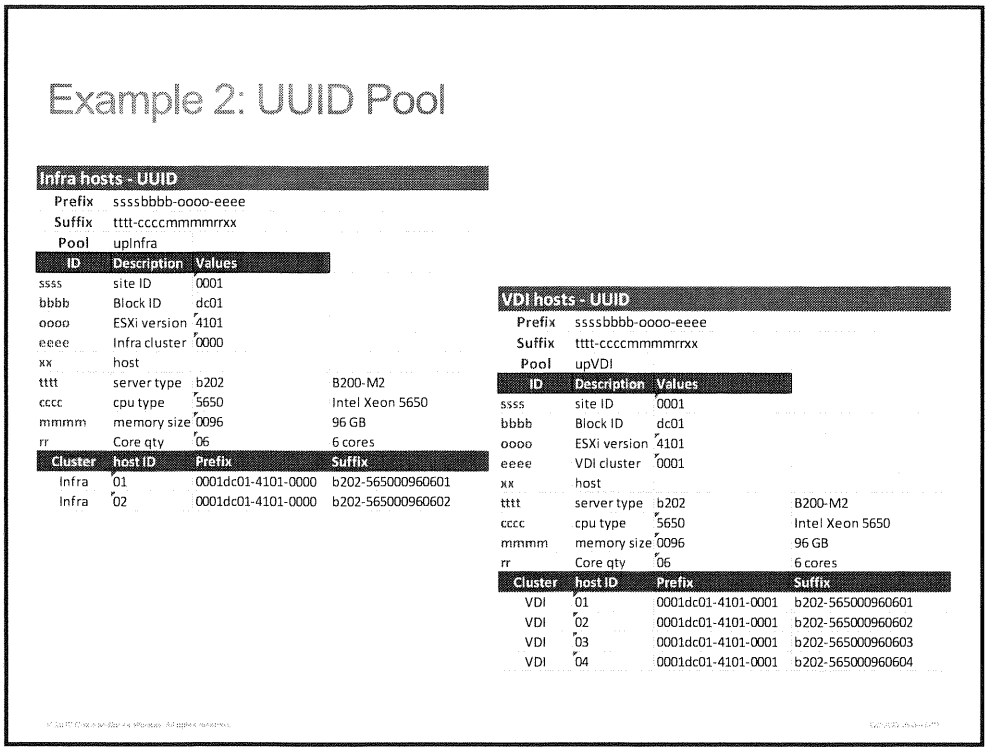


The UUID pool and addresses are defined as the UUID template that was defined earlier. The UUID addresses encode the operating system type (3 for vSphere host) and vSphere cluster ID (001 for the first one).

As with the MAC and pWWN addresses, the NICs, and the HBAs, the vCenter configuration is easier to read due to encoded fields.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets



UUIDs are used to uniquely identify a sever personality. In the VDI solution example, UUIDs are administratively defined similar to other addresses. Individual service profiles are assigned a UUID from the UUID pool with embedded identifiers as follows for the infrastructure hosts.

UUID is comprises a prefix and a suffix.

VDI Solution: Infrastructure Hosts

Prefix

ssssbbbb-oooo-eeee

ssss	site ID = 0001
bbbb	block ID = dc01
oooo	ESXi version = 4101
eeee	cluster type, infra cluster = 0000

Suffix

tttt-ccccmmmmrrxx

tttt	server type, B200 M2 = b202
cccc	cpu type, Intel Xeon 5650 = 5650
mmmm	memory size, 96 GB = 0096
rr	core quantity, 6 cores = 06
xx	Host portion of the address assigned by the UCS

The following table lists the UUID pool used by the service profiles for infrastructure hosts.

https://t.me/learningnets

UUID Pool for Infrastructure Hosts

Pool	Quantity	UUID prefix	UUID suffix
uplnfra	2	0001dc01-4101-0000	b202-5650009606xx

VDI Solution: VDI Hosts

Prefix

ssssbbbb-oooo-eeee

ssss site ID = 0001

bbbb block ID = dc01

oooo ESXi version = 4101

eeee cluster type, VDI cluster = 0001

Suffix

tttt-ccccmmmmrrxx

tttt server type, B200 M2 = b202

cccc cpu type, Intel Xeon 5650 = 5650

mmmm memory size, 96 GB = 0096

rr core quantity, 6 cores = 06

xx Host portion of the address assigned by the UCS

The following table lists the UUID pool used by the service profiles for VDI hosts.

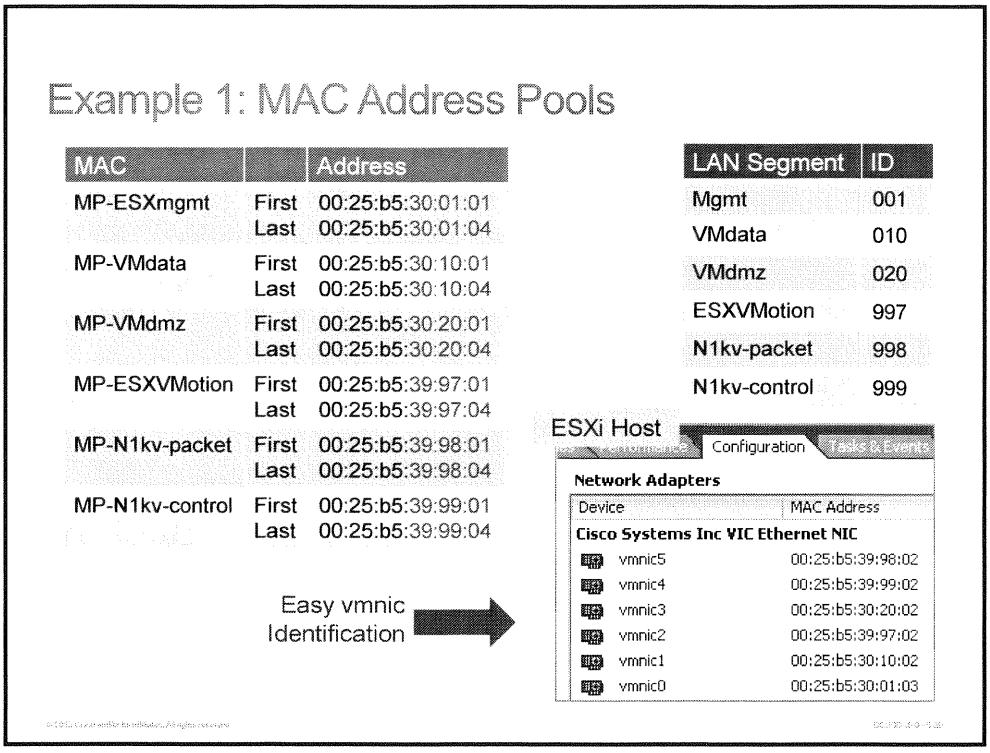
UUID Pool for VDI Hosts

Pool	Quantity	UUID prefix	UUID suffix
upVDI	20	0001dc01-4101-0001	b202-5650009606xx

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

Design MAC Addressing

This topic describes how to define the MAC addressing for given solutions.



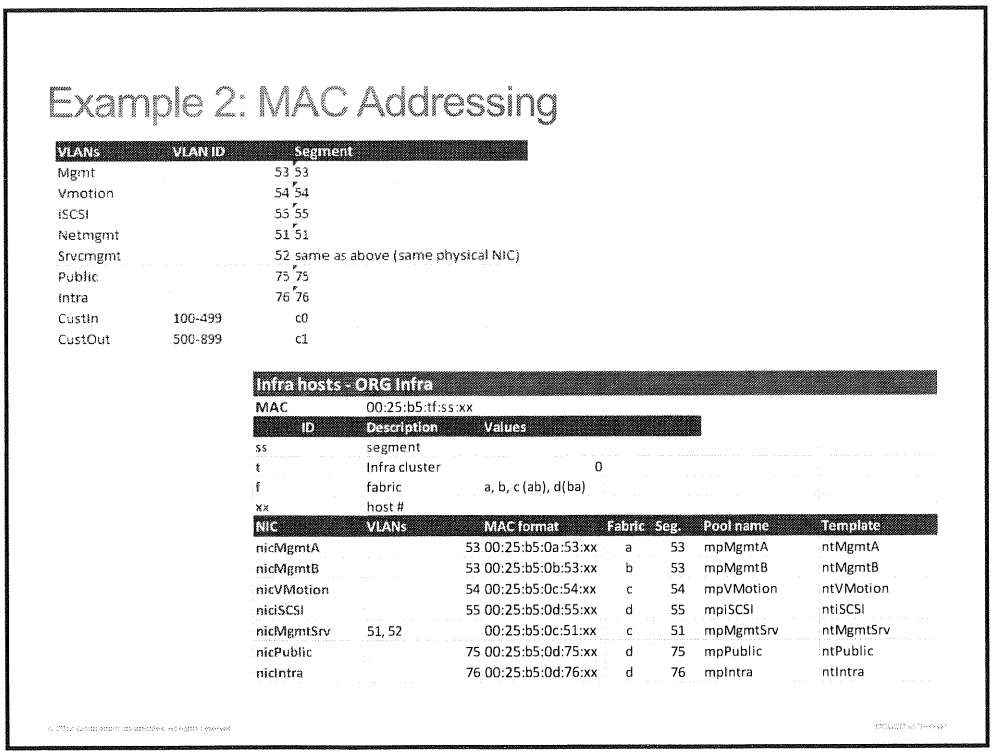
MAC address pools are designed for each NIC that will be available to the vSphere host. It has been determined that individual hosts will have six NICs, meaning six MAC pools are required.

- **MP-ESXmgmt pool:** MAC addresses that are used by the NIC that is connected to the management segment. Encodes VLAN 1, which is used for the management segment.
- **MP-VMdata pool:** MAC addresses that are used by the NIC that is connected to the VM data segment, which combines the VMdata1 and VMdata2 subsegments (VLANs 11 and 12). Thus, the encoded field is 10 (the least common denominator).
- **MP-VMdmz pool:** MAC addresses that are used by the NIC that is connected to the VM DMZ segment, which combines the VMdmz1 and VMdmz2 subsegments (VLANs 21 and 22). As with the VM data pool, this one encodes field value 20, which is the least common denominator for the given segments.
- **MP-ESXVMotion pool:** MAC addresses that are used by the NIC that is connected to the VMotion vSphere segment, which is utilized by the VM migration as well as advanced functionalities like high availability, fault tolerance, DRS, and DPM.

You can see from the tables in the figure that the MAC addresses in an individual pool are created based on the MAC address template that was specified earlier.

Notice the NIC appearance in the vSphere vCenter configuration. There is no similarity between the name of the vmnic (the vSphere host physical NIC) and the NIC defined by Cisco UCS. However, because the MAC addresses have encoded fields, it is easy to determine which vmnic should be used for which segment. Thus, deployment is easier.

https://t.me/learningnets



The connectivity provided via vNICs is configured under the service profiles for each individual host.

vNICs are Cisco UCS objects configured within a service profile to describe how ESXi host is connected to the Ethernet network.

VDI Solution: Infrastructure Hosts

The infrastructure hosts use the following vNICs connected to the segments as listed in the following table.

Infrastructure Host vNICs and Segments

vNIC	Segment	Fabric	VLANs
nicIntra	Intra	A,B	Intra
nicMgmtA	Mgmt	A	Mgmt
nicMgmtB	Mgmt	B	Mgmt
nicMgmtSrv	MgmtSrv	A,B	Netmgmt, Srvcmgmt
nicPublic	Public	A,B	Public
nicVMotion	VMotion	A,B	VMotion
niciSCSI	iSCSI	A,B	iSCSI

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

To speed and ease the deployment of service profiles, the vNICs use MAC addresses acquired from MAC pools. MAC addresses incorporate identifiers and are constructed in the following manner:

00:25:b5:TF:SS:XX

- **T** Infrastructure cluster = 0
- **F** Fabric = a, b, c (ab), d(ba)
- **SS** Segment
- **XX** Host portion of the address assigned by the Cisco UCS—up to 255 different MAC addresses

Because individual segments can carry multiple VLANs, the following segment encoding is used:

LAN Segment Encoding

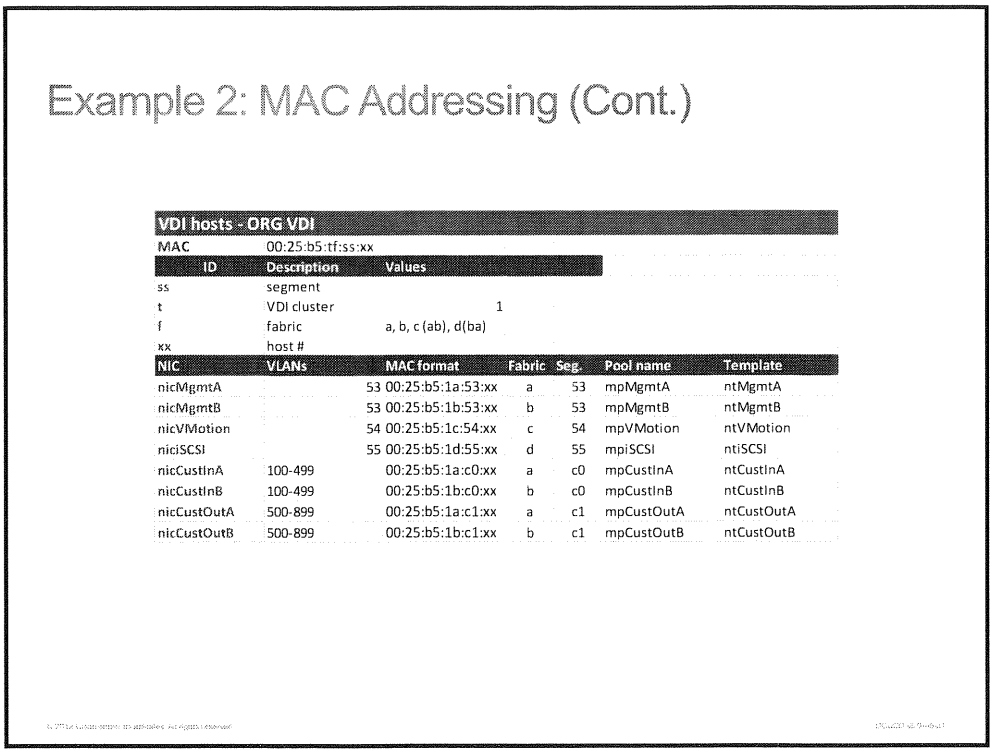
Segment	Segment ID	VLANs
Intra	76	Intra
Mgmt	54	Mgmt
MgmtSrv	51	Netmgmt, Srvcmgmt
Public	75	Public
VMotion	54	VMotion
iSCSI	55	iSCSI

The following table lists the MAC pools used by the vNICs of infrastructure hosts.

MAC Pools for vNICs of Infrastructure Hosts

Pool	Prefix	Qty	First MAC	Last MAC
mpMgmtA	00:25:b5:0a:53:xx	2	00:25:b5:0a:53:01	00:25:b5:0a:53:02
mpMgmtB	00:25:b5:0b:53:xx	2	00:25:b5:0b:53:01	00:25:b5:0b:53:02
mpVMotion	00:25:b5:0c:54:xx	2	00:25:b5:0c:54:01	00:25:b5:0c:54:02
mpiSCSI	00:25:b5:0d:55:xx	2	00:25:b5:0d:55:01	00:25:b5:0d:55:02
mpMgmtSrv	00:25:b5:0c:51:xx	2	00:25:b5:0c:51:01	00:25:b5:0c:51:02
mpPublic	00:25:b5:0d:75:xx	2	00:25:b5:0d:75:01	00:25:b5:0d:75:02
mpIntra	00:25:b5:0d:76:xx	2	00:25:b5:0d:76:01	00:25:b5:0d:76:02

https://t.me/learningnets



VDI Solution: VDI Hosts

The VDI hosts use the following vNICs connected to the segments as listed in the table.

VDI Hosts vNICs and Segments

vNIC	Segment	Fabric	VLANs
nicMgmtA	Mgmt	A	Mgmt
nicMgmtB	Mgmt	B	Mgmt
nicVMotion	VMotion	A,B	VMotion
niciSCSI	iSCSI	A,B	iSCSI
nicCustA	nicCust	A	CustIn1—CustIn250,
DCust1—DCust6			
nicCustB	nicCust	B	CustIn1—CustIn250,

To speed and ease the deployment of service profiles, the vNICs use MAC addresses acquired from MAC pools. MAC addresses incorporate identifiers and are constructed in the following manner:

00:25:b5:TF:SS:XX

- **T** VDI cluster = 1
- **F** Fabric = a, b, c (ab), d(ba)
- **SS** Segment
- **XX** Host portion of the address assigned by the Cisco UCS—up to 255 different MAC addresses

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

Because individual segments can carry multiple VLANs, the following segment encoding is used.

LAN Segment Encoding

Segment	Segment ID	VLANs
Mgmt	54	Mgmt
VMotion	54	VMotion
iSCSI	55	iSCSI
CustIn	c0	100-350, 401-406

The following table lists the MAC pools used by the vNICs of VDI hosts.

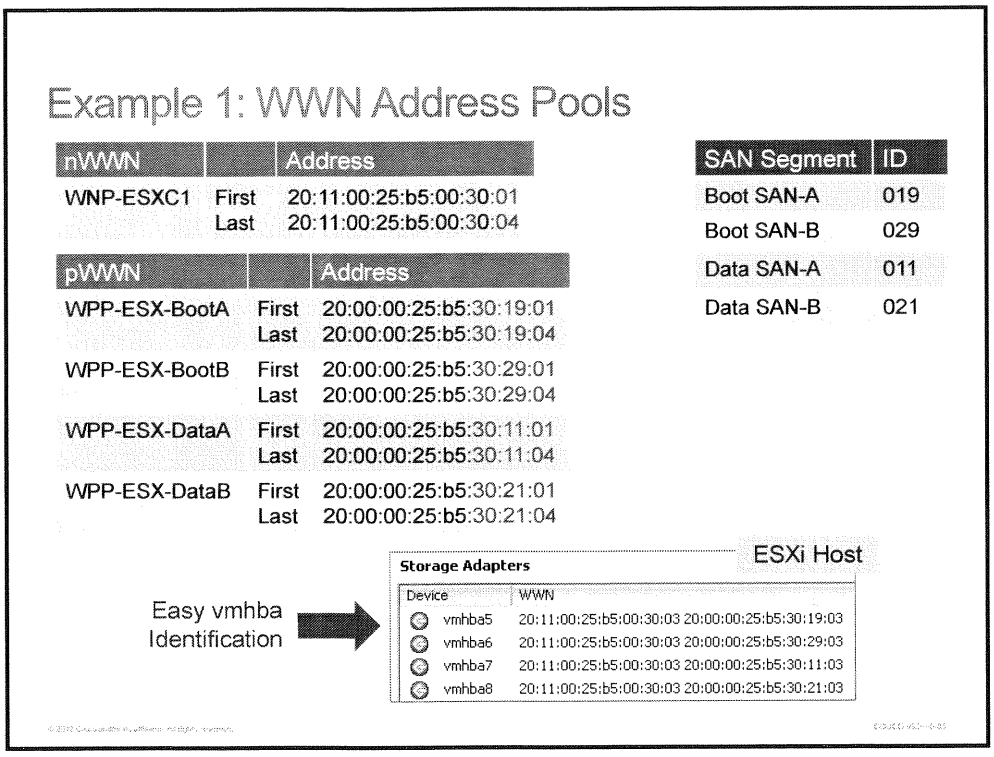
MAC Pools for vNICs of VDI Hosts

Pool	Prefix	Qty	First MAC	Last MAC
mpMgmtA	00:25:b5:1a:53:xx	20	00:25:b5:1a:53:01	00:25:b5:1a:53:14
mpMgmtB	00:25:b5:1b:53:xx	20	00:25:b5:1b:53:01	00:25:b5:1b:53:14
mpVMotion	00:25:b5:1c:54:xx	20	00:25:b5:1c:54:01	00:25:b5:1c:54:14
mpiSCSI	00:25:b5:1d:55:xx	20	00:25:b5:1d:55:01	00:25:b5:1d:55:14
mpCustInA	00:25:b5:1a:c0:xx	20	00:25:b5:1a:c0:01	00:25:b5:1a:c0:14
mpCustInB	00:25:b5:1b:c0:xx	20	00:25:b5:1b:c0:01	00:25:b5:1b:c0:14
mpMgmtA	00:25:b5:1a:53:xx	20	00:25:b5:1a:53:01	00:25:b5:1a:53:14

https://t.me/learningnets

Design WWN Addressing

This topic describes how to define the WWN addressing for given solutions.



nWWN pools are designed for each vSphere host that will be connected to the SAN-attached storage. Because there are only four hosts initially, the WNP-ESXC1 nWWN pool is designed with only four nWWN addresses. However, the addresses have an encoded operating system type (3 for vSphere hypervisor) and host ID. The VSAN segment is not encoded because the host has only one nWWN address in the fabric.

The pWWN address pools are designed for each HBA that will be available to the vSphere host. It has been determined that individual hosts will have four HBAs, so four pWWN pools are required.

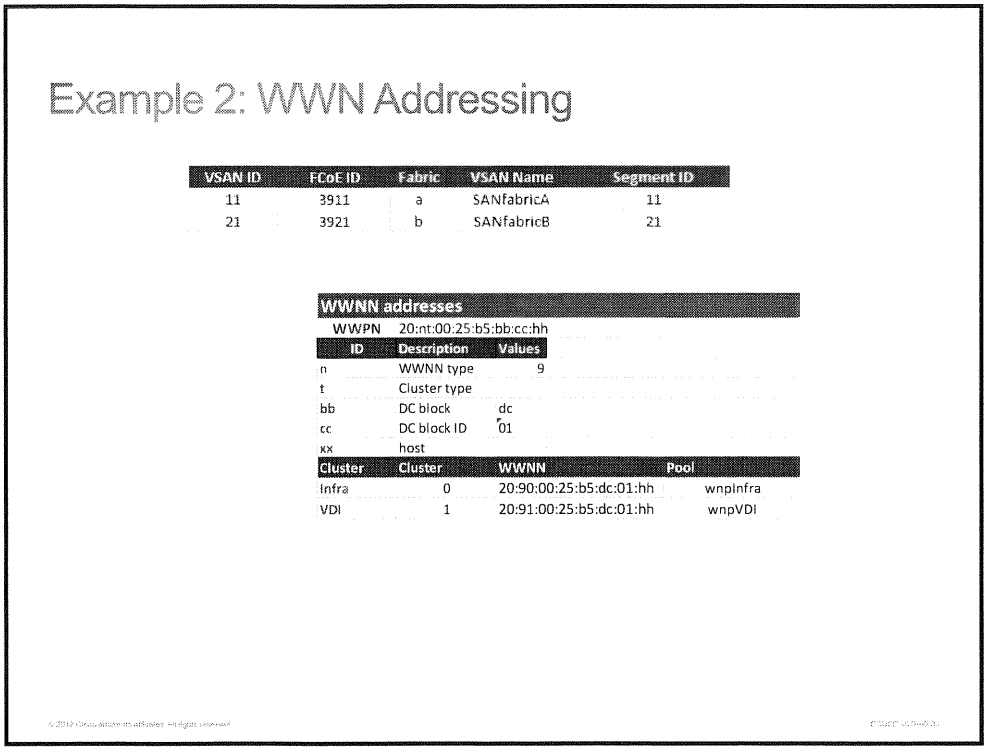
- **WPP-ESX-BootA pool:** pWWN addresses that are used by the HBA that is connected to the boot segment in fabric A. Uses field value 19 to encode VSAN 10 and identify the boot HBA.
- **WPP-ESX-BootB pool:** pWWN addresses that are used by the HBA that is connected to the boot segment in fabric B. Uses field value 29 to encode VSAN 20 and identify the boot HBA.
- **WPP-ESX-DataA pool:** pWWN addresses that are used by the HBA that is connected to the data segment in fabric A (the virtual machine file system [VMFS]). Uses field value 11 to encode VSAN 10 and identify the data HBA.
- **WPP-ESX-DataB pool:** pWWN addresses that are used by the HBA that is connected to the data segment in fabric B (the same VMFS as fabric A). Uses field value 21 to encode VSAN 20 and identify the boot HBA.

You can see from the tables in the figure that the pWWN addresses in an individual pool are created based on the pWWN address template that was specified earlier.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Notice the HBA appearance in the vSphere vCenter configuration. There is no similarity between the name of the vmhba (the vSphere host physical HBA) and the HBA defined by Cisco UCS. However, because the pWWN addresses have encoded fields, it is easy to determine which vmhba should be used for which segment. Thus, deployment is easier.



The connectivity provided via vHBAs is configured under the service profiles for each individual host. vHBAs are Cisco UCS objects configured within the service profile to describe how the ESXi host is connected to the Fibre Channel SAN network.

VDI Solution: Infrastructure Hosts

The infrastructure hosts are using the vHBAs connected to the segments as listed in the following table.

Infrastructure Host vHBAs and Segments

vHBA	Segment	Fabric	VSAN
hbaA0	SANFabricA	A	SANFabricA
hbaA1	SANFabricA	A	SANFabricA
hbaB0	SANFabricB	B	SANFabricB
hbaB1	SANFabricB	B	SANFabricB

To speed up and ease the deployment of service profiles, the vHBAs use WWN addresses acquired from WWN pools.

https://t.me/learningnets

Because individual segments can carry multiple VSANs, the following segment encoding is used:

SAN Segment Encoding

Segment	Segment ID	VLANs
SANFabricA	11	SANFabricA
SANFabricB	21	SANFabricB

nWWN Addresses

nWWN addresses incorporate identifiers and are constructed in the following manner:

20:NT:00:25:b5:BB:CC:HH

N WWN type—nWWN = 9

T Infrastructure cluster = 0

BB DC block = dc

CC DC block ID = 01

XX Host portion of the address assigned by the UCS

The following table lists the nWWN pools that are used by the service profiles for infrastructure hosts.

nWWN Pool for Infrastructure Hosts

Pool	Prefix	Qty	First WWN	Last WWN
wnpInfra	20:90:00:25:b5:dc:01:hh	2	20:90:00:25:b5:dc:01:01	20:90:00:25:b5:dc:01:02

pWWN Addresses

pWWN addresses incorporate identifiers and are constructed in the following manner:

20:NT:00:25:b5:FF:SS:HH

N WWN type—first HBA=0, second hba=1

T Infrastructure cluster = 0

FF fabric= 0A or 0B

SS Segment

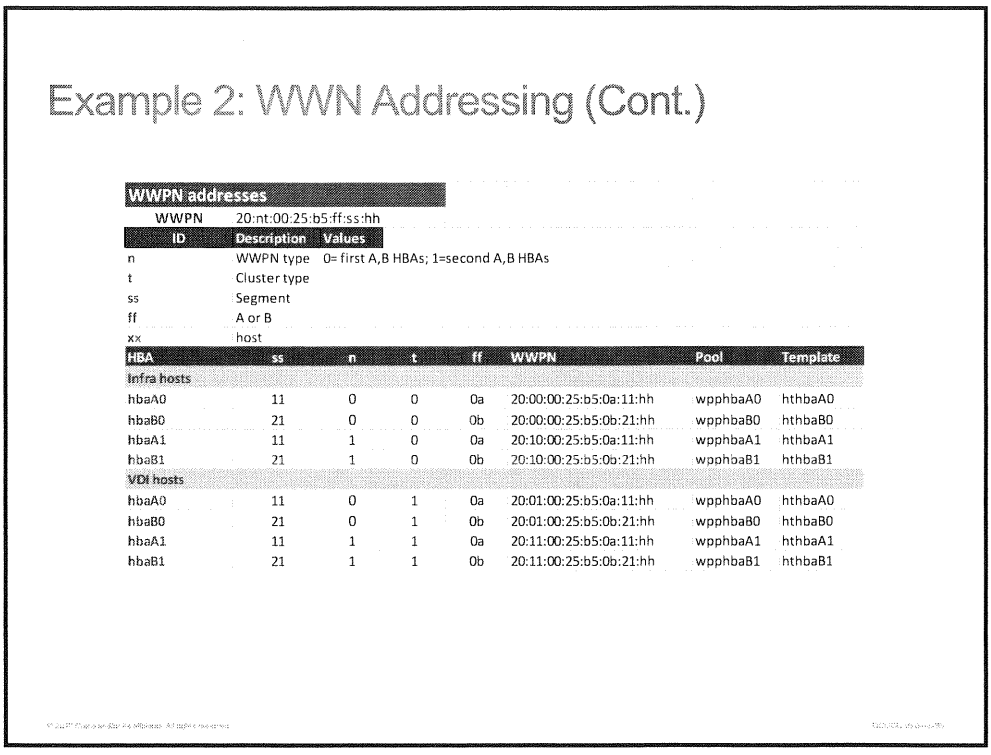
XX Host portion of the address assigned by the UCS

The following table lists the pWWN pools used by the vHBAs of infrastructure hosts.

pWWN Pools for Infrastructure Hosts

Pool	Prefix	Qty	First WWN	Last WWN
wpphbaA0	20:00:00:25:b5:0a:11:hh	2	20:00:00:25:b5:0a:11:01	20:00:00:25:b5:0a:11:02
wpphbaA1	20:10:00:25:b5:0a:11:hh	2	20:10:00:25:b5:0a:11:01	20:10:00:25:b5:0a:11:02
wpphbaB0	20:00:00:25:b5:0b:21:hh	2	20:00:00:25:b5:0b:21:01	20:00:00:25:b5:0b:21:02
wpphbaB1	20:10:00:25:b5:0b:21:hh	2	20:10:00:25:b5:0b:21:01	20:10:00:25:b5:0b:21:02

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



The Connectivity VDI Solution: VDI Hosts

The VDI hosts are using vHBAs connected to the segments as listed in the following table.

VDI Hosts vHBAs and Segments

vHBA	Segment	Fabric	VSAN
hbaA0	SANFabricA	A	SANFabricA
hbaA1	SANFabricA	A	SANFabricA
hbaB0	SANFabricB	B	SANFabricB
hbaB1	SANFabricB	B	SANFabricB

To speed and ease the deployment of service profiles, the vHBAs use WWN addresses acquired from WWN pools.

Because individual segments can carry multiple VSANs, the following segment encoding is used.

SAN Segment Encoding

Segment	Segment ID	VLANs
SANFabricA	11	SANFabricA
SANFabricB	21	SANFabricB

https://t.me/learningnets

nWWN Addresses

nWWN addresses incorporate identifiers and are constructed in the following manner:

20:NT:00:25:b5:BB:CC:HH

N WWN type—nWWN = 9

T VDI cluster = 1

BB DC block = dc

CC DC block ID = 01

XX Host portion of the address assigned by the Cisco UCS

The following table lists the WWNN pools used by the service profiles for VDI hosts.

Pool	Prefix	Qty	First WWN	Last WWN
wnpVDI	20:91:00:25:b5:dc:01:hh	20	20:91:00:25:b5:dc:01:01	20:91:00:25:b5:dc:01:14

pWWN Addresses

pWWN addresses incorporate identifiers and are constructed in the following manner:

20:NT:00:25:b5:FF:SS:HH

N WWN type—first HBA=0, second hba=1

T VDI cluster = 1

FF fabric= 0A or 0B

SS Segment

XX Host portion of the address assigned by the Cisco UCS

The following table lists the pWWN pools used by the vHBAs of VDI hosts.

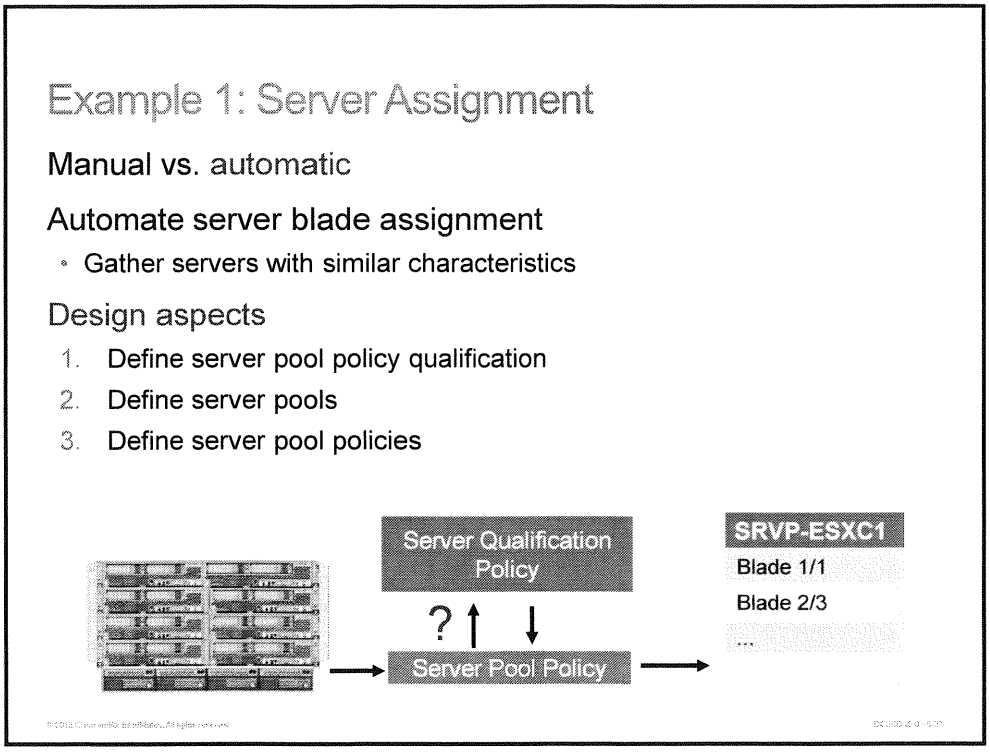
pWWN Pools for VDI Hosts

Pool	Prefix	Qty	First WWN	Last WWN
wpphbaA0	20:01:00:25:b5:0a:11:hh	20	20:01:00:25:b5:0a:11:01	20:01:00:25:b5:0a:11:14
wpphbaA1	20:01:00:25:b5:0b:21:hh	20	20:01:00:25:b5:0b:21:01	20:01:00:25:b5:0b:21:14
wpphbaB0	20:11:00:25:b5:0a:11:hh	20	20:11:00:25:b5:0a:11:01	20:11:00:25:b5:0a:11:14
wpphbaB1	20:11:00:25:b5:0b:21:hh	20	20:11:00:25:b5:0b:21:01	20:11:00:25:b5:0b:21:14

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

Design Common Policies

This topic describes how to define the common policies for hosts of a given solution.

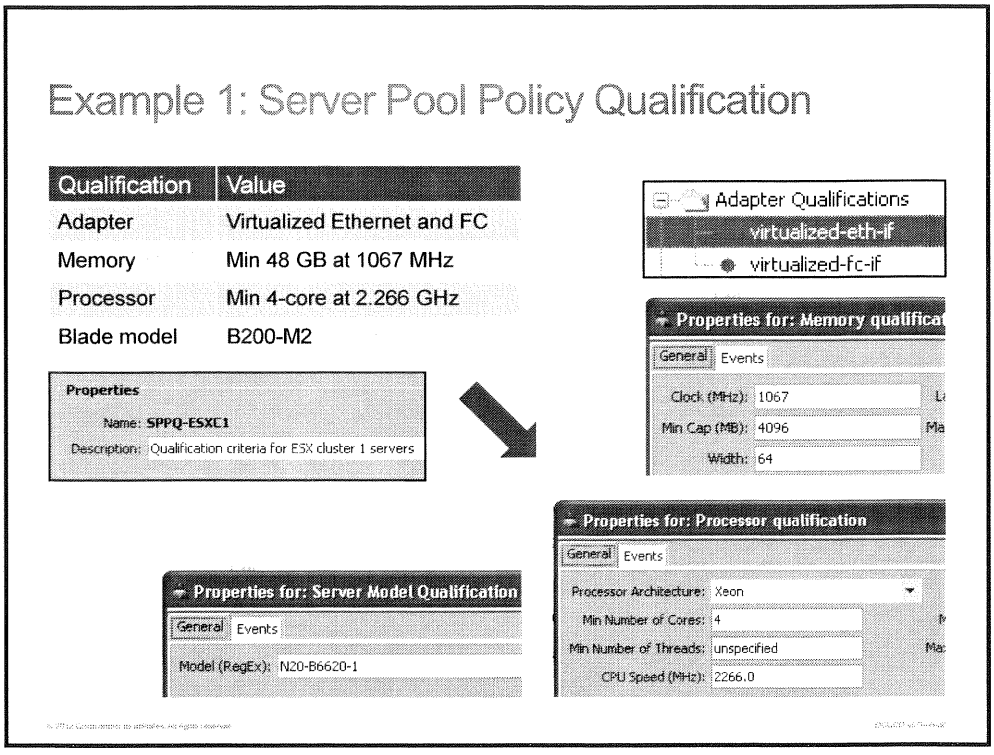


vSphere hosts in this example have minimum requirements for the server hardware. When the service profile is created, the blade can be assigned manually or automatically from the pool. In this example, because the minimum hardware characteristics are identified, the automatic assignment with policy enforcement can be used—that is, the server blade will be automatically assigned from the pool to the service profile upon creation.

For this purpose, the design will specify the following objects upon which the server blade will be automatically assigned to the service profile upon creation:

- **Server pool policy qualification:** This object specifies the hardware characteristics against which server blades are matched when discovered by the system.
- **Server pool:** This object groups the blades of similar characteristics (in this example).
- **Server pool policy:** This object binds together qualification policy and server pool, meaning it populates the server pool with blades matched against criteria that are defined by the qualification policy.

https://t.me/learningnets



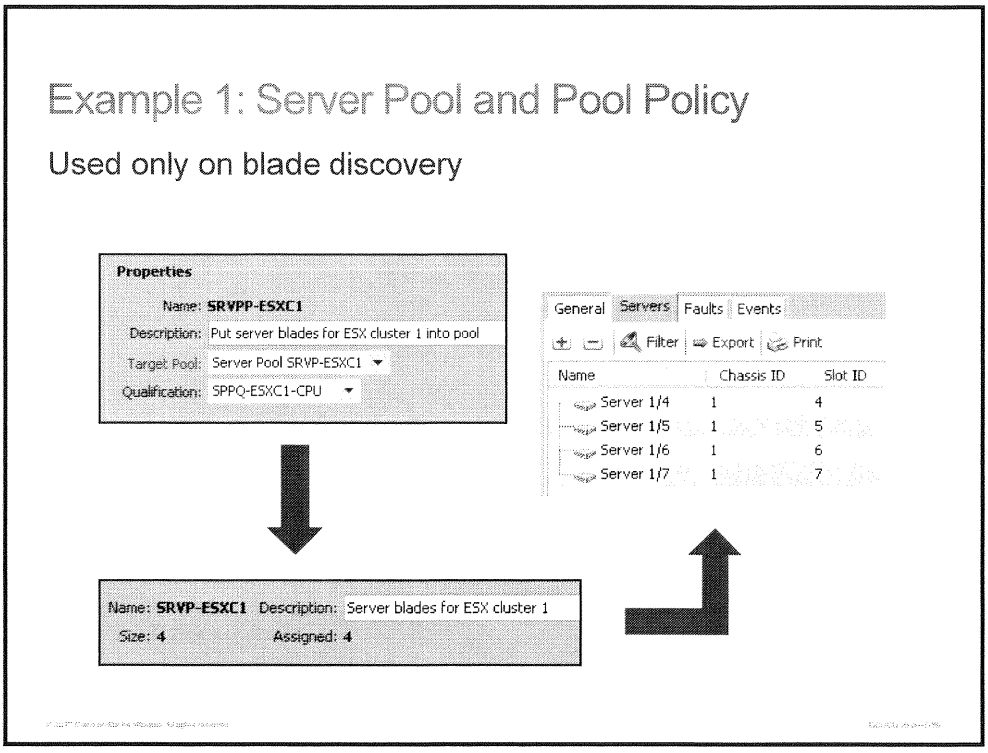
In this example, the qualification policy criteria are as follows:

- **Adapter** that must be virtualized with Ethernet and Fibre Channel support
- **Memory** that is at least 48 GB in size and at least running at 1067 MHz
- **CPU** with a minimum of four cores running at 2.266 GHz
- **Blade size** should be half-width

The name of the qualification policy for the vSphere cluster hosts is **SPPQ-ESXC1.**

© 2012 Cisco Systems, Inc.

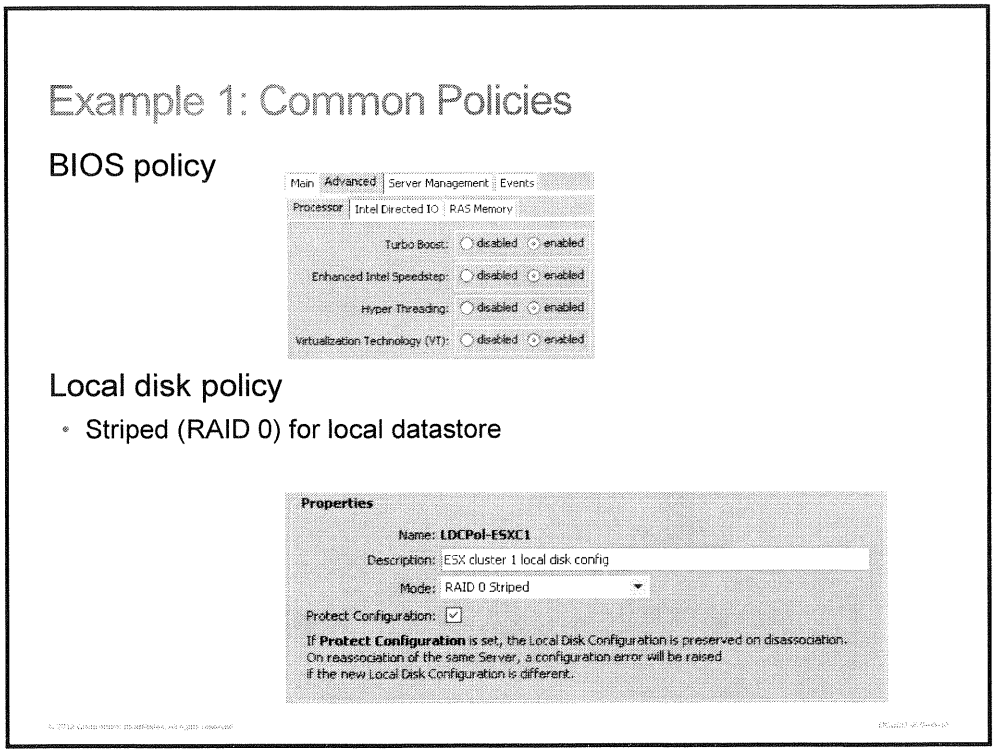
https://t.me/learningnets



The server pool that will be used by service profiles for auto blade assignment will be named SRVP-ESXC1 and will be bound to the SRVPP-ESXC1 pool policy that combines it with the qualification policy discussed earlier.

The design has to include a note that the automatic pool population is done only upon server blade discovery. In other words, if the server blades are already present, but the pool, pool policy, and qualification policy are defined later, the implementation engineer should start blade rediscovery for all the blades that are not used. This way, the pool is populated.

https://t.me/learningnets



Apart from pools and related policies, there are also other policies that need to be defined.

The first one is the BIOS policy (or BIOS defaults for the entire system). Once it is defined, it ensures that the necessary settings are applied (for example, Intel Virtualization Technology [VT] bit).

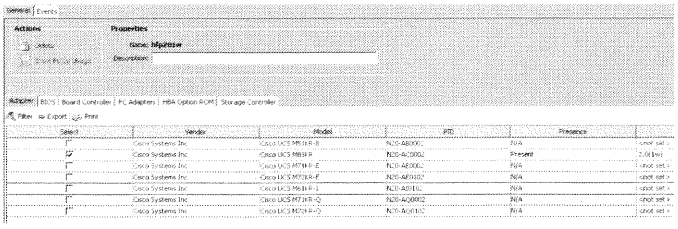
The second is the local storage policy, which requires a local disk to be used as a local datastore in VMware and to use the entire space. Thus, the LDPol-ESXC1 is defined with the Redundant Array of Independent Disks (RAID) stripe disk configuration option.

© 2012 Cisco Systems, Inc.

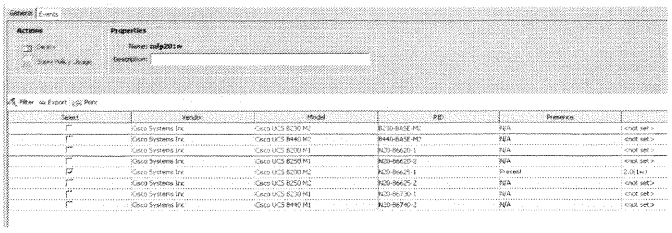
https://t.me/learningnets

Example 1: Common Policies (Cont.)

Host firmware package – adapter firmware = 2.0(1w)



Management firmware package = 2.0(1w)

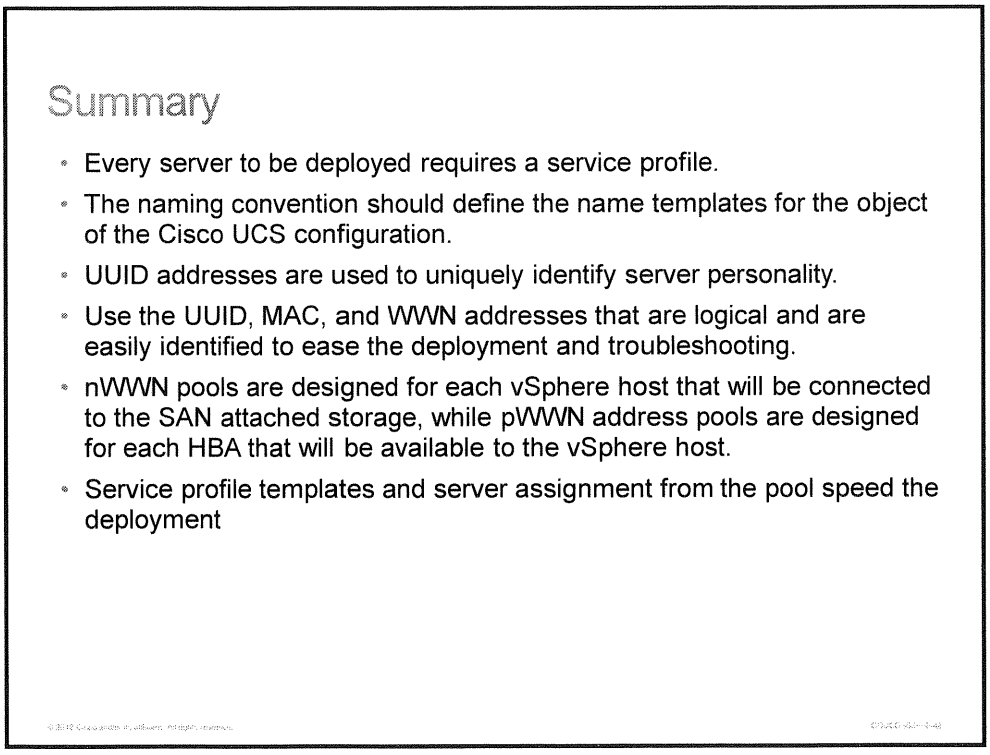


Likewise, the host and management firmware packages are defined to ensure that the vSphere host service profiles (and, consequently, the blades) will use proper firmware for the adapters—specifically for the HBAs. This is important for the drive array vendor certification and support.

https://t.me/learningnets

Summary

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



© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Lesson 2

Designing Unified Computing Management

Overview

This lesson discusses management aspects of the given Cisco Unified Computing System (UCS) solution.

Objectives

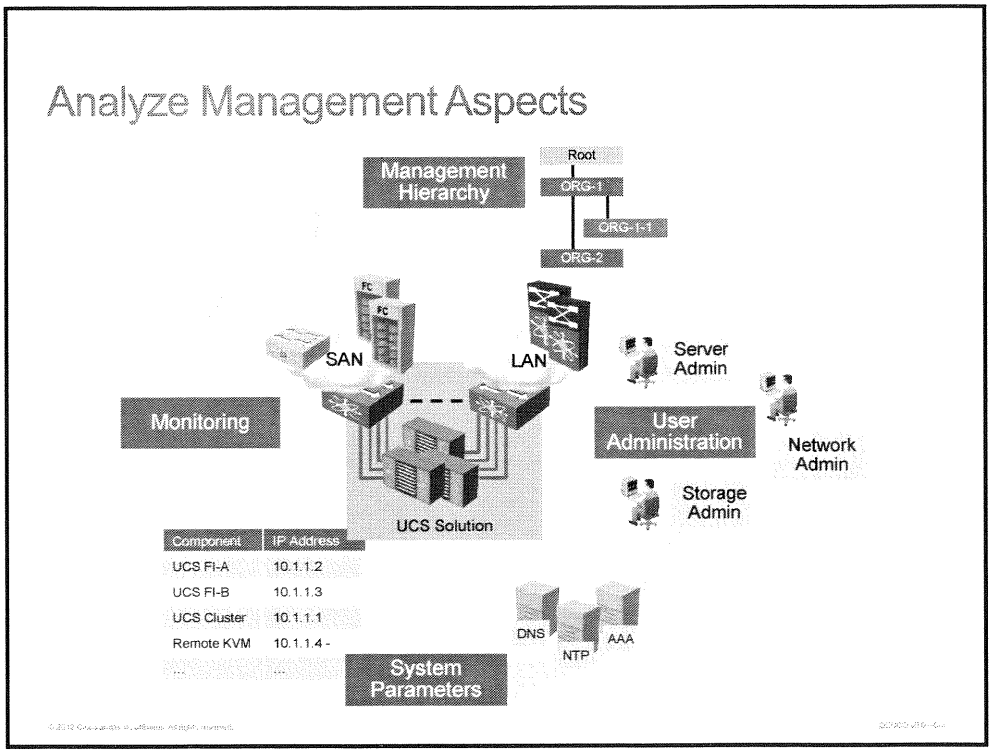
Upon completing this lesson, you will be able to design management aspects of the Cisco UCS solution. You will be able to meet these objectives:

- Define the Cisco UCS management access
- Define the organizational hierarchy within Cisco UCS configuration
- Define the remote management connectivity characteristics

https://t.me/learningnets

Design Management Access

This topic describes how to define the Cisco UCS management access.



The management aspects of Cisco UCS consist of:

- Management hierarchy within organizations
- User management access
- System-related policies and parameters
- Monitoring policies

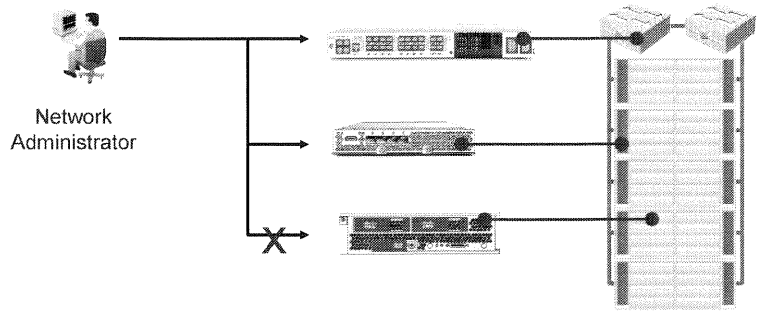
The first three are necessary for Cisco UCS deployment and need to be agreed upon by the customer, especially when the equipment is to be deployed in the existing environment.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

RBAC Overview

- RBAC = method of authorizing user system access
  - Based on user roles and privileges
  - Ability to grant privileges based on user responsibilities
  - Delegates management rights
- Does not control access to server-deployed operating system

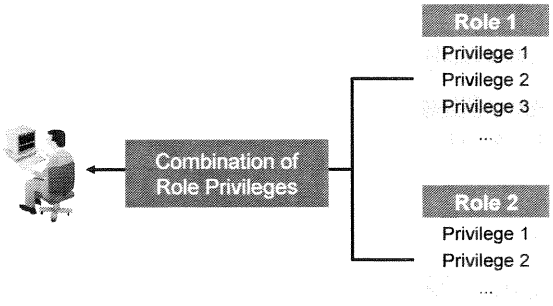


Role-based access control (RBAC) is a method of restricting or authorizing system access for users based on user roles and locales.

https://t.me/learningnets

Roles and Privileges

- Privilege—individual rule
- Role—collection of privileges
  - Grants read-only or read/write access
- User
  - Assigned one or more roles
  - Has superset of privileges of assigned roles



A role defines the privileges of a user in the system. Privileges give users that are assigned to user roles access to specific system resources and permission to perform specific tasks. There are default privileges that are assigned to user roles in Cisco UCS.

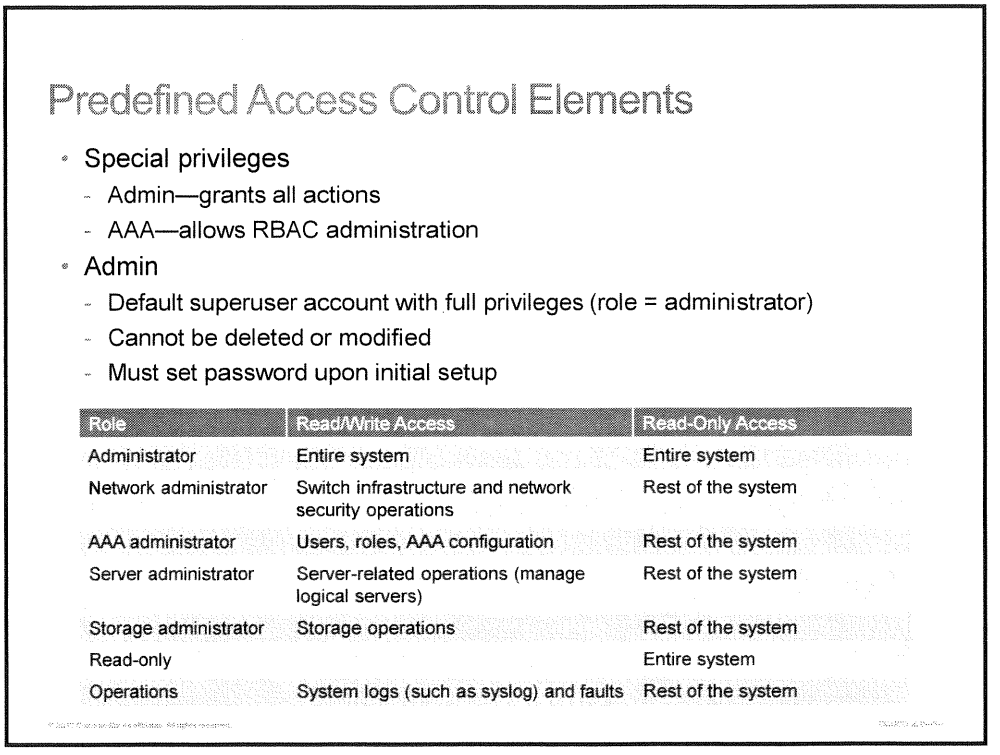
Roles can be created, modified to add new privileges or remove existing privileges, or deleted. When a role is modified, the new privileges are applied to all users assigned to that role. Privilege assignment is not restricted to the privileges defined for the default roles, meaning you can use a custom set of privileges to create a unique role.

A user can be assigned to one or more roles. A user assigned to multiple roles has the combined privileges of all assigned roles. For example, if Role 1 has storage-related privileges and Role 2 has server-related privileges, then users who are assigned to both Role 1 and Role 2 have both storage- and server-related privileges.

All roles include "read" access to all configurations on the system. The difference between the read-only role and other roles is that a user who is assigned to only the read-only role cannot modify the system state. A user assigned to another role can modify the system state in an assigned area or areas specific to that user.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets



There is a predefined set of privileges within the Cisco UCS solution. The following two privileges grant certain special actions:

- **Admin:** A user-assigned role with this privilege can perform any action in Cisco UCS within any organization.
- **Authentication, authorization, and accounting (AAA):** A user-assigned role with this privilege can perform administration of the RBAC features.

Apart from predefined privileges, the system also includes predefined roles:

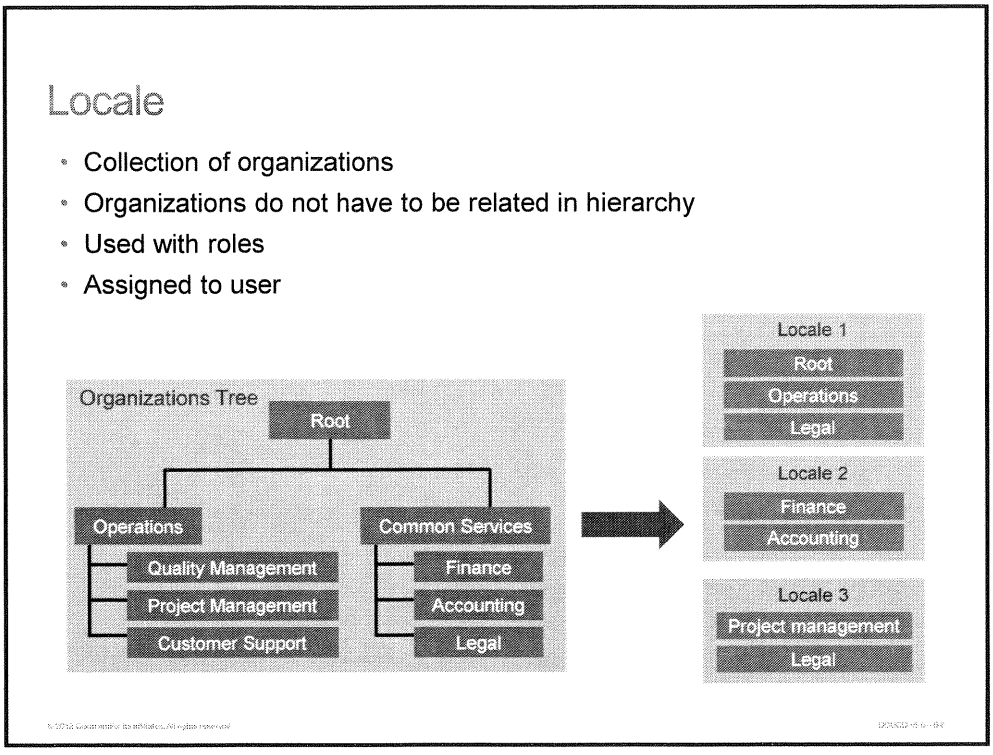
- **Administrator:** Enables complete read/write access to the entire system.
- **Network administrator:** Enables read/write access to switch infrastructure and network security operations. Enables read-only access to the rest of the system.
- **AAA administrator:** Enables read/write access to users, roles, and AAA configuration. Enables read-only access to the rest of the system.
- **Server administrator:** Enables read/write access to server-related operations. A user with this role has privileges to manage logical servers. Enables read-only access to the rest of the system.
- **Storage administrator:** Enables read/write access to storage operations. Enables read-only access to the rest of the system.

https://t.me/learningnets

- **Read-only:** Enables read-only access to system configuration but with no privileges to modify the system state.
- **Operations:** Enables read/write access to system logs, including the syslog servers, and faults. Enables read-only access to the rest of the system.

Note The default administrator account is assigned the administrator role by default, and this assignment cannot be changed.

Note If a role is deleted after it has been assigned to users, all users with that role are assigned the default read-only role.



Locale defines the organizations (domains) that a user is allowed to access.

A user can be assigned one or more locales. Each locale defines one or more organizations that the user is allowed to access. Access is limited to the organizations that are specified in the locale.

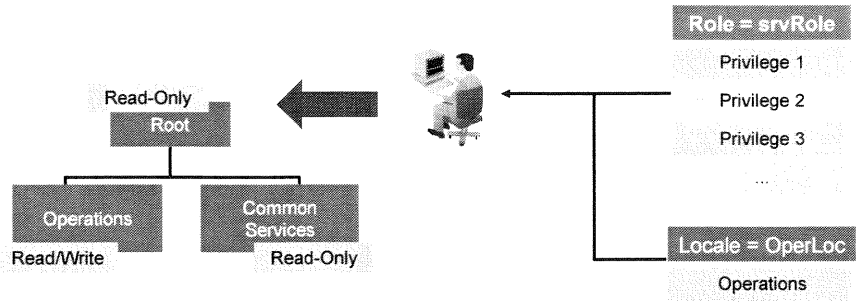
An exception to this rule is a locale without any organizations, which allows unrestricted access to system resources in all organizations.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

Using RBAC

- User is assigned role-based privileges upon login.
- Privilege sets for all roles are applied to the user-authorized organizations.

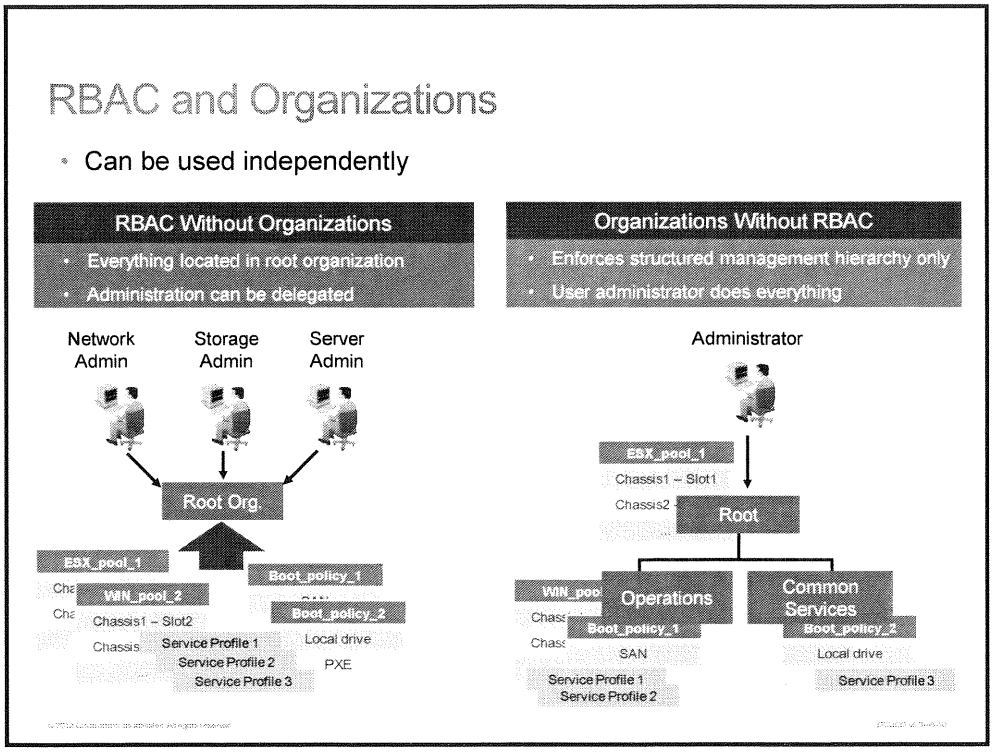


Users are not directly assigned privileges. Rather, management of individual user privileges is simply a matter of assigning the appropriate roles and locales.

A user is granted write access to desired system resources only if the assigned role grants the access privileges and the assigned locale allows access.

For example, a user with the srvRole role in the Operations organization could update server configurations in the Operations organization, but that user would not be able to update server configurations in the Common Services organization unless the locale for the srvRole role also included that organization.

https://t.me/learningnets



RBAC and organizations can be used independently. It is not necessary to use RBAC if you are managing a hierarchical set of organizations and vice versa.

Using RBAC Without Organizations

In cases where RBAC is used without organizations—when everything is in the root organization—the advantages of the RBAC style of delegated management can still be used. Users can be segregated based on their responsibilities. For example, there can be users that have control over the logical server structure and other users that have control over the border (LAN and SAN) configuration.

Using Organizations Without RBAC

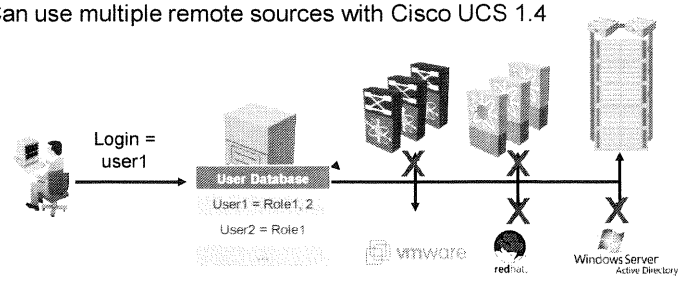
In cases where organizations are used without RBAC, the organization structural management hierarchy can be utilized to better manage large amounts of equipment and large numbers of logical objects. In addition, some of the pool inheritance features can scale the design logic.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

User Authentication

- User authentication used in conjunction with RBAC
  - Applies administrative control
- Centralized authentication
  - Simplifies administration of multiple RBAC schemes
  - Cisco UCS, Nexus 7000, MDS 9000, etc.
- Cisco UCS authentication methods
  - Local to Cisco UCS
  - Remote via protocol (LDAP, RADIUS, TACACS+)
    - Can use multiple remote sources with Cisco UCS 1.4



User authentication is used to apply administrative control. It can be applied with RBAC. When a user logs into the Cisco UCS Manager, he or she is authenticated against the appropriate database and assigned privileges that are based on his or her roles.

Centralized authentication is typically used to simplify administration of multiple RBAC schemes.

Cisco UCS supports two methods to authenticate user logins:

- Local authentication with a user database deployed locally to the Cisco UCS Manager
- Remote authentication by using one of the following protocols:
  - Lightweight Directory Access Protocol (LDAP)
  - RADIUS
  - TACACS+

The operation of the user authentication occurs as follows:

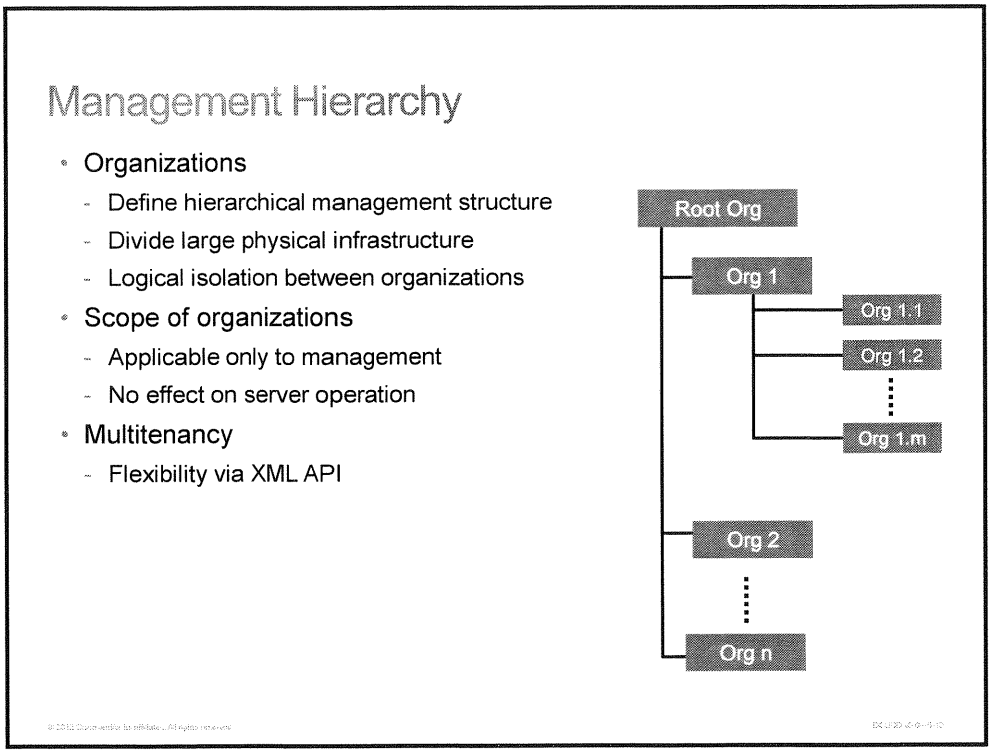
- User logs into the Cisco UCS system.
- The Cisco UCS system queries the authentication server for the user approval.
- The authentication server grants or denies user access based on these credentials. If the user is granted access, the authentication server passes the approved role or roles to the Cisco UCS system.
- Cisco UCS grants or denies user access based on the authentication server response.

Cisco UCS version 1.4 introduces several enhancements, among which is the ability to define multiple remote sources for the authentication.

https://t.me/learningnets

Design Configuration Hierarchy

This topic describes how to define the organizational hierarchy within the Cisco UCS configuration.



Organizations

Multitenancy allows you to divide the large physical infrastructure of an instance into logical entities that are known as organizations. You can achieve a logical isolation between organizations without providing a dedicated physical infrastructure for each organization.

The designer must (with the help of the customer) analyze and define the organizational structure—the hierarchy that will be implemented in Cisco UCS.

If you set up a multitenant environment, all organizations are hierarchical. The top-level organization is always the root. The policies and pools that you create in the root are systemwide and are available to all organizations in the system. However, any policies and pools that are created in other organizations are only available to organizations that are above the organization in which the policies or pools have been created.

Administrative Access

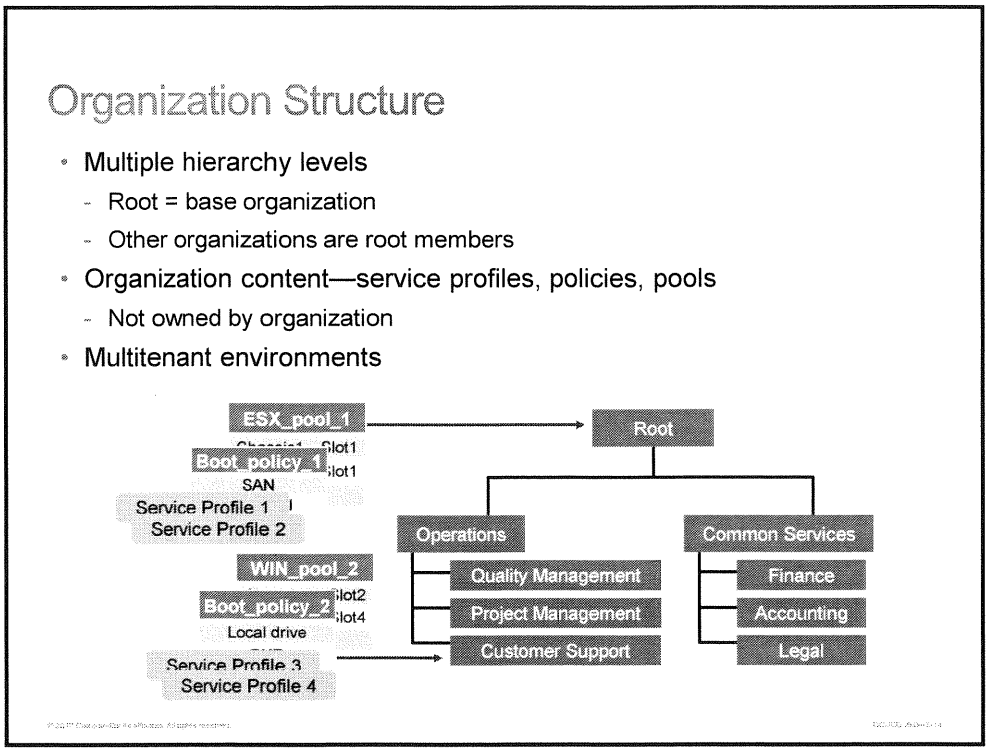
You can assign unique resources to each tenant through the related organization in the multitenant environment. These resources can include different policies, pools, and quality of service (QoS) definitions. You can also implement locales to assign or restrict user privileges and roles to organizations if you do not want all users to have access to all organizations.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Cisco UCS supports two methods to authenticate user logins:

- Local database
- Remote authentication using one of these protocols:
  - LDAP
  - RADIUS
  - TACACS+



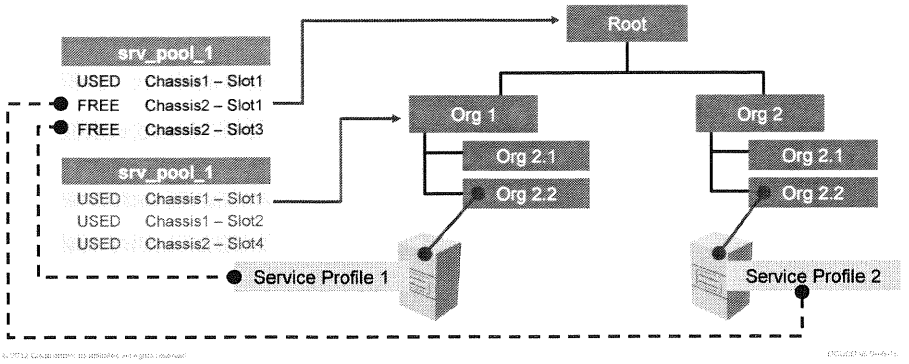
Organizations are hierarchically organized into a tree, where the top-level organization is always the root. Cisco UCS organizations enable an administrator to assign unique resources via related organizations. The resources can include policies, pools, and service profiles and templates.

The policies and pools that you create in the root are systemwide and available to all organizations in the system. However, any policies and pools that are created in other organizations are available only to organizations that are above them in the same hierarchy. For example, if a system has organizations named "Finance" and "HR" that are not in the same hierarchy, Finance cannot use any policies in the HR organization, and HR cannot access any policies in the Finance organization. However, both Finance and HR can use policies and pools in the root organization.

https://t.me/learningnets

Organization Pool Resolution

- Service profile associated with own or parent organization pools
- Hierarchical "bottom to top" tree resolution
  - Search tree for available resource in pools with identical name
  - "First come, first serve" principle
- All pool types (blade server, MAC, WWN, UUID)



Organizations use a hierarchical pool resolution method, meaning a pool is resolved to a service profile from the bottom to the top of the organization tree. This approach is used for any pool type—whether a blade server, MAC, world wide name (WWN), or Universally Unique Identifier (UUID) pool.

The resolution process is performed in the following way:

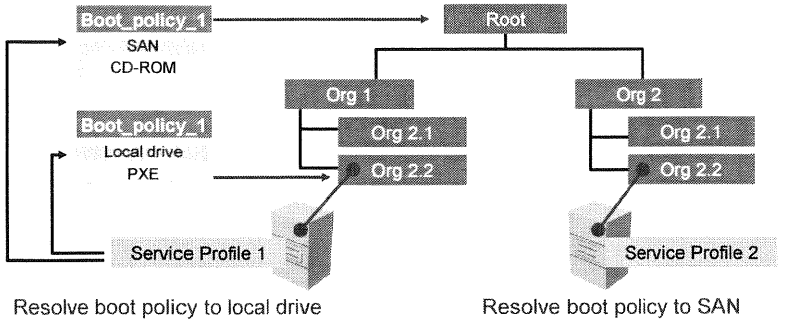
- Cisco UCS Manager checks within the organization for pools with the specified name that are assigned to the service profile or policy.
- If an available resource is found inside a pool, Cisco UCS Manager uses that resource.
- If the pool does not have any available resources at the local level, Cisco UCS Manager moves up in the hierarchy to the parent organization and searches for a pool with the same name. This is repeated until the search reaches the root organization.
- If the search reaches the root organization and has not found an available resource, Cisco UCS Manager returns to the local organization and begins to search for a default policy or available resource within the default pool.
- If an available resource in a default pool is found, it is used. If the pool does not have any available resources, Cisco UCS Manager moves up in the hierarchy to the parent organization and searches for a default pool. Again, this step is repeated until the search reaches the root organization.
- If Cisco UCS Manager cannot find an available resource in the hierarchy, it returns an allocation error.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Organization Policy Resolution

- Service profile associated with own or parent organization policies
  - Only policies with different names are applicable
- Hierarchical "bottom to top" tree resolution
  - "First match" policy principle
- All policy types (boot, firmware, etc.)



Organizations use a hierarchical policy resolution method. In other words, a policy is resolved to a service profile from the bottom to the top of the organization tree. If there are policies of the same type using the same name but different settings, only the policy from the same organization level can be applied to a service profile.

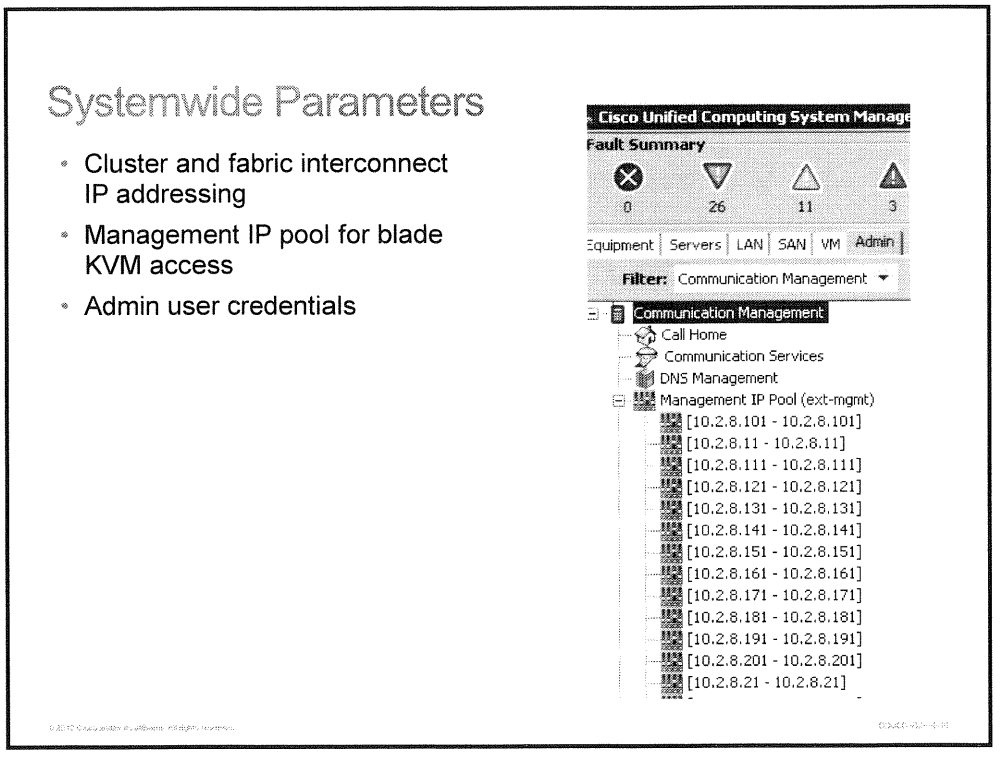
Alternately, a service profile on a certain organization level can be associated with policies from different, higher organizational levels as long as the policies use different names.

The resolution process is similar to the one used for pools, with the difference being that the policy used is the first one in the organizational tree that matches the assigned policy name, starting from the same level as the service profile. If no policy is available, the default policy is used.

https://t.me/learningnets

Design Remote Management Connectivity

This topic describes how to define the remote management connectivity characteristics.



Management IP Pool

Each Cisco UCS cluster also requires a pool of IP addresses that are used for management access to the individual server blade. An individual Cisco UCS cluster can have up to 320 server blades, each requiring an IP address, a cluster IP address, and the addresses of the fabric interconnect switches A and B. Thus, a maximum of 323 IP addresses are required for a cluster.

System Parameters

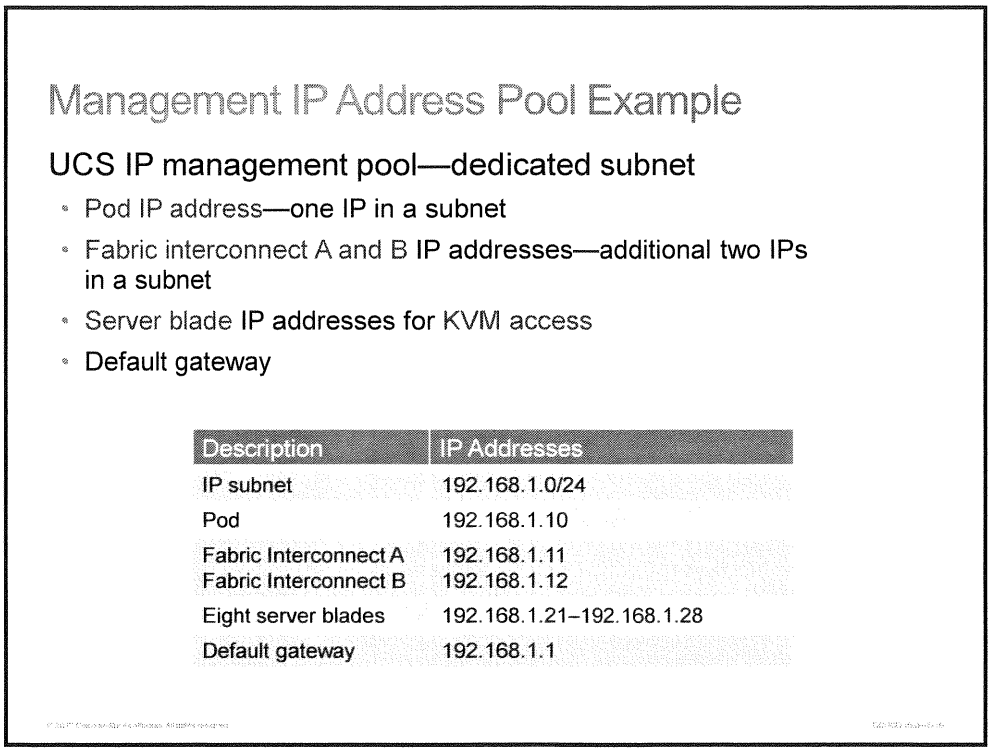
Cisco UCS provides several diagnostic tools to aid in troubleshooting and monitoring the environment. These tools include command-line debug statements, syslog, and Simple Network Management Protocol (SNMP). (Call Home is another diagnostic and reporting tool.) Syslog is the mechanism used for processes and scripts to write log entries. The callers can fully specify all of the characteristics of the log entries.

Network Time Protocol (NTP) is important for configuration and maintenance.

Cisco UCS can be accessed via different management interfaces, including Secure Shell (SSH), HTTPS, and Telnet. The design must specify which interface should be enabled and used.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets



During the Cisco UCS sizing process, the following quantities and server classes that you will need to deploy have been established:

- 18 PS-class1 server blades requiring redundant LAN and SAN connectivity
- 17 PS-class2 server blades requiring redundant LAN and SAN connectivity, with SAN boot functionality
- 15 VS-class1 server blades requiring redundant LAN and SAN connectivity, with twice the number of LAN interfaces

Because the servers will be deployed in a single data center, and there is no special requirement that the servers need to be identified by the MAC or WWN addresses, you can create a simple naming convention and pools.

For the site ID, you will use 01 in case the implementation increases in the future.

The name of the Cisco UCS cluster will be set to sanjose_01, where 01 stands for the first Cisco UCS cluster in the San Jose data center. This allows for the flexibility of introducing new Cisco UCS clusters, if needed, while still being able to identify the cluster from the name alone.

For the management IP pool, you need 53 IP addresses in all:

- 18+17+15 = 50 IP addresses for the server blades
- Three IP addresses for the Cisco UCS fabric interconnects and cluster IP

The 192.168.1.0/24 subnet will be used, which is wide enough to allow future growth. From this management IP address pool, the IP addresses will be given in the following fashion:

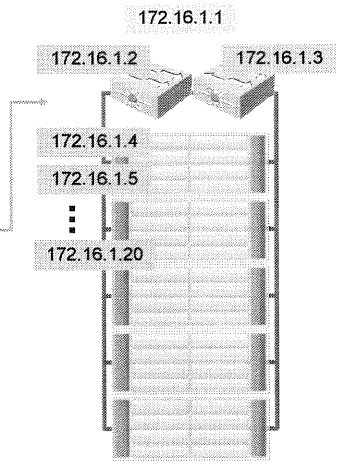
- **Cluster IP:** 192.168.1.10
- **Switch A:** 192.168.1.11
- **Switch B**: 192.168.1.12
- **Blade management IP addresses:** 192.168.1.21–192.168.1.28
- **Gateway for the segment:** 192.168.1.1

https://t.me/learningnets

Management IP Address Pool Convention

- Per cluster IP management pool—dedicated subnet
  - Cluster IP address—first in subnet
  - IP addresses of fabric interconnect A and B—second and third in the subnet
  - Blade IP addresses—remaining in subnet
- Maximum 323 addresses needed

Site	clusterID	Mgmt IP Pool
Athens	01	172.16.1.0/24
	02	172.16.2.0/24
	03	172.16.3.0/24
	04	172.16.4.0/24
Sparta	01	172.17.1.0/24
	02	172.17.2.0/24
Crete	01	172.18.1.0/24
Troy	01	172.19.1.0/24



Each Cisco UCS cluster also requires a pool of IP addresses that are used for management access to the individual server blade. An individual Cisco UCS cluster can have up to 320 server blades—each requiring an IP address, a cluster IP address, and the addresses of the fabric interconnect switches A and B. Thus, a maximum of 323 IP addresses are required for a cluster.

The first vital part of the management configuration presents the IP addressing setup. First, Cisco UCS requires management IP addresses for remote access (such as HTTPS or SSH) as follows:

- IP address of fabric interconnect switch A
- IP address of fabric interconnect switch B
- Subnet mask = 255.255.255.0
- Default gateway = 10.240.253.1
- High-availability configuration = cluster

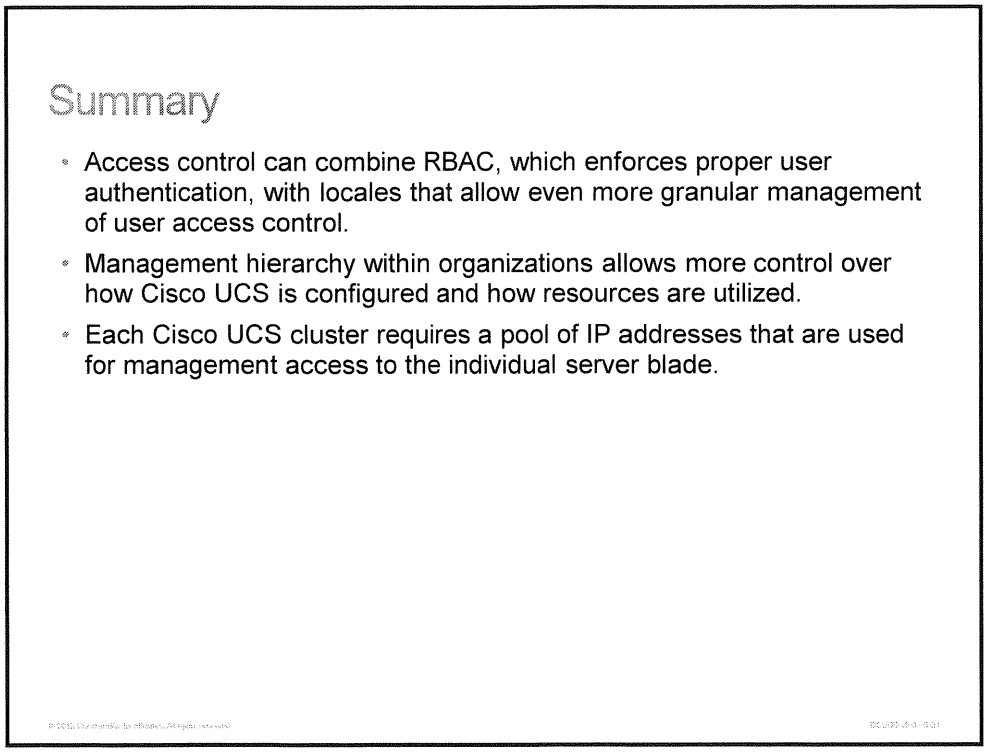
Second, the pool of IP addresses required for remote keyboard, video, mouse (KVM) access needs to be defined under the management IP address pool. Note that the IP addresses are typically dynamic in their nature, meaning there is no control over what server blade gets which IP address.

Third, each server blade—that is, service profile bound to server blade—typically requires a static IP address (an address that does not change) for features like VMware Distributed Resource Scheduler (DRS).

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

Summary

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

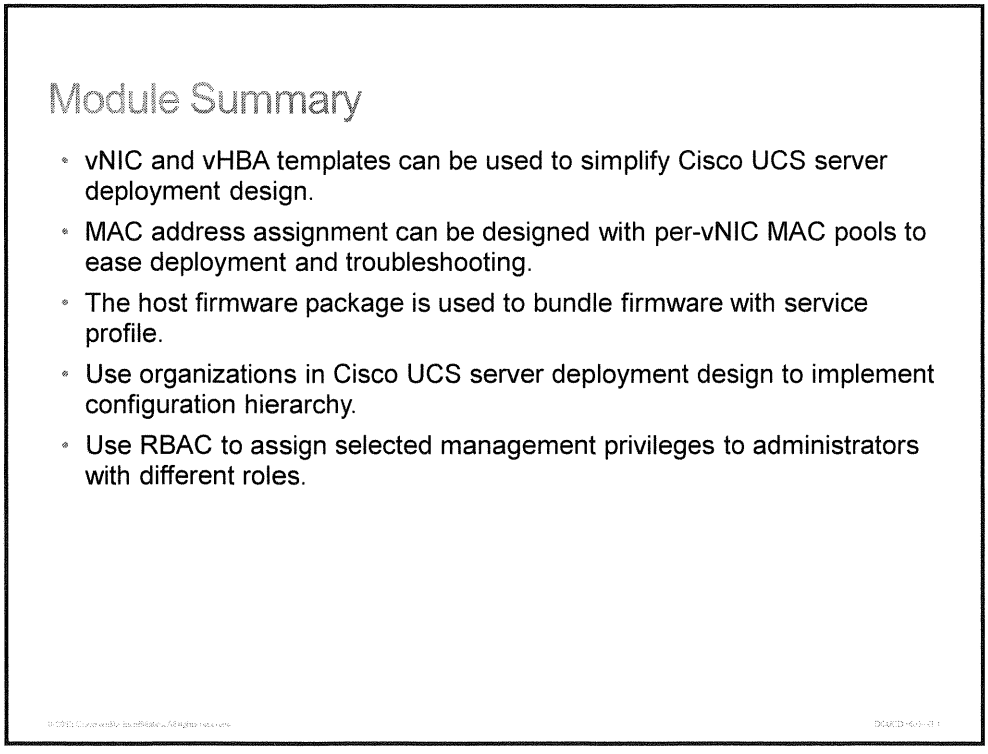


https://t.me/learningnets

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

Module Summary

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



https://t.me/learningnets

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

Module Self-Check

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

Q1) Which option makes server deployment design easier? (Source: Designing Cisco UCS Server Deployment)

A) scrub policy
B) autodiscovery policy
C) vNIC template
D) manual address assignment

Q2) In which two ways can you assign server blades to the server pool? (Choose two.) (Source: Designing Cisco UCS Server Deployment)

A) manually
B) using qualification and pool policy
C) automatically
D) in EHM mode

Q3) Which option is the default organization? (Source: Designing Unified Computing Management)

A) admin
B) system
C) root
D) main

Q4) What happens when the MAC pool is empty when a new service profile is created? (Source: Designing Unified Computing Management)

A) A burned-in address is used.
B) MAC overlapping occurs.
C) The MAC address from the default pool is used if present.
D) The MAC pool creation dialog box appears.

https://t.me/learningnets

Module Self-Check Answer Key

Q1) C

Q2) A, B

Q3) C

Q4) C

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Module 6

Cisco Unified Computing Solution Applications

Overview

The Cisco Unified Computing System (UCS) is a very flexible server infrastructure that can be used for various application needs and deployments. There are many applications with specific needs and design requirements that Cisco UCS can address. To design a solution for such applications using Cisco UCS, the architect needs to understand the application specifics.

This module describes and explains two distributed computing applications, unified communications and Hadoop, as well as their specifics and Cisco UCS options that fit the application needs.

Module Objectives

Upon completing this module, you will be able to recognize two specific applications that can be run on Cisco UCS. This ability includes being able to meet these objectives:

- Identify Cisco Unified Communications on the Cisco UCS solution and its specifics
- Identify the distributed computing applications

https://t.me/learningnets

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

Lesson 1

Designing Cisco Unified Communications on Cisco UCS

Overview

Unified communications presents one of the most widely used solutions, which today is run on data center resources. Designing, deploying, and running such a solution has its own requirements that need to be properly addressed.

This lesson identifies the Cisco Unified Communications solution and its characteristics that govern the selection and design of the Cisco Unified Computing System (UCS) solution for unified communications.

Objectives

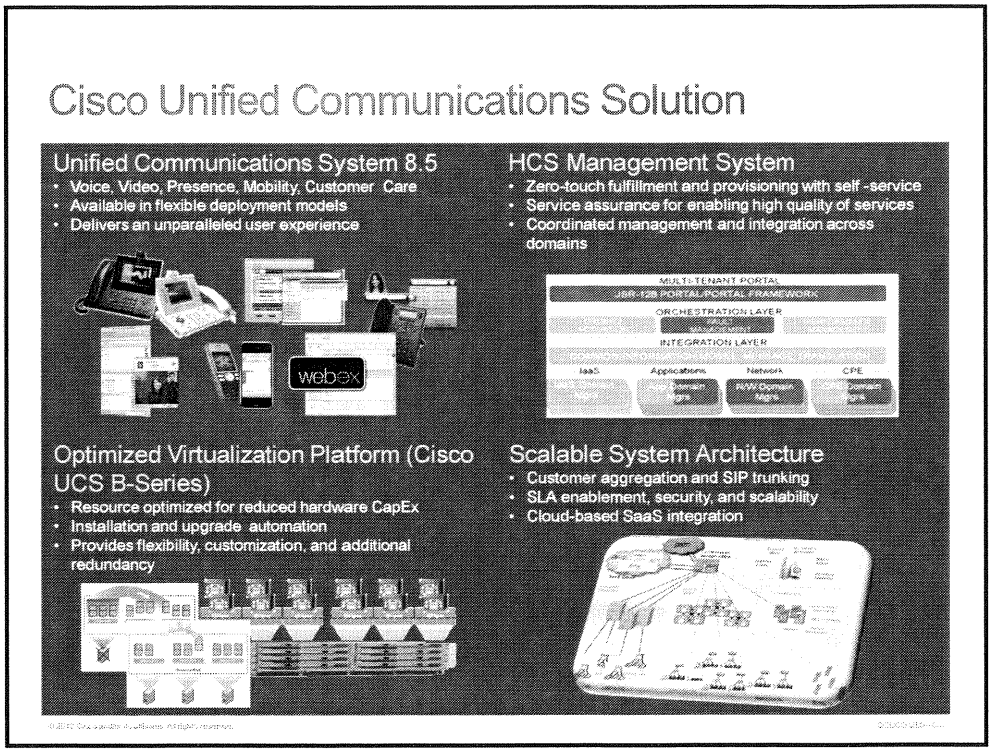
Upon completing this lesson, you will be able to identify Cisco Unified Communications on the Cisco UCS solution and its specifics. You will be able to meet the following objectives:

- Recognize Cisco Unified Communications solutions
- Assess Cisco Unified Communications characteristics
- Describe the two options for deploying Cisco Unified Communications on Cisco UCS

https://t.me/learningnets

Cisco Unified Communications Overview

This topic identifies Cisco Unified Communications solutions.

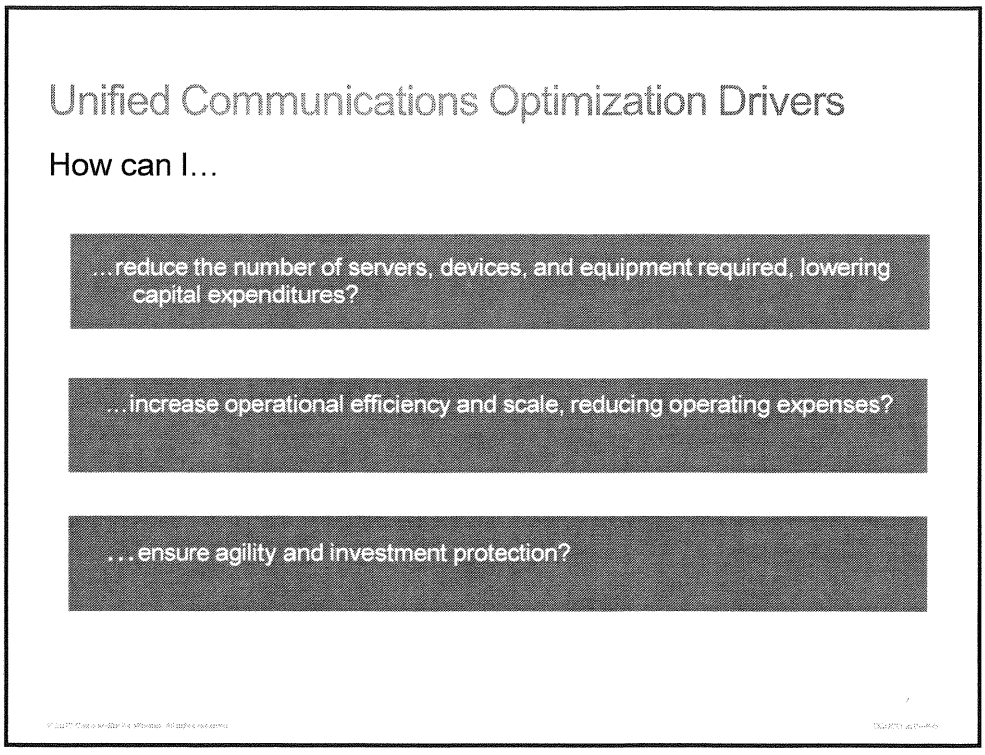


Cisco Unified Communications uses the network as a platform for collaboration. Applications can be flexibly deployed onsite, on-demand, and in blended deployment models. Intercompany and intracompany collaboration is facilitated with a wide array of market-leading solutions:

- **Conferencing:** Give your workforce compelling, productive, and cost-effective conferencing experiences with high-quality voice, video, and web conferencing.
- **Customer care:** Get closer to customers while increasing satisfaction and loyalty. Proactively connect people with the information, expertise, and support they need.
- **IP communications:** Enhance workforce productivity by extending consistent communications services to employees in all workspaces by using a full suite of IP communications solutions and endpoints.
- **Messaging:** Communicate effectively, and with a high level of security, within and between companies. View real-time presence information and communicate using email, instant messaging, and voicemail.
- **Mobile applications:** Increase mobile employee productivity and responsiveness to customers while controlling mobile costs by making mobile devices extensions of the enterprise network itself.

© 2012 Cisco Systems, Inc.

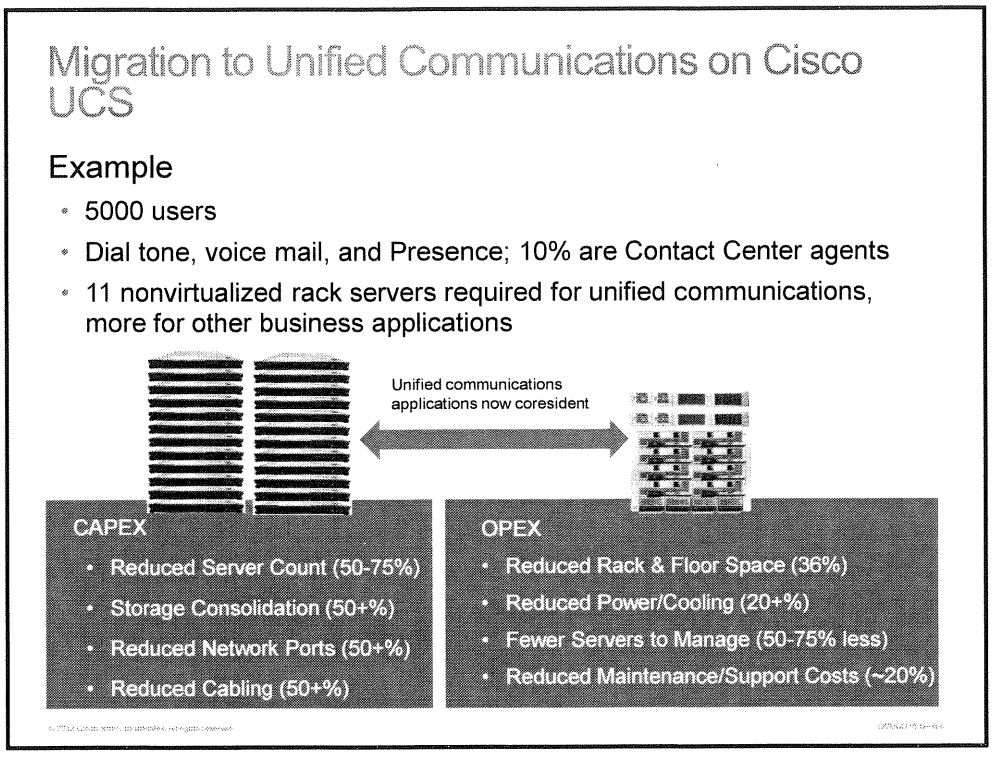
https://t.me/learningnets



The Cisco Unified Communications architects are pursuing how to optimize Cisco Unified Communications deployments, which encompasses the following questions:

- How can we reduce the number of servers, devices, and equipment required, thereby lowering capital expenditures?
- How can we increase operational efficiency and scale, thereby reducing operating expenses?
- How can we ensure agility and investment protection?

https://t.me/learningnets



The Cisco Unified Communications on Cisco UCS enables customers to deploy the applications in a virtualized environment.

Customers can deploy Cisco Unified Communications applications on:

- Supported Cisco UCS setups
- VMware vSphere ESXi server virtualization platform

This enables customers to ensure performance, reliability, management, and high availability for the Cisco Unified Communications.

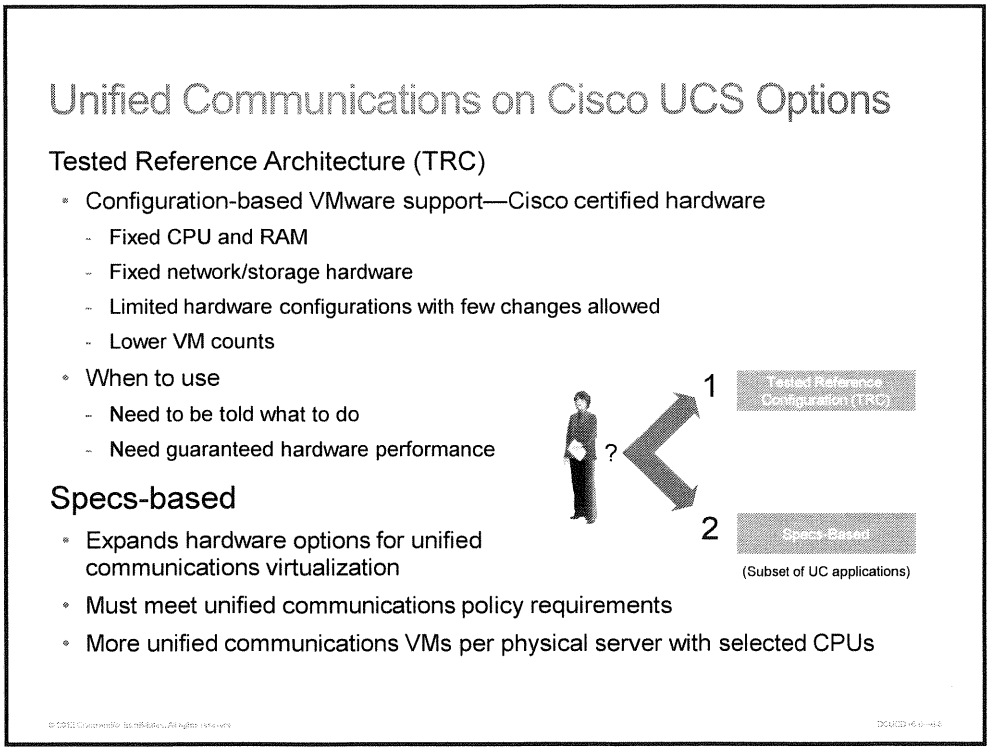
The figure details the savings the customer can achieve by deploying Cisco Unified Communications on Cisco UCS.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Cisco Unified Communications Characteristics

This topic assesses Cisco Unified Communications characteristics.



Tested Reference Configuration

Tested Reference Configurations (TRCs) are select hardware configurations tested and documented for specific guaranteed performance, capacity and application coresidency scenarios running "full-load" Cisco Unified Communications virtual machines (VMs). These are intended for customers who want a packaged solution from Cisco that is pre-engineered for a specific deployment scenario and customers who are not sufficiently experienced with hardware virtualization.

TRCs are intended for these types of customers:

- Lower levels of virtualization proficiency
- Require guaranteed Cisco Unified Communications application VM performance when the solution elements under Cisco Unified Communication' control are configured correctly per Cisco Unified Communications support policy
- Desire Cisco to manage hardware sizing, thereby determining hardware specifications and translating those specifications into a hardware parts list to ensure the guaranteed performance
- Those willing to accept for the guarantee a tradeoff of fewer choices for hardware and fewer choices for VM deployment options.

https://t.me/learningnets

TRCs do cover the following:

- Definition of hardware specifications down to the Cisco UCS part number level
- Direct-attached storage (DAS)/Redundant Array of Independent Disks (RAID) options to use when DAS is part of the definition
- Expectations for supportable mixture of VM quantity and size when VMs are at full load, based on Cisco Unified Communications sizing rules and coresidency support policy.

TRCs do not cover the following:

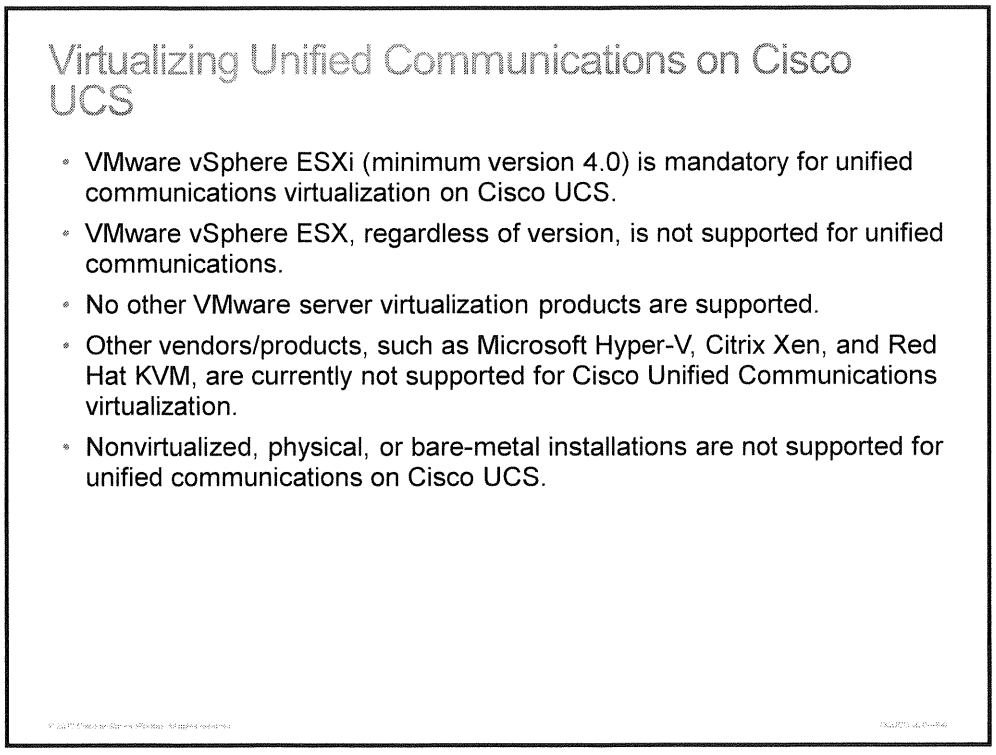
- Configuration of solution elements outside Cisco Unified Communications control, such as storage-area networking (SAN)/network-attached storage (NAS)-attached storage arrays
- Configuration settings or step-by-step procedures for hardware BIOS, firmware, drivers, and RAID options
- Configuration settings, patch recommendations or step-by-step procedures for VMware (use Cisco UCS and VMware best practices subject to Cisco Unified Communications support policy rules)
- Configuration settings or step-by-step procedures for quality of service (QoS) parameters, virtual-to-physical network mapping or adapters (these will vary by deployment, so use Cisco UCS and VMware best practices subject to Cisco Unified Communications support policy rules)

Specs-Based Configuration

Specs-based configuration is intended for customers with extensive expertise in virtualization and server/storage sizing who wish to use their own hardware standards. This hardware has not been directly tested or documented by Cisco Unified Communications, but it meets certain specifications for which Cisco Unified Communications applications are supported. Performance is not guaranteed, but the TRC may be used for guidance when sizing hardware.

Caution Customers who choose this option are responsible for sufficiently sizing hardware to meet Cisco Unified Communications VM performance needs at the level of the load deployed. Customers are also responsible for resolving performance problems by adjusting either the hardware configuration or the VM deployment model. Cisco is not responsible for performance problems resulting from use of specs-based support when the problem can be resolved by deployment changes such as migrating the VM to a faster hardware platform, powering off or migrating some of the other virtual machines to a different physical server, or splitting up a dense Cisco Unified Communications VM into multiple, less-dense Cisco Unified Communications VMs on different physical servers. It is recommended to use the TRC for guidelines on hardware sizing. Deploying specifications that are better than TRC *should* provide better performance but is not guaranteed. Likewise, deploying specifications that are worse than TRC *should* provide worse performance, but it is not guaranteed. Use of a processor with slower speed than the TRC is not supported as the deployment will have unknown and unspecified reduction in capacity and performance.

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



Supported Virtualization Hypervisor Vendors

- At this time, the only vendor supported for Cisco Unified Communications is VMware.
- Other vendors/products, such as Microsoft Hyper-V, Citrix Xen, and Red Hat KVM, are not supported for Cisco Unified Communications virtualization.

Supported VMware Products

VMware vSphere ESXi is mandatory for Cisco Unified Communications virtualization and Cisco Unified Communications on Cisco UCS:

- Cisco Unified Communications is only supported on Cisco UCS servers with VMware vSphere ESXi.
- Nonvirtualized, physical, or bare-metal installations are not supported for Cisco Unified Communications on Cisco UCS.

In all cases, the customer is required to have a valid service contract (either with Cisco or with VMware, depending on the purchase) to be supported.

ESX and ESXi are architecture options for VMware vSphere. Cisco Unified Communications only supports ESXi. Recall that vSphere 3.x and 4.x offered both options, but 5.0+ only offers ESXi. VMware vSphere ESX, regardless of the version, is not supported for Cisco Unified Communications. This is due to technical reasons and VMware transitioning from ESX to ESXi. Note that a VMware ESX cluster can contain VMware ESXi servers running Cisco Unified Communications. No other VMware server virtualization products are supported.

Supported Versions of VMware vSphere ESXi

The minimum required version for VMware vSphere ESXi is 4.0. But a higher minimum version may be required depending on what Cisco Unified Communications applications and versions you run and what hardware you select.

https://t.me/learningnets

VMware Performance Indicators

Performance Area	Object	Counter	Acceptable range
CPU	Host	Usage	Less than 80%
CPU	Virtual Machine	Ready	Less than 3%
Memory	Host	Consumed	General trend is stable
Memory	Host	Balloon/Swap used	0 Kb
Disk	Specific datastore	Kernel command latency	Less than 3 ms
Disk	Specific datastore	Physical device command latency	Less than 20 ms
Disk	Specific datastore	Average commands issued per second	Less than LUN capacity

The figure lists the performance indicators to monitor and view when a VM is having suboptimal (or bad) performance. Most counters are from the ESXi host, which can give a perspective of VM interactions and overall host and data store utilization.

Performance indicators are still valid from within VMs. For the Cisco Unified Communications applications that support it, use Cisco UCS Real-Time Monitoring Tool (RTMT) or the Perfmon data for to analyze the performance of the Cisco Unified Communications application. Data from these tools provides a view of the guest performance: disk, CPU, memory, and other details.

Move to the VMware infrastructure when there is a need to get the ESXi host perspective. Use the vSphere Client to view data.

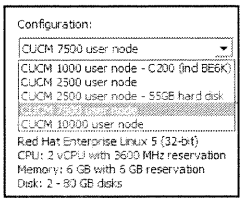
- If vCenter is available, historical data is available through the client
- If vCenter is not available, live data from the host is available through the client

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

VM Sizing

- VM virtual hardware defined by a VM template:
  - vCPU, vRAM, vDisk, vNICs
- Capacity
  - A VM template is associated with a specific capacity
- VM templates are packaged in an OVA file
- There are usually different VM templates per release. For example:
  - CUCM_8.0_vmv7_v2.1.ova
  - CUCM_8.5_vmv7_v2.1.ova
  - CUCM_8.6_vmv7_v1.5.ova
  - Includes product, product version, VMware hardware version, template version



VMs still have resources, only that they are now virtual resources, virtual CPU, and memory. There are also templates that define the specifications of a specific VM, thereby defining the exact virtual resources that are applicable to the VM. For each VM template, there is associated capacity that goes with it.

The following are identical regardless of the support model chosen:

- Virtual machine Open Virtualization Archive (OVA) requirements
- VMware product, version, and feature support
- VMware configuration requirements for Cisco Unified Communications
- Application/VM coresidency policy (specifically regarding application mix, third-party support, reservations, and oversubscription). Max VM count per server with specs-based depends on the physical server configuration and the OVAs deployed. For example, Cisco UCS C210 M2 TRC#1 deployed with multiple Cisco Unified Communications Manager 7.5K user OVAs can support 4:1. If specs-based was used with a supported dual 6-core CPU and appropriate increases in memory and storage, the same server model running multiple Cisco Unified Communications Manager 7.5K user OVAs could support 6:1.

https://t.me/learningnets

Examples of Supported VM Configurations (OVAs)

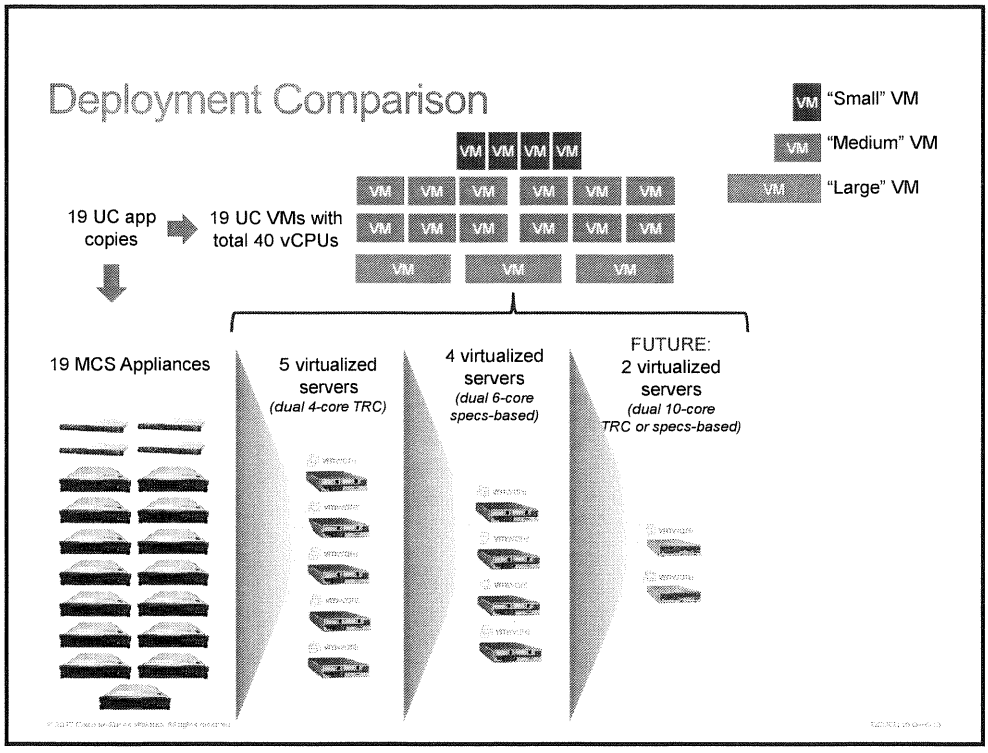
Product	Scale (users)	vCPU	vRAM (GB)	vDisk (GB)	Notes
Unified CM 8.6	10,000	4	6	2 x 80	Not for C200/BE6k
	7500	2	6	2 x 80	Not for C200/BE6k
	2500	1	4	1 x 80 or 1x55GB	Not for C200/BE6k
	1000	2	4	1 x 80	For C200/BE6k only
Unity Connection 8.6	20,000	7	8	2 x 300	Not for C200/BE6k
	10,000	4	6	2 x 146	Not for C200/BE6k
	5000	2	6	1 x 200	
	500	1	4	1 x 160	
Unified Presence 8.5	5000	4	6	2 x80	Not for C200/BE6k
	2500	2	4	1 x 80	Not for C200/BE6k
Unified CCX 8.5	400 agents	4	8	2 x 146	Not for C200/BE6k
	300 agents	2	4	2 x 146	Not for C200/BE6k
	100 agents	2	4	1 x 146	For C200/BE6k only

Up-to-date list of OVAs is available at http://www.cisco.com/go/uc-virtualized

This is an example of the different OVAs that Cisco supports:

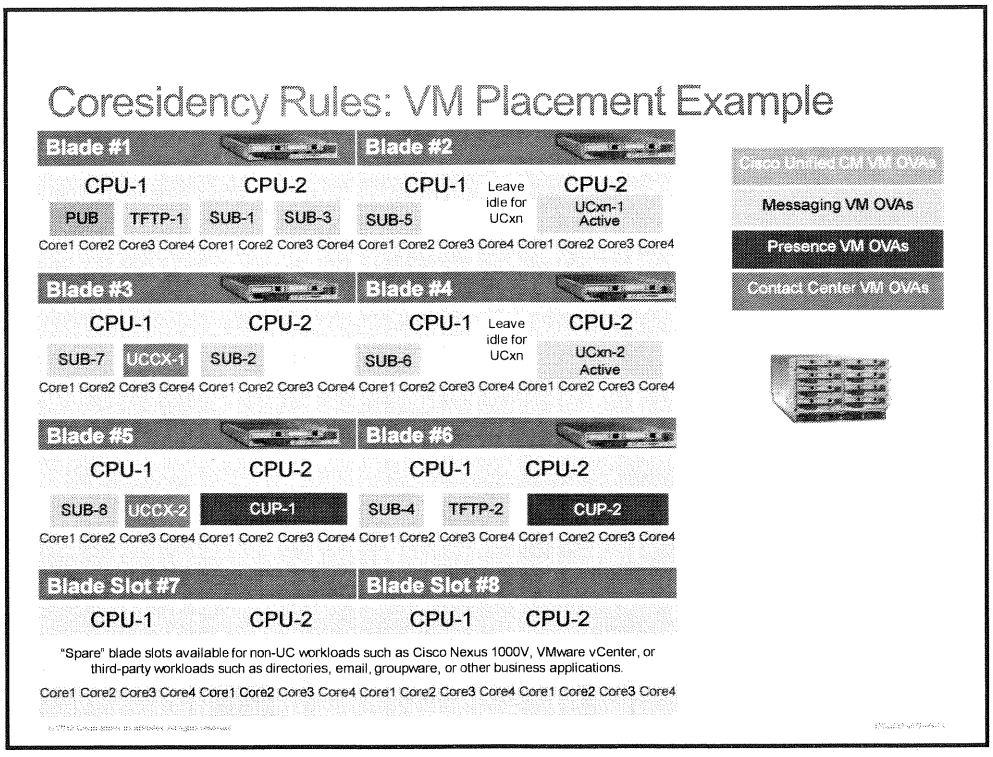
- OVA for 7.5k users (two vCPUs):
  - More "efficient" than the 10k OVA, the 10k OVA requires 100 percent more vCPU but can process only 33 percent more users (10k versus 7.5k users)
  - For small deployments—typically less than 30k users—the 7.5k OVA fits well and may require a lower number of required physical servers.
  - Example: 30k users, 1:1 redundancy:
    - With 7.5k OVA: 8 * 2 vCPU = 16 vCPU for subs (two servers if using TRC)
    - With 10k OVA: 6 * 4 vCPU = 24 vCPU for subs (three servers if using TRC)
- OVA for 10k users (four vCPUs):
  - May be preferable for large deployments if the customer wants to reduce the number of clusters
  - May be worth adding hardware to have fewer clusters. This could simplify deployment, administration, monitoring, upgrades, and backups and can reduce cost, especially with newer processors that have more cores.
  - Example 1: Cisco Unified Communications Manager cluster with 40,000 users
    - With 7.5k user OVA: two clusters or one megacluster required
    - With 10k user OVA: one cluster
  - Example 2: Cisco Unified Communications Manager cluster with 80,000 users
    - With 7.5k user OVA: three clusters or two megaclusters
    - With 10k user OVA: one megacluster

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



When sizing the Cisco Unified Communications on the Cisco UCS, solution the number of VMs plus the coresidency rules govern the hardware and Cisco UCS server quantity selection.

https://t.me/learningnets



There are rules on how to place a different VM. The sum of all the vCPUs of the Cisco Unified Communication application, virtual CPU (vCPU) must not exceed the physical cores that are available on your server. There is a 1:1 mapping between vCPU and the physical core. If a user has VMs and requires one vCPU, then this user can have up to eight VMs (if using a TRC). If a VM requires two vCPUs, it is possible to have up to four VMware chassis.

There is a similar rule for the memory. The sum of all virtual RAM (vRAM) must not exceed the actual vScale RAM that a user has. Remember to reserve 2 GB for ESXi (TRC or spec-based).

Only Cisco Unified Communications applications can be coresident on the same server or blade. There cannot be coresidency with third-party applications on the same server or blade; however, coresidency on the same chassis is acceptable). The Cisco Unified Communications application in this case means the standard Cisco Unified Communications application plus the attendant console, and the Network and Service Management (NMTG) Interoperability Program Cisco Unified Communications products (Cisco Unified Communications applications include CUxAC and NMTG Cisco Unified Communications products). Third-party applications include some Cisco products that are not part of the Cisco Unified Communications portfolio, such as the Nexus 1000V and vCenter.

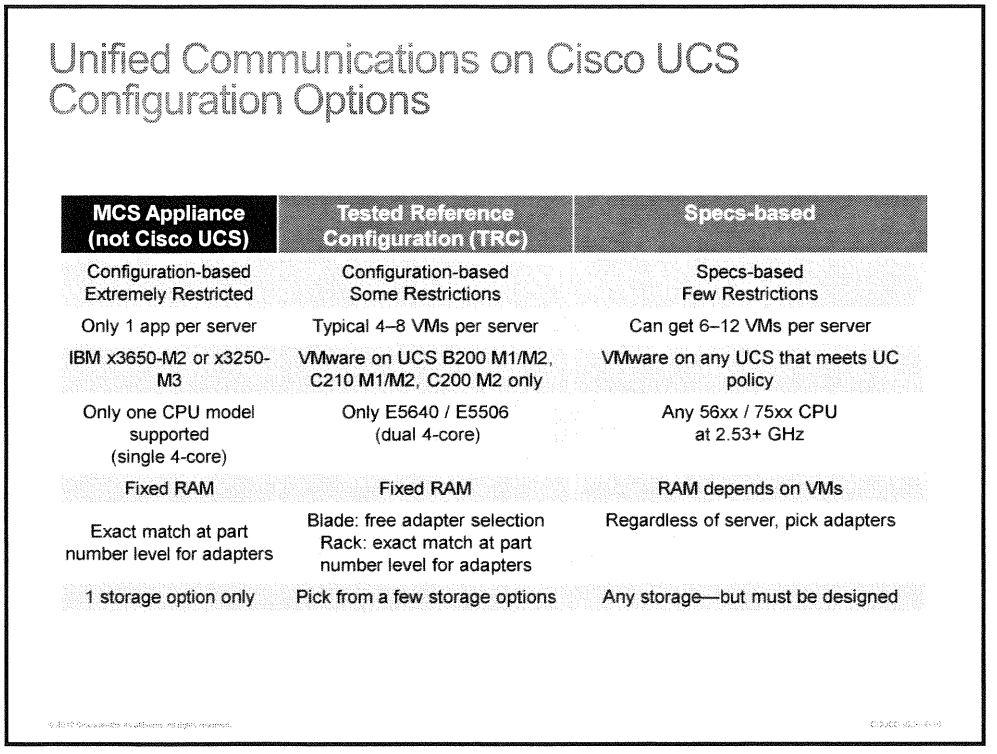
Remember to distribute Cisco Unified Communications application nodes across Cisco UCS blades, chassis, and sites in order to minimize failure impact.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

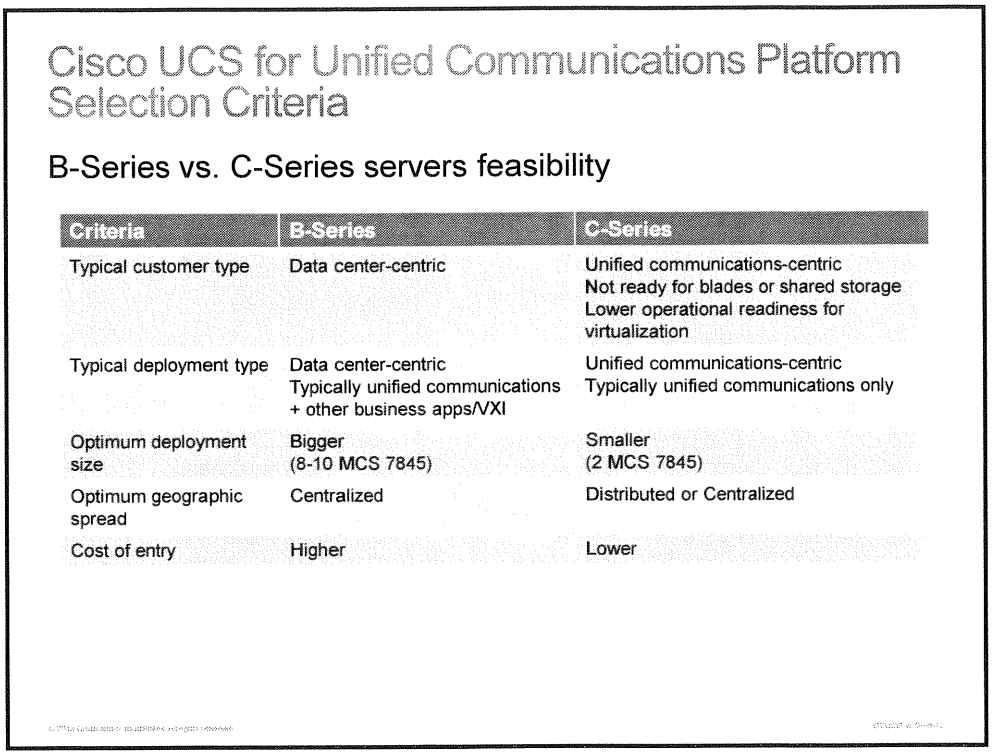
Selecting Cisco UCS Hardware for Cisco Unified Communications

This topic describes the two options for deploying Cisco Unified Communications on Cisco UCS: TRC and specs-based deployment.



The figure compares the TRC and specs-based hardware configurations for Cisco Unified Communications on Cisco UCS design.

https://t.me/learningnets



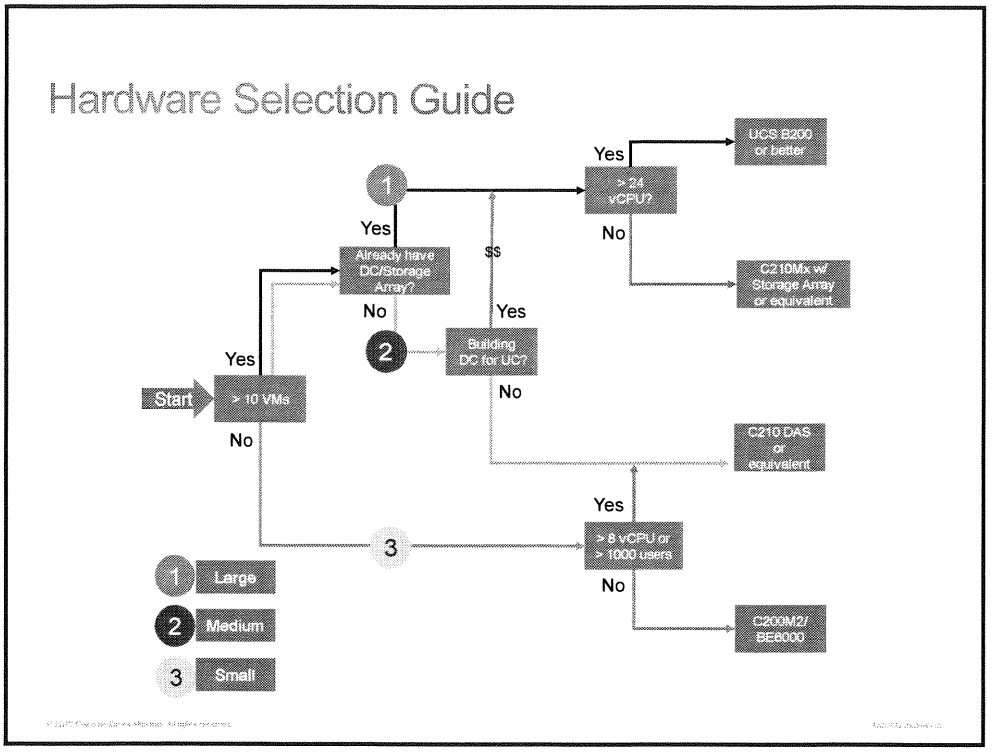
Cisco has introduced software application versions from the Cisco Unified Communications portfolio (Versions 8.0(2) and later) that are supported for production use on specific configurations of VMware vSphere (ESXi), Cisco UCS servers, and various storage configurations.

Either UCS B-Series or C-Series can be used for running Cisco Unified Communications on Cisco UCS.

The figure lists the decision criteria.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets



Platform Options

There are two platform options:

- The first option is traditional tested reference configurations, which are very specific hardware configurations where the performance is guaranteed.
- The second option is the specific hardware-based support; customers can use their own hardware, but performance of that hardware is not guaranteed.

TRC:

- Tested and documented by Cisco
- Performance is guaranteed
- For customers who want a packaged solution from Cisco with guaranteed performance
- Based on very specific hardware configurations; if customers deviate from these configurations, the hardware is not considered TRC anymore.

TRCs are based on three platforms: the B200, C210, and C200. The types of storage that are supported are NAS, directly attached storage with local drive, and Fibre Channel SAN.

With the B200, the Cisco Unified Communications application has to run on the SE10 as it cannot run locally on the blade itself. VMware can run locally on the blade if you choose a configuration with hard drives. The other possibility is that VMware can run directly on the SAN that is called Boot From SAN. With that configuration, a customer does not need any disc on the B200 blade.

With the C210, your Cisco Unified Communications application can run locally on a hard drive. It is also possible to put a Cisco Unified Communications application on the SAN (for VMware). VMware Hypervisor can either be run locally or off of the SAN.

For the third platform, the C200, there is only one option: all the Cisco Unified Communications applications and the VMware Hypervisor have to run locally on the server.

https://t.me/learningnets

For Cisco Unified Communications without virtualization, customers need a separate physical server for each Cisco Unified Communications application. For example, having Cisco Unified Communications Manager, Cisco Unity Connection, or Cisco Connect Express, on the same physical server is not possible.

With virtualization, users can have multiple Cisco Unified Communications applications on the same physical server. So instead of having four physical machines with dedicated CPU/memory/hard disk drive (HDD), a virtual deployment will have the same "server" components but will be deployed on virtual hardware that will provide the appropriate vCPU/vMemory/virtual HDD (vHDD) required by the application. Different applications will have different vCPU/vMemory/vHDD requirements, but those requirements are abstracted from the physical server and allocated via an OVA definition.

All VMware vSphere editions, including a free hypervisor, are supported. Some vSphere editions might be required for large VMs or for support of some VMware features. Other hypervisors (such as Microsoft Hyper-V and Citrix XenServer) are not planned or roadmapped.

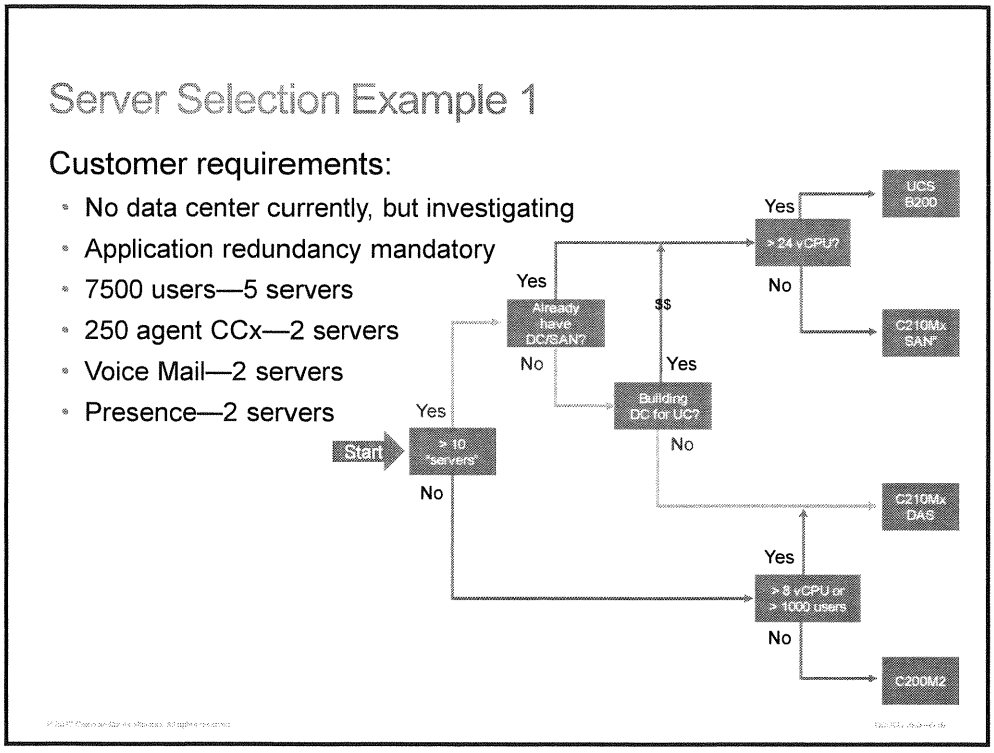
Hardware Selection

If the customers already have a data center established, then the B-Series chassis is going to be a suitable deployment. B-Series servers are more data center-centric.

Cisco UCS C-Series servers are good for customers used to the physical server environment. If customers are not familiar with data centers, they are probably much more comfortable with everything being local (storage-wise) on the server. The C210 is a very popular choice for customers looking to go virtual but do not have the data center set up yet.

If a customer uses few actual physical servers—if, for example, two servers are physically needed to support a site—then a C-Series is the best option. If a customer uses eight to 10 servers or is replacing eight to 10 physical servers, the B-Series is going to be better as it scales better.

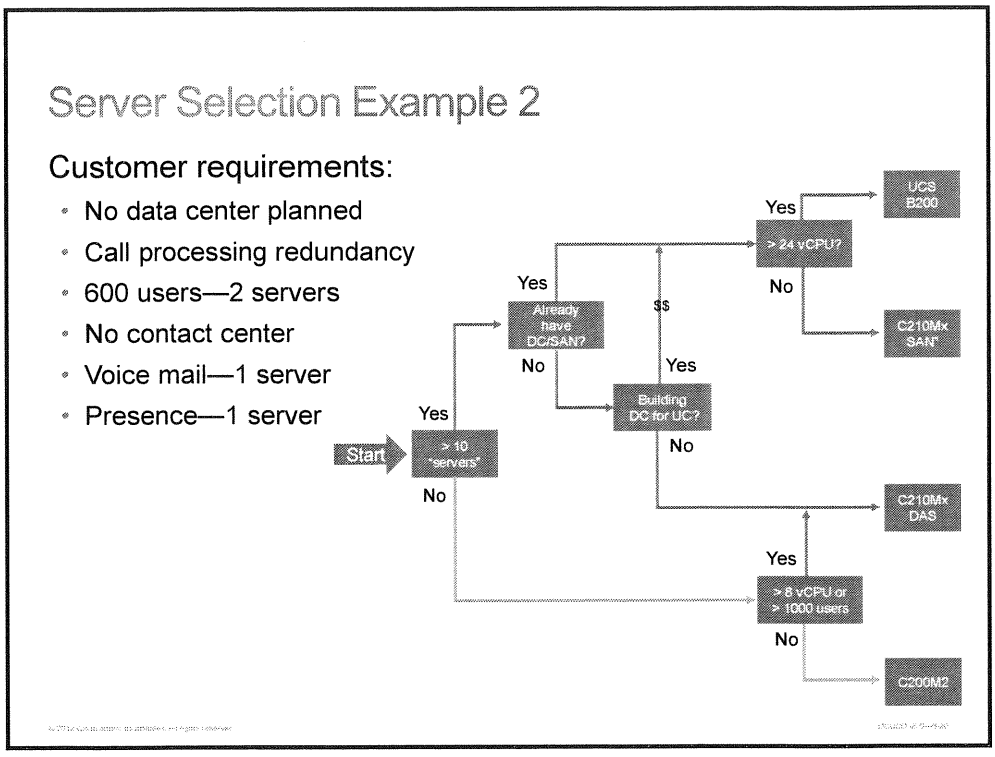
© 2012 Cisco Systems, Inc.
https://t.me/learningnets



The figure shows the Cisco UCS selection decision process for the following customer requirement:

- No data center currently, but investigating
- Application redundancy mandatory
- 7500 users – five servers
- 250 agent CCx – two servers
- Voice mail – two servers
- Presence – two servers.

https://t.me/learningnets

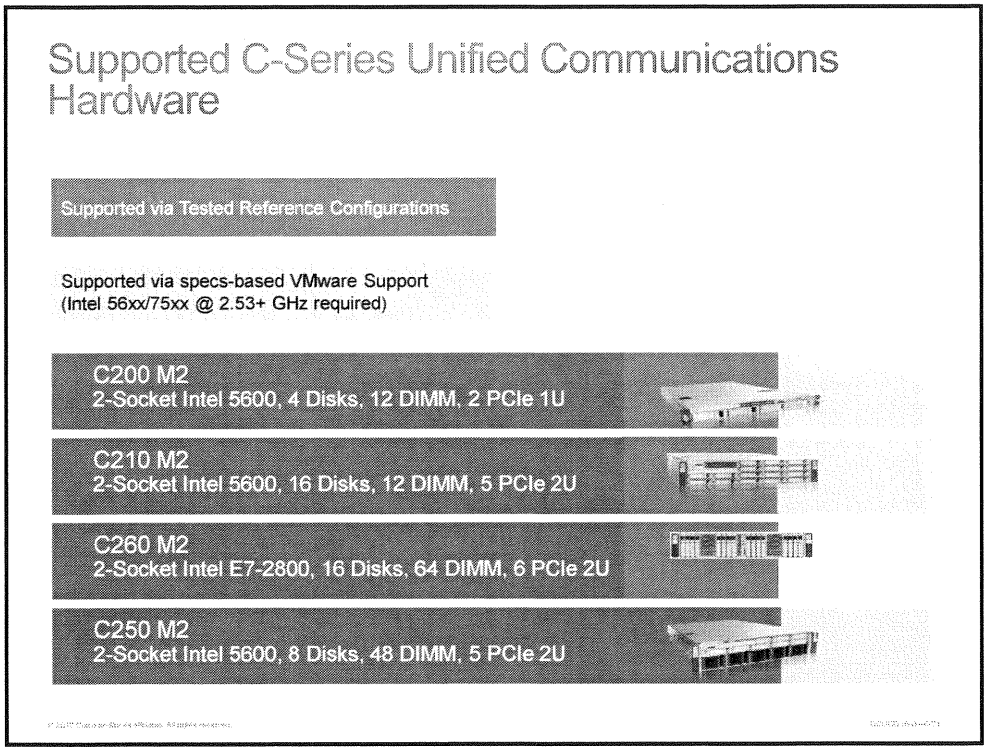


This is another example of selecting the Cisco UCS hardware for the customer requirements:

- No data center planned
- Call processing redundancy
- 600 users – two servers
- No contact center
- Voice mail – one server
- Presence – one server.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

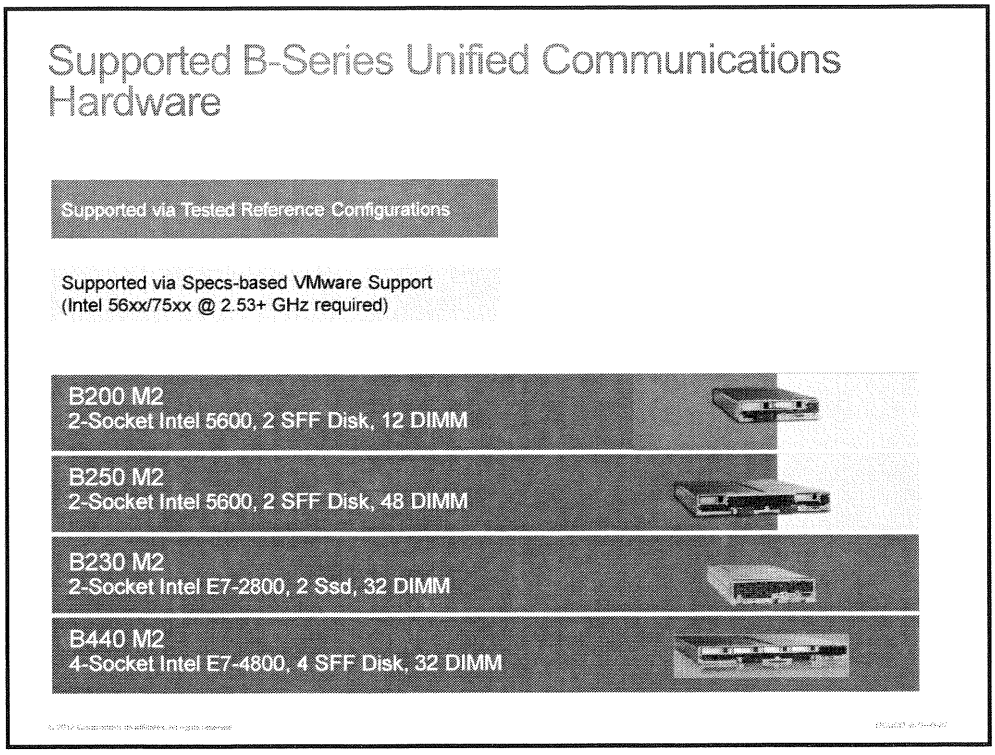


TRCs for C-series rack-mount servers are provided for the following servers and described as follows:

- Cisco UCS C260 M2 Rack-Mount Server with a specific configuration of DAS is supported.
- Cisco UCS C210 M2 General-Purpose Rack-Mount Servers with specific configurations of either DAS or Fibre Channel SAN are supported (including a disk-less server for VMware "boot from SAN" option).
- The Cisco UCS C200 M2 General-Purpose Rack-Mount Server with specific configurations of DAS is supported.

Cisco Unified Communications applications require the use of VMware vSphere 4 or 5 (ESXi) or Cisco Unified Communications Virtualization Foundation to support Cisco UCS hardware. No applications run or install directly on the server hardware; all applications run only as VMs. Cisco Unified Communications does not support a physical, bare-metal, or nonvirtualized installation on this server hardware.

https://t.me/learningnets



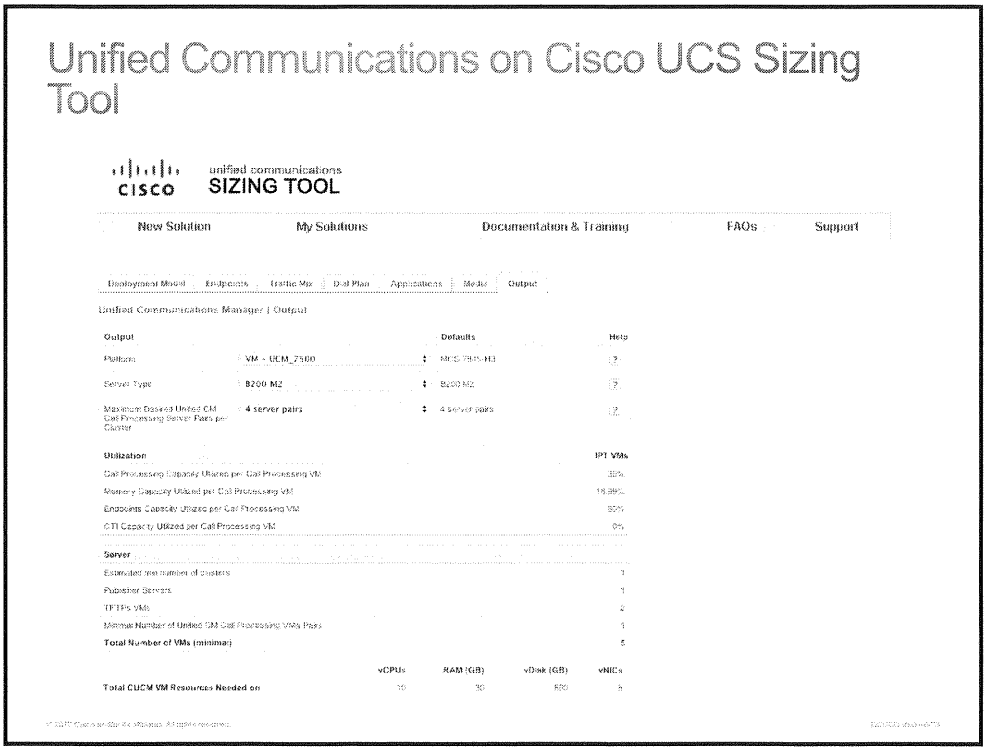
One example of TRC is a solution containing B-200 M2 servers sized only for VMs supporting up to 1000 users. Hardware requirements for Cisco Unified Communications on Cisco UCS B-200 M2 are as follows:

- CPU:
  - Dual E5506 (2.13 GHz, eight total physical cores available to apps)
- RAM:
  - 24 GB (6x4 GB—after ESXi 22 GB of total physical RAM available to apps)
- Storage:
  - DAS (four disks with RAID 10) for VMware and Cisco Unified Communications applications
  - 1.8 TB available to apps after RAID/VMware virtual file system (VMFS) overhead, not counting VM overhead
- Adapter: 1GbE network interface card (NIC)

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Unified Communications on Cisco UCS Sizing Tool



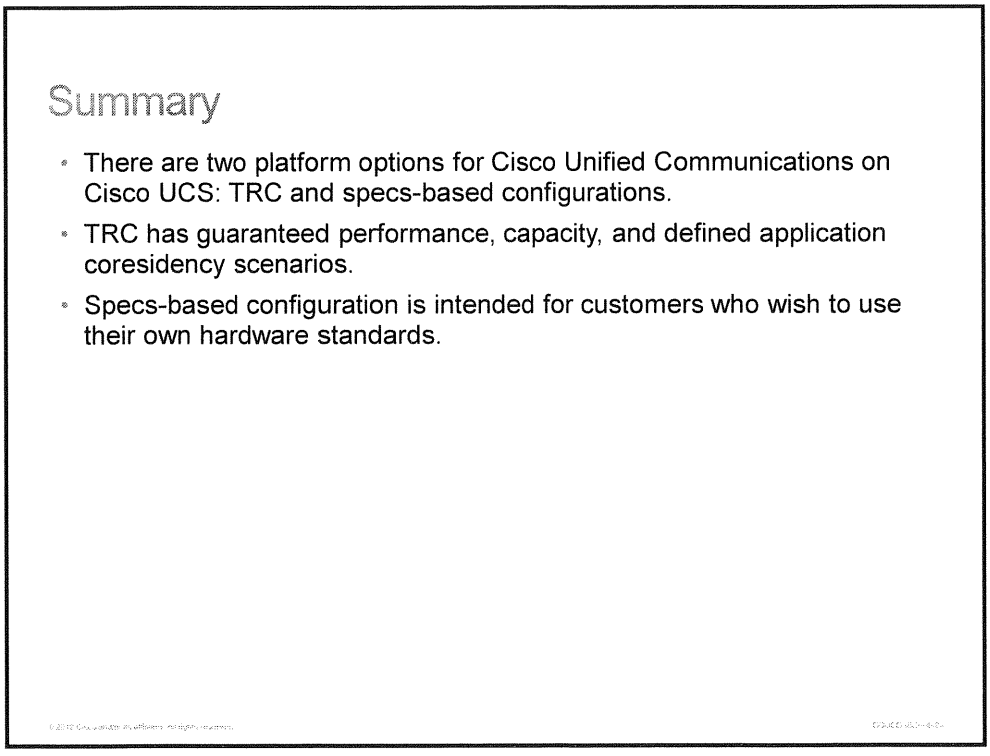
Use the Cisco Unified Communications sizing tool (http://tools.cisco.com/cucst) to size your deployment. (The number of endpoints is not always the limiting factor; it could be CPU, memory, or something else.)

The Cisco Unified Communications sizing tool has fields for the template as well. The sizing tool will tell a user how many VMs are needed. It will not tell the user how many servers to use; the user has to do the VM placement.

https://t.me/learningnets

Summary

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



References

For additional information, refer to these resources:

- http://docwiki.cisco.com/wiki/Unified_Computing_System_Hardware
- http://docwiki.cisco.com/wiki/Specification-Based_Hardware_Support
- http://docwiki.cisco.com/wiki/Unified_Communications_Virtualization
- http://tools.cisco.com/cucst/faces/landing.jsp

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Lesson 2

Designing Distributed Computing on Cisco UCS

Overview

The amount of data in the world is rising exponentially, and so the ways to manage such data in order to extract useful information are changing as well. To process vast amounts of data, a distributed divide-and-conquer model is used wherein data is divided into smaller chunks and processed by individual nodes in a large cluster in a manner known as distributed computing.

This lesson identifies the distributed computing applications, recognizes the Hadoop architecture and performance characteristics, and provides a Cisco Unified Computing System (UCS) design example for the Greenplum MR Hadoop solution.

Objectives

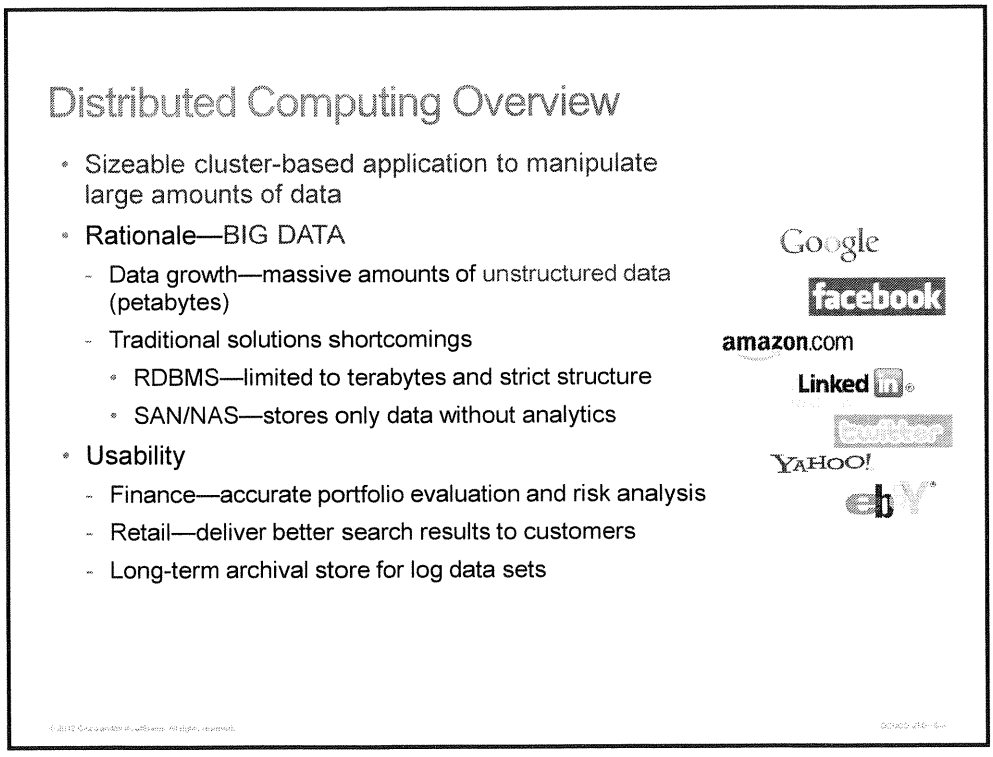
Upon completing this lesson, you will be able to identify the distributed computing applications. You will be able to meet the following objectives:

- Recognize distributed computing solutions and applications
- Understand general Hadoop architecture
- Assess general Hadoop performance characteristics
- Design Cisco UCS for the Greenplum MR Hadoop solution

https://t.me/learningnets

Distributed Computing Overview

This topic identifies distributed computing solutions and applications.



Large data is a foundational element of social networking and Web 2.0-based information companies. The enormous amount of data is generated as a result of democratization and ecosystem factors such as the following:

- **Mobility trends:** Mobile devices, mobile events and sharing, and sensory integration
- **Data access and consumption:** Internet, interconnected systems, social networking, and convergent interfaces and access models (Internet, search and social networking, and messaging)
- **Ecosystem capabilities:** Major changes in the information processing model and the availability of an open-source framework; the general-purpose computing and unified network integration

Data generation, consumption, and analytics have provided competitive business advantages for Web 2.0 portals and Internet-centric firms that offer services to customers and services differentiation through correlation of adjacent data.

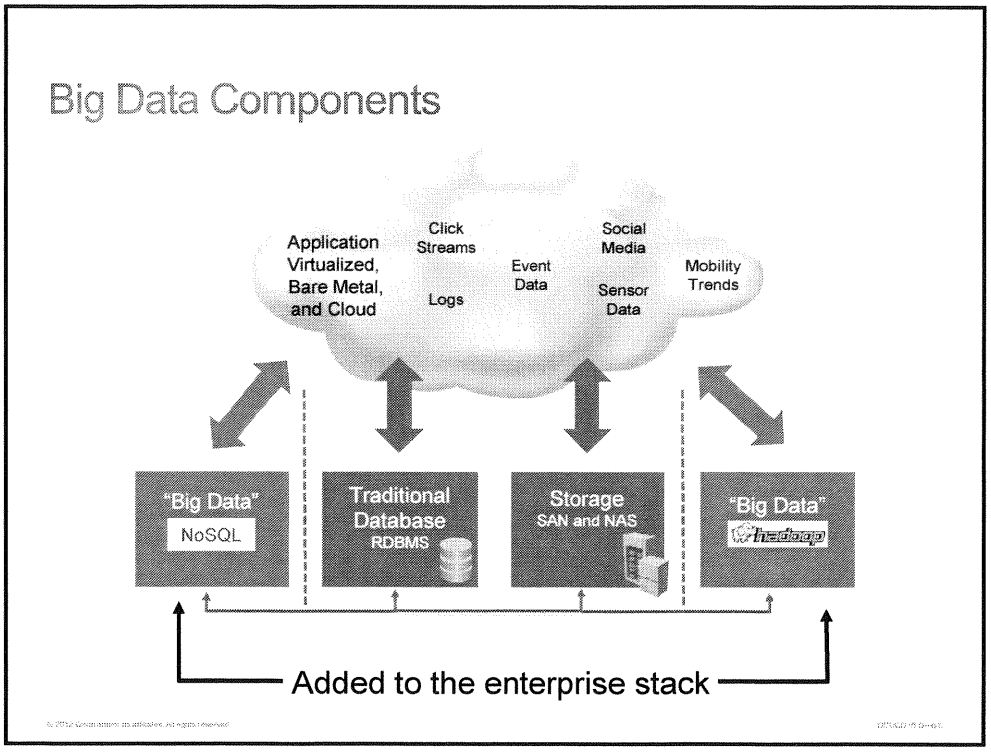
With the rise of business intelligence data mining and analytics spanning market research, behavioral modeling, and inference-based decisions, data can be used to provide a competitive advantage. These are just a few of the nearly limitless use cases of large data for the companies with large Internet presence:

- Targeted marketing and advertising
- Related attached sales promotions
- Analysis of behavioral social patterns
- Metadata-based optimization of workload and performance management for millions of users

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

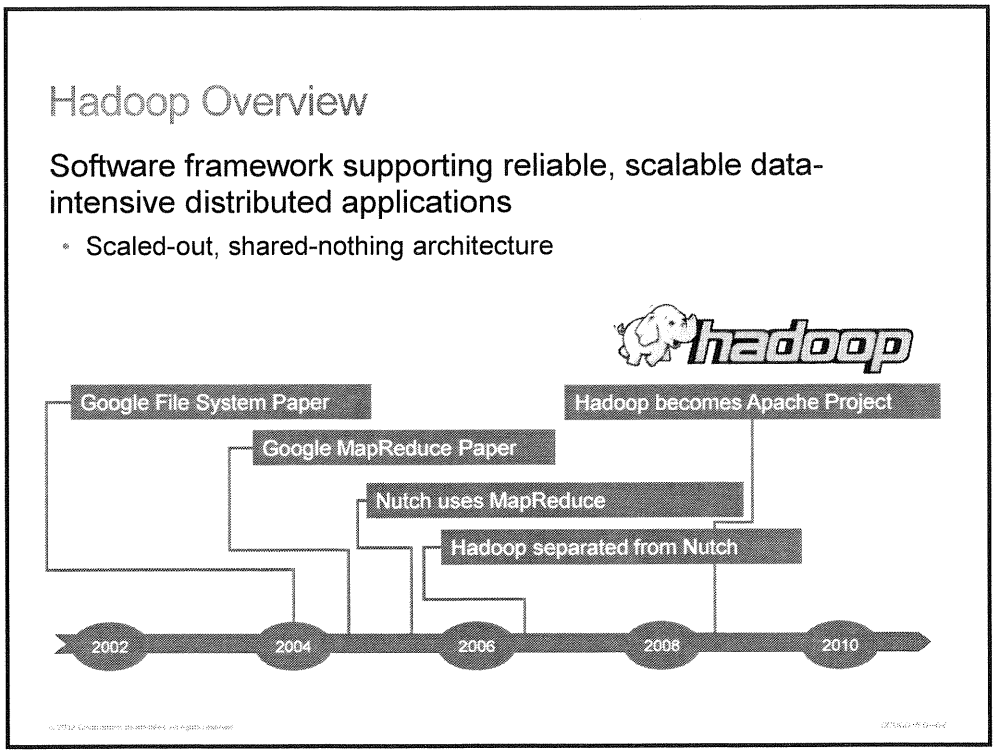
The requirements of traditional enterprise data models for application, database, and storage resources have grown over the years, and the cost and complexity of these models have increased along the way to meet the needs of significant data. This rapid change has prompted changes in the fundamental models that describe how large data is stored, analyzed, and accessed. The new models are based on a scaled-out, shared-nothing architecture, bringing new challenges to enterprises deciding what technologies to use, where to use them, and how to use them. The traditional model is now being expanded to incorporate new building blocks to address the challenges of large data with new information-processing frameworks purpose-built in order to meet the requirements of the significant data. However, these purpose-built systems must also meet the inherent requirements for integration into current business models, data strategies, and network infrastructures.



Two main building blocks are being added to the enterprise stack to accommodate large data:

- **Hadoop:** Provides storage capability through a distributed, shared-nothing file system as well as analysis capability through MapReduce.
- **NoSQL:** Provides the capability to capture, read, and update, in real time, the large influx of unstructured data and data without schemas. Examples include click streams, social media, log files, event data, mobility trends, and sensor and machine data.

https://t.me/learningnets



The challenge facing companies today is how to analyze this massive amount of data in order to best find those critical pieces of information that provide that company with a competitive edge. Hadoop provides the framework to process massive amounts of data: to either transform it into a more usable structure and format, or analyze and extract valuable analytics from it.

Hadoop History

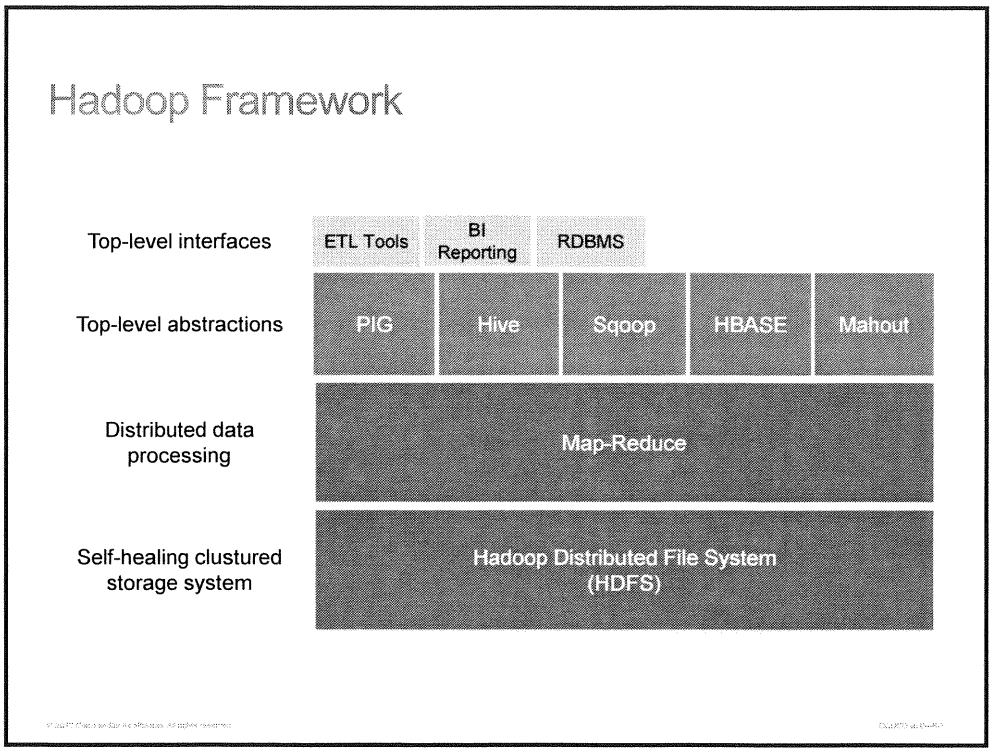
During the upturn in Internet traffic in the early 2000s, the scale of data transmitted reached terabyte and petabyte levels on a daily basis for many companies. At those levels, standard databases could no longer scale enough to manage the so-called "big data."

In 2004, Google published a paper on the Google File System (GFS) and another paper on MapReduce, the Google patented software framework for distributed computing and large data sets on a scaled-out shared-nothing architecture. These papers were published to address the challenge of sorting and scaling large data. The concepts in these papers were then implemented in Nutch, an open-source web-search software that enabled processing based on the sort-and-merge approach.

Hadoop, also based on open-source and Java architecture, was later derived from Nutch as an Apache project. Enabling cost-effective bulk computing, Hadoop is both a distributed file system modeled on GFS and a distributed processing framework using the MapReduce metaphor, according to its creator, Doug Cutting of Yahoo!

© 2012 Cisco Systems, Inc.

https://t.me/learningnets



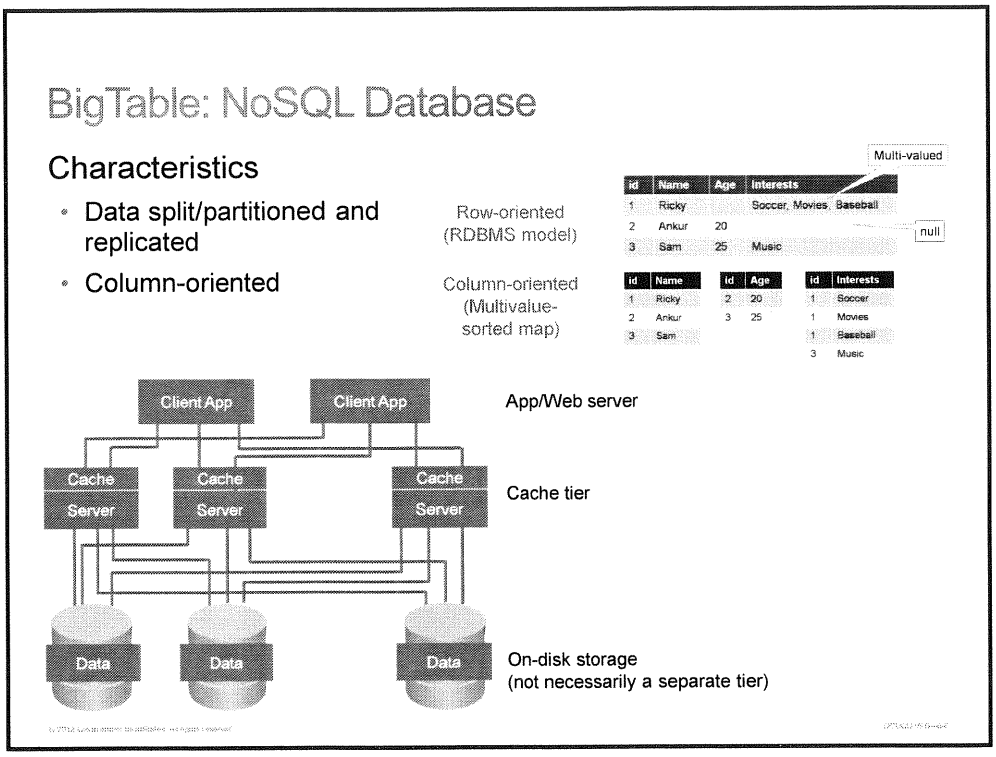
The Apache Hadoop is a software framework that can be used for processing large amounts of data in a distributed manner using clusters of computers exploiting a simple programming model.

The Hadoop is designed to scale up from one to thousands of servers, with each server capable of both local computation and storage. The high availability is part of the application layer (meaning the library itself is designed to detect and manage failures at the application layer) instead of relying on the hardware. This enables delivery of a highly available service in addition to a cluster of commodity computers that may be prone to failures.

The Apache Hadoop project includes the following subprojects:

- **Hadoop Common:** The common utilities that support the other Hadoop subprojects
- **Hadoop Distributed File System (HDFS):** A distributed file system that provides high-throughput access to application data
- **Hadoop MapReduce:** A software framework for distributed processing of large data sets on compute clusters
- **Other Hadoop-related projects:**
  - **Avro:** A data serialization system
  - **Cassandra:** A scalable multimaster database with no single points of failure
  - **Chukwa:** A data-collection system for managing large distributed systems
  - **HBase:** A scalable, distributed database that supports structured data storage for large tables
  - **Hive:** A data warehouse infrastructure that provides data summarization and ad hoc querying
  - **Mahout:** A scalable machine learning and data mining library
  - **Pig:** A high-level data-flow language and execution framework for parallel computation
  - **ZooKeeper:** A high-performance coordination service for distributed applications

https://t.me/learningnets



BigTable is using the highly distributed, shared-nothing architecture. The data is stored in many unreliable, general-purpose servers using partitioning and replication to achieve proper redundancy, high availability, and scalability levels. Data is partitioned by its key and stored in different servers, while at the same time being replicated multiple times at different servers.

Compared to the traditional relational database management system (RDBMS) systems that contiguously store each row on a disk, with BigTable each column is stored contiguously on disk. Such approach is used with the assumption that in most cases, not all columns are required for accessing the data. As a consequence, this means that more records are stored in a disk block, in turn resulting in disk I/O reduction.

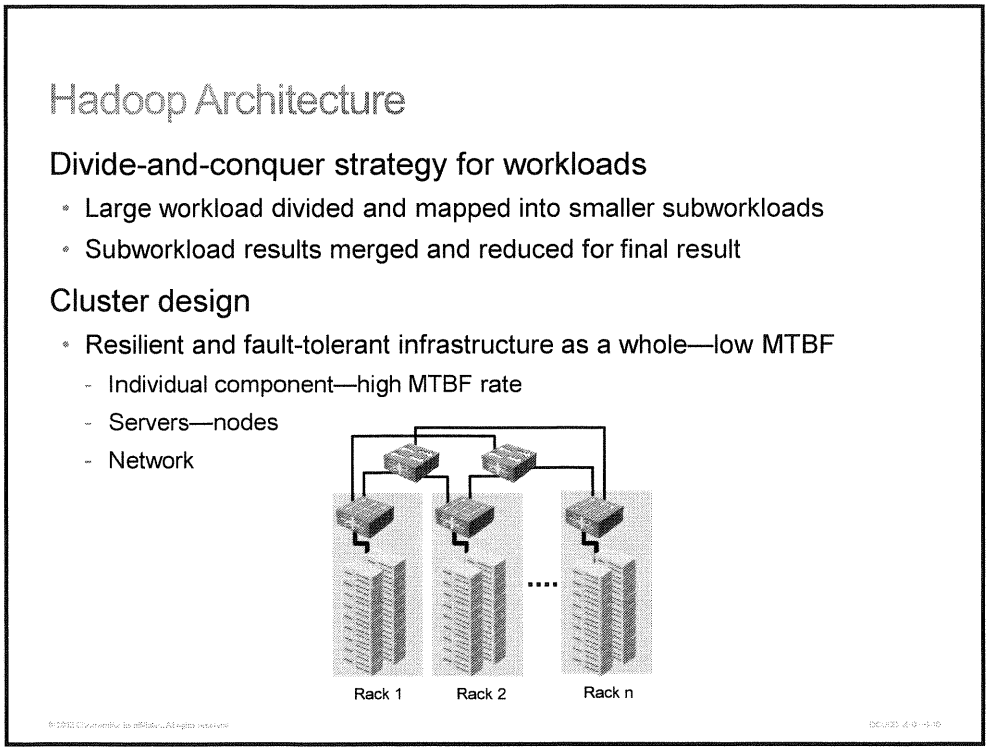
Such column-oriented layout is a natural fit and very effective for storing very sparse data (such as cells having Null values) and also multicell values.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Hadoop Architecture

This topic identifies general Hadoop architecture.



Divide-and-conquer strategies can be quite effective for several kinds of workloads that manage massive amounts of data. A single large workload can be divided or mapped into smaller subworkloads, and the results from the subworkloads can be merged, condensed, and reduced in order to obtain the final result.

The idea behind Hadoop was to exploit this feature of the workload and assign the smaller subworkloads to a large cluster of inexpensive nodes built with general-purpose hardware, rather than use expensive, fault-tolerant hardware. Furthermore, processing massive amounts of data requires storing the same massive amounts of data.

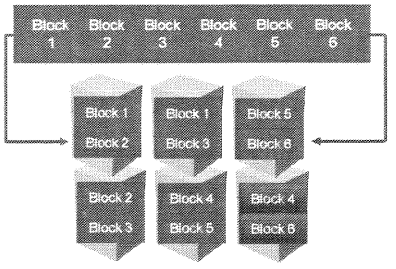
Hadoop has a distributed, cluster file system that scales to store these massive amounts of data. The cluster is built so that the entire infrastructure is resilient and fault-tolerant—even though individual components can fail—dramatically lowering the systemwide mean time between failures (MTBF) rate despite a higher component MTBF rate.

https://t.me/learningnets

Hadoop Distributed File System

Characteristics of data storage

- No central location—data stored across data nodes in the cluster
- Stored in large blocks—128 MB or larger
- Stored reliably by replication



The HDFS is the first building block of a Hadoop cluster. To efficiently process massive amounts of data, it is important to move computing to where the data is, thereby using a distributed file system rather than a central system for the data.

A single large file is split into blocks, and the blocks are distributed among the nodes of the Hadoop cluster.

The blocks used in HDFS are large, typically 128 MB or more, compared to the small blocks associated with traditional file systems. The larger blocks allow the system to scale without increasing the size and complexity of HDFS metadata.

Files in HDFS are "write once" files, meaning input data is streamed or loaded into HDFS and processed by the MapReduce framework (described later in this document), and any generated results are stored back in HDFS. The original input data is not modified during its life in HDFS. With this approach, HDFS is not intended to be used as a general-purpose file system.

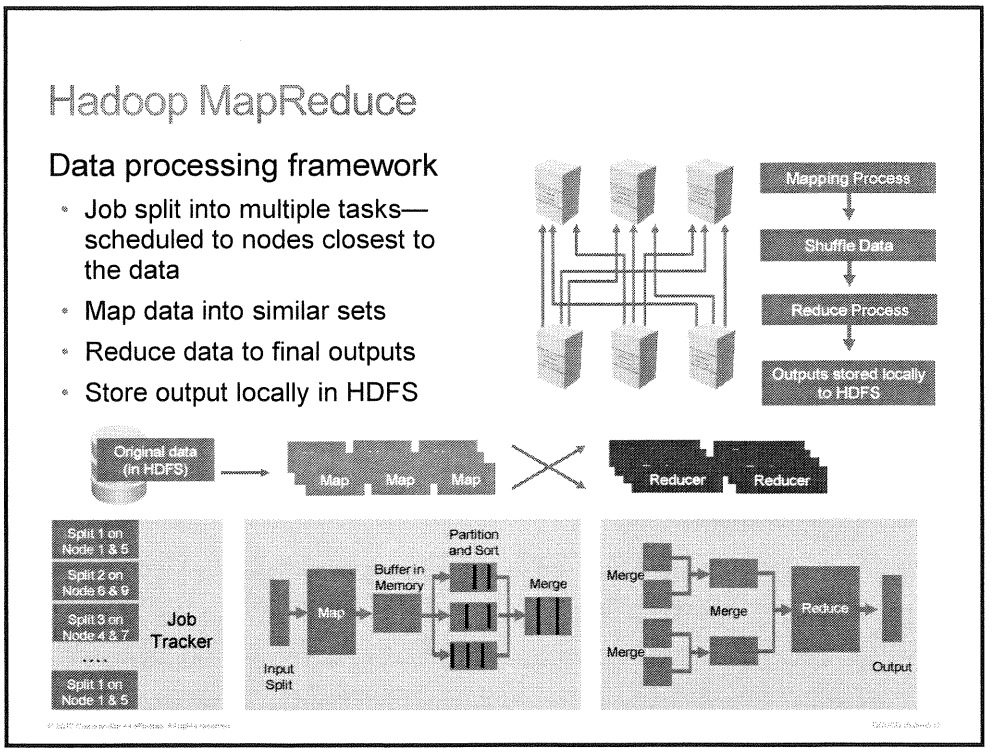
To enhance reliability and availability of the data in HDFS, the data assigned to one node is replicated among the other nodes. The default replication is threefold. This replication helps ensure that the data can survive the failure or nonavailability of a node.

The nodes in a Hadoop cluster serve one of the following HDFS functions:

- **Name node:** The name node is a single node in the cluster that is the brain of HDFS. It is responsible for keeping track of the file system metadata. It maintains a list of all of the blocks in an HDFS file and a list of data nodes that host these blocks. Because it can be a single point of failure, it is generally provisioned with a resilient, highly available server.
- **Data node:** The data node is a shared-nothing cluster of computers capable of executing the workload components. These nodes operate independently of each other and are built with general-purpose hardware storing the data blocks of workload within HDFS.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets



The MapReduce component of Hadoop is a framework for processing huge data sets on the Hadoop cluster. The MapReduce framework provides a clean abstraction of the underlying Hadoop cluster infrastructure so that programmers can use the power of the infrastructure without dealing with the complexities of a distributed system. MapReduce workloads can be divided into two distinct phases:

- **Map phase:** The submitted workload is divided into smaller subworkloads and assigned to mapper tasks. Each mapper processes one block of the input file. The output of the mapper is a sorted list of key-and-value pairs. These key-value pairs are distributed or shuffled to reducers.
- **Reduce phase:** The input for the reduce phase is the list of key-value pairs received from mappers. The job of a reducer task is to analyze, condense, and merge the input in order to produce the final output. The final output is written to a file in HDFS.

Relationship Between MapReduce and HDFS

The unit of input for a map task is an HDFS data block of the input file. The map task functions most efficiently if the data block it has to process is available locally on the node on which the task is scheduled. This approach is called HDFS data localization. An HDFS data locality miss occurs if the data needed by the map task is not available locally. In such a case, the map task will request the data from another node in the cluster. Such an operation is expensive and time-consuming, leading to inefficiencies and delay in job completion.

Clients, Data Nodes, and HDFS Storage

Input data is uploaded to the HDFS file system in one of two ways:

- An HDFS client has a large amount of data to place into HDFS.
- An HDFS client is constantly streaming data into HDFS.

Both these scenarios have the same interaction with HDFS, except that in the streaming case, the client waits for enough data to fill a data block before writing to HDFS. Data is stored in HDFS in large blocks of generally 64 to 128 MB or more in size. This storage approach allows easy parallel processing of data.

https://t.me/learningnets

During the process of writing to HDFS, the blocks are generally replicated to multiple data nodes for redundancy. The number of copies, or the replication factor, is set to a default of three and can be modified by the cluster administrator. When a new data block is stored on a data node, the data node initiates a replication process to replicate the data onto a second data node. The second data node, in turn, replicates the block to a third data node, thus completing the replication of the block.

Anatomy of a MapReduce Job

MapReduce workloads are split into two distinct phases: a map phase and a reduce phase.

The following are the broad steps in running a MapReduce job:

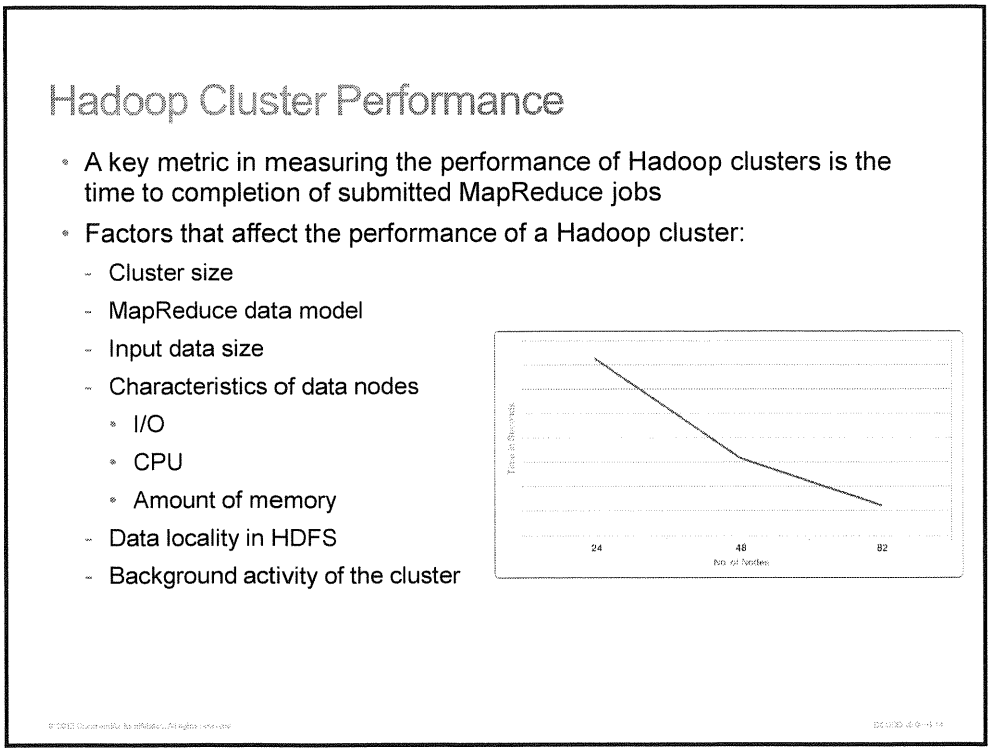
1. **Input setup:** Load the input data into HDFS.
   - The input data can be bulk loaded or streamed into HDFS.
   - Input data is split into large blocks (128 MB or larger) and distributed to data nodes in the cluster.
   - The blocks are replicated to help ensure availability of the input data in the event of failures in the cluster during MapReduce processing.
   - The name node keeps track of the list of data nodes and the blocks they hold.
2. **Job setup:** Submit a MapReduce job to the JobTracker. A job definition consists of:
   - Path to the input file in HDFS.
   - Path to the output file in HDFS in which the results should be stored.
   - Class that defines the map function.
   - Class that defines the reduce function.
   - Driver code that sets up the MapReduce job.
3. **Job initialization:** The JobTracker interacts with the TaskTracker on each data node to schedule the map and reduce tasks, consisting of:
   - The JobTracker interacts with the name node to get a list of data nodes that hold the blocks of the input file. It then schedules map tasks on these data nodes.
   - The JobTracker also schedules reducer tasks on nodes in the cluster.
4. **Map phase:** The mapper processes input splits, or chunks of an HDFS block of an input file, and generates intermediate key-value pairs.
5. **Sort phase:** The mapper performs a sort of the intermediate key-value pairs. Each mapper then partitions the intermediate data into smaller units—one unit per reducer—typically using a hash function for the partitioning.
6. **Shuffle phase:** In *Hadoop: The Definitive Guide* (2010), shuffle is described as follows: "MapReduce makes the guarantee that the input to every reducer is sorted by key. The process by which the system performs the sort—and transfers the map outputs to the reducers as inputs—is known as the shuffle."
7. **Reduce phase:** Each reducer merges all the units received from mappers and processes the merged list of key-value pairs to generate the final result.
8. **Result storage and replication:** The results generated by reducers are stored as files in HDFS. This HDFS write operation again triggers replication of blocks of the result file for the sake of redundancy.
9. **Result extraction:** A client reads the HDFS file to export the results from HDFS.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

Hadoop Performance Characteristics

This topic discusses the general Hadoop performance characteristics.



A key metric in measuring the performance of Hadoop clusters is the time to completion of submitted MapReduce jobs. Numerous factors can affect the performance of a Hadoop cluster. Some of these factors, in order of their impact on performance, are:

- Cluster size
- MapReduce data model
- Input data size
- Characteristics of data nodes
  - I/O
  - CPU
  - Amount of memory
- Data locality in HDFS
- Background activity of the cluster

https://t.me/learningnets

Cluster Size

Most Hadoop clusters start out with a modest size and grow as the demand on the cluster increases. You can increase the cluster size on an as-needed basis because the Hadoop infrastructure can exploit the parallel nature of MapReduce algorithms. As the number of jobs in the cluster grows, or as the size of the input data increases, you can, in most cases, simply add more nodes to the cluster and scale almost linearly. When nodes are added to the cluster, you must ensure that the network infrastructure and the name node can scale to match the increased size of the cluster.

A general characteristic of an optimally configured cluster is the capability to decrease job completion times by scaling out the nodes. For example, doubling the number of servers should decrease the completion time by roughly half for the same job and data set size.

MapReduce Data Model

MapReduce jobs consume and produce data. They also produce intermediate data (key-value pairs). The amount of input data, result data, and intermediate data depends on the purpose and structure of the MapReduce algorithm. Typically, an extract, transform, and load (ETL)-type workload produces the same amount of intermediate data and result data that it consumes as input. This behavior occurs because the ETL algorithms simply transform the input data without adding new data to the input data set. However, some algorithms do generate much smaller amounts of intermediate and result data than the amount of input data. For example, business intelligence workloads typically are in this category. Business intelligence algorithms process a large amount of input data and extract a small but very valuable set of intelligence data from the input data. This feature of the data model of the algorithm dictates the amount of intermediate data that is produced, sorted, and shuffled by mappers and consumed by reducers. The more intermediate and result data there is, the longer the job takes to complete.

The MapReduce algorithm structure can have a significant impact on job completion time. It is possible to give too much of the computation to mappers and not enough to the reducers (or the reverse), thus making the algorithm inefficient. For example, a business intelligence workload with very little computation performed in the mapping function will potentially generate more data for reducers to process, and hence more network traffic between mappers and reducers. It is important to divide the algorithm efficiently—keeping in mind the functional differences of the map and reduce phases—in order to achieve optimal results. Also, it is possible, for certain workloads, to create a pipeline of small MapReduce jobs instead of a single large MapReduce job, with the entire pipeline of jobs completing faster than the single job. The algorithmic complexity, the data model of the algorithm, the size of the input data set, and the characteristics of the cluster have to be taken into account in order to make these optimizations.

Input Data Size

Given the same algorithm, the larger the input data set, the longer it will take to complete the processing and produce the results.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Characteristics of Data Nodes

Nodes are the building blocks of a Hadoop cluster. The performance of the cluster is strongly influenced by the capabilities of the nodes in the cluster.

MapReduce workloads process large amounts of data—typically, a few gigabytes to as much as several terabytes—while the entire cluster may store data into multiple petabytes. Even with many nodes in a cluster and decent data distribution among the nodes, each node must be able to read and write very large amounts of data to and from HDFS on the local disk. With typical disk data (SATA) transfer rates of about 80 Mb/s, the time required to read the HDFS blocks needed by the MapReduce algorithm can be significant and can contribute to the total job time.

Many MapReduce jobs can be very CPU- and memory-intensive. The total processing capacity (CPU architecture, number of cores, and processing power of the cores) and memory (amount of memory and memory latency) available for such jobs dictates the speed with which the map or reduce phase can be completed. Faster processing and more memory generally decrease the amount of time needed for job completion.

Note The amount of memory allocated per data node should depend on the importance of job-completion time considerations and price-to-performance tradeoffs for those gains. If a Hadoop cluster is used for a single job, it is sufficient if a data node has adequate memory to run that single job. However, most Hadoop clusters run multiple jobs at a time. In such cases, having more memory will allow more tasks to be accommodated on the data node, thereby increasing the possibility of faster job completion.

When determining the characteristics of the data node, you need to consider the benefits and costs of the available options. For example, would it be more appropriate to use a 1-rack unit (1 RU) server with eight disks and two CPUs, or a 2-RU server with 16 disks and two CPUs? A 2-RU server with 16 disks gives the node more storage space and more disk transfer parallelism, helping it to cope better with workloads that are I/O-intensive. However, a 2-RU server also uses more space, thereby decreasing the number of CPUs that can be fit into a single rack (CPU rack density), a factor that can increase the number of racks in a cluster and the total infrastructure needs (space, power, cooling, and networking resources) of the cluster. From the perspective of a top-of-rack (ToR) networking device, the ports on the device could be sparsely used, leading to the use of more ToR devices and, hence, more core ports and devices. A 1-RU server, while limited in storage, offers greater CPU density and a lower per-server power footprint. More 1-RU servers can be fit into a single rack, thus requiring a higher-density ToR switch or fabric extender. It is useful to consider some of these tradeoffs in the context of whether the applications are CPU-bound or I/O-bound. A CPU-bound application might benefit from greater CPU density per rack, whereas an I/O-bound application might benefit from larger storage capacity and greater disk transfer parallelism.

Data Locality in HDFS

A task that is scheduled on a node can finish more quickly if the data it requires is available locally. The Hadoop JobTracker has enough intelligence power to schedule jobs optimally on nodes on which the data is available. Furthermore, because typical Hadoop clusters use a replication factor of greater than one (>1), the JobTracker has more than one node to choose from in order to schedule the task. However, in some instances, all the nodes that host a particular data block may be experiencing heavy load. In such instances, the JobTracker will be forced to schedule tasks in any node that is not experiencing heavy load. Because the data blocks needed by the task are not available locally, the task will have to request the data block from a node that has the data. Every such data transfer consumes resources on the node requesting the data, on the node providing the data, and on the network, which can lead to slower job completion.

https://t.me/learningnets

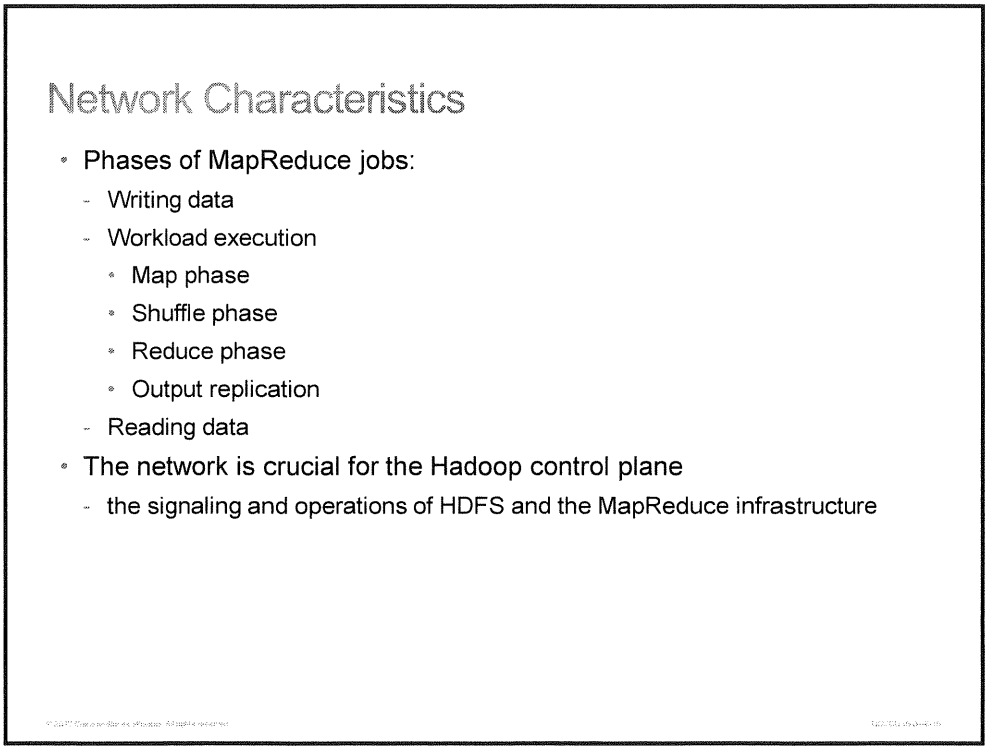
Background Activity of the Cluster

Hadoop clusters are typically used for numerous MapReduce workloads for jobs that run in parallel. The total existing activity of the cluster influences the resources available for any newly submitted jobs because the newly submitted jobs have to be scheduled alongside existing jobs and will compete for task slots within the cluster. Although you can increase the priority of the new job, this change will adversely affect the other existing jobs. Background activity can also impair the ability of a cluster to accelerate jobs by applying strategies such as speculative execution, through which the Hadoop infrastructure can speed up job completion by aggressively and intelligently spawning multiple instances of the same task and using the results only from the task that finishes first, then discarding results from the rest. Less sharing of the cluster among MapReduce jobs will lead to better performance for the scheduled jobs.

Other background activities, such as importing and exporting data from HDFS, and internal HDFS operations such as data rebalancing when nodes are added or removed from the cluster, can consume the resources of the cluster and affect the completion times of submitted jobs.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets



The nodes in a Hadoop cluster are interconnected through the network. Typically, one or more of the following phases of MapReduce jobs transfers data over the network:

1. **Writing data:** This phase occurs when the initial data is either streamed or bulk-delivered to HDFS. Data blocks of the loaded files are replicated, transferring additional data over the network.

2. **Workload execution:** The MapReduce algorithm is run.

   — **Map phase:** In the map phase of the algorithm, almost no traffic is sent over the network. The network is used at the beginning of the map phase only if an HDFS locality miss occurs (the data block is not locally available and has to be requested from another data node).

   — **Shuffle phase:** This is the phase of workload execution in which traffic is sent over the network, the degree of which depends on the workload itself. Data is transferred over the network when the output of the mappers is shuffled to the reducers.

   — **Reduce phase:** In this phase, almost no traffic is sent over the network because the reducers have all of the data they need from the shuffle phase.

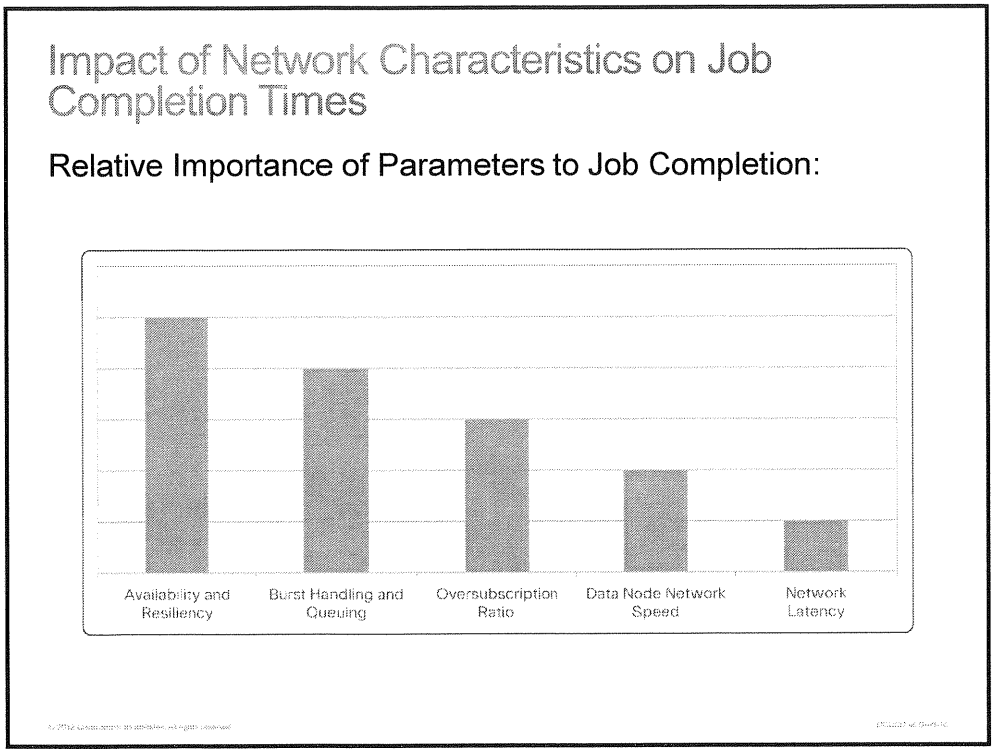
   — **Output replication:** MapReduce output is stored as a file in HDFS. The network is used when the blocks of the result file have to be replicated by HDFS for redundancy.

3. **Reading data:** This phase occurs when the final data is read from HDFS for consumption by the end application, such as the website, indexing, or Structured Query Language (SQL) database.

In addition, the network is crucial for the Hadoop control plane—the signaling and operations of HDFS and the MapReduce infrastructure.

Be sure to consider the benefits and costs of the choices available when designing a network: network architectures, network devices, resiliency, oversubscription ratios, and so on.

https://t.me/learningnets



A functional and resilient network is a crucial part of a good Hadoop cluster. However, an analysis of the relative importance of the factors shows that other factors in a cluster have a greater influence on the performance of the cluster than does the network. Nevertheless, you should consider some of the relevant network characteristics and their potential effects. The figure shows the relative importance of the primary parameters.

Availability and Resiliency

The failure of a networking device can affect multiple data nodes of a Hadoop cluster. The tasks on the affected nodes may then need to be rescheduled on other functioning nodes, thus increasing the load on the functioning nodes. Furthermore, the Hadoop infrastructure may start certain maintenance activities (such as data rebalancing and replication factor) to compensate for the failed nodes, thereby increasing the load on the cluster even more. These events are critical in the degradation of the cluster performance. Running jobs will take a lot longer to complete, and the ability to schedule new jobs will be impaired.

It is important to build a network that is available and resilient. First, you should focus on the network architecture: to deploy architectures that provide the required redundancy and that can also scale as the cluster grows. Technologies that allow network designs with multiple redundant paths between the data nodes are inherently better than technologies that have one or two points of failure alone.

After the architectural framework is laid out, you should consider the availability aspects of individual devices. Switches and routers that run operating systems proven in the industry to be resilient provide better network availability to servers. Switches and routers that can be upgraded without any disruption to the data nodes provide higher availability. Furthermore, devices that are proven to be easy to manage, troubleshoot, and upgrade help ensure less network downtime and also increase the availability of the network and, hence, the cluster.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Burst Handling and Queuing Depth

Several HDFS operations and phases of MapReduce jobs are bursty:

- Loading input files into HDFS or writing result files to HDFS uses the network resources. In addition, these write operations trigger replication of data blocks of the file, leading to higher network use. These operations, because they occur in a short period of time, show up as bursts of traffic in the network.
- The output of mappers, as they complete, is shuffled to reducers over the network. If many mappers finish at the same time, they will try to send their output to reducers simultaneously, leading to bursts of network traffic.

A network that cannot process bursts effectively will drop packets, so optimal buffering is needed in network devices to absorb bursts. Any packet dropped because a buffer is not available results in a retransmission. If the retransmissions are excessive, this can lead to longer job completion times. Be sure to choose switches and routers with architectures that employ buffer and queuing strategies that can manage bursts effectively.

Oversubscription Ratio

A good network design will consider the possibility of unacceptable congestion at critical points in the network under realistic loads. A ToR device that accepts 20 Gb/s of traffic from the servers but that has only two 1-Gb/s uplinks (a total of 2 Gb/s) provisioned (a 20:2 [or 10:1] oversubscription ratio) can drop packets, leading to poor cluster performance. However, overprovisioning the network can be costly. Generally accepted oversubscription ratios are around 4:1 at the server access layer and 2:1 between the access layer and the aggregation layer or core. Lower oversubscription ratios can be considered if higher performance is required. You should consider how oversubscription increases when certain devices fail and be sure to provision critical points in the network (such as the core) adequately. Network architectures (Layer 2 multipathing technologies such as Cisco FabricPath or Layer 3 Equal-Cost Multipath [ECMP]) that deliver a linear increase in oversubscription with each device failure are better than architectures that degrade dramatically during failures.

Data Node Network Speed

Be sure to provision data nodes with enough bandwidth for efficient job completion. Also continue to keep in mind the price-to-performance ratio entailed in adding more bandwidth to nodes.

Recommendations for a cluster depend on workload characteristics. Typical clusters are provisioned with one or two 1-GB uplinks per data node. Cluster management is made easier by choosing network architectures that are proven to be resilient and easy to manage and that can scale as your data grows. The use of 10 Gb/s server access is largely dependent on the cost/performance tradeoff. The workload characteristics and business requirement to complete the job in required time will drive the 10 Gb/s server connectivity. As 10-Gb/s Ethernet LAN on motherboard (LOM) connectors become more commonly available on servers in the future, more clusters likely will be built with 10-GB Ethernet data node uplinks.

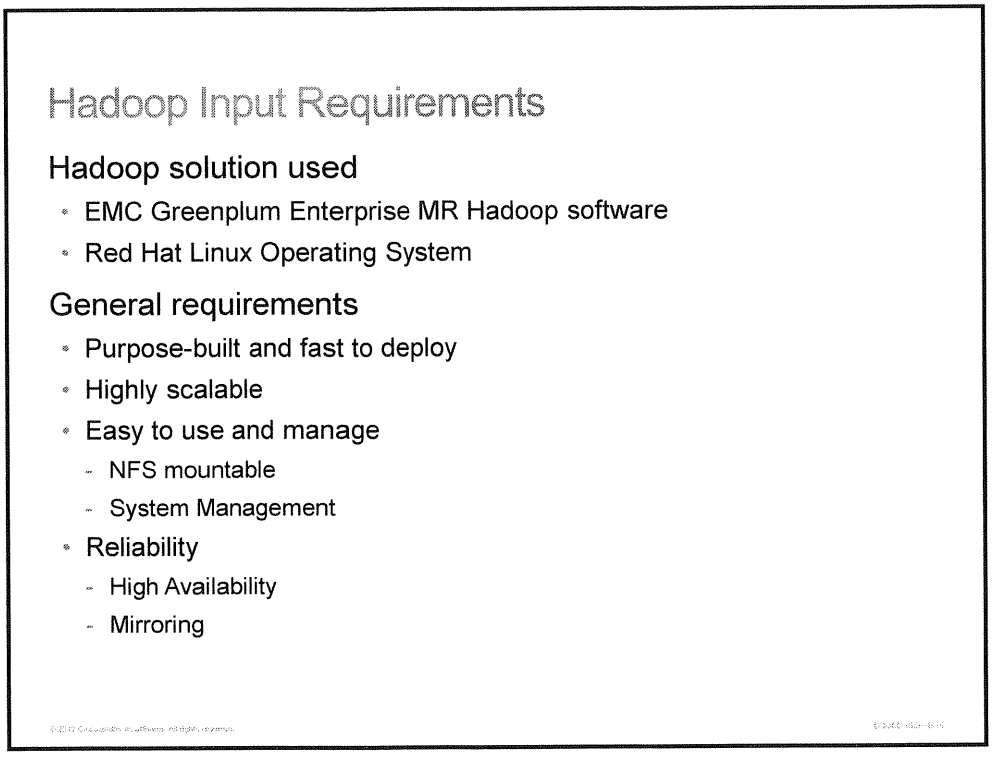
Network Latency

Variations in switch and router latency have been shown to have only limited impact on cluster performance. From a network perspective, any latency-related optimization should start with a networkwide analysis. "Architecture first and device next" is an effective strategy. Architectures that deliver consistently lower latency at scale are better than architectures with higher overall latency but lower individual device latency. The latency contribution to the workload is much higher at the application level, contributed to by the application logic (Java Virtual Machine software stack, socket-buffer, and so on) than network latency. Slightly more or less network latency will not noticeably affect job completion times.

https://t.me/learningnets

Designing Cisco UCS for Greenplum MR Hadoop

This topic describes the Cisco UCS design for the Greenplum MR Hadoop solution.

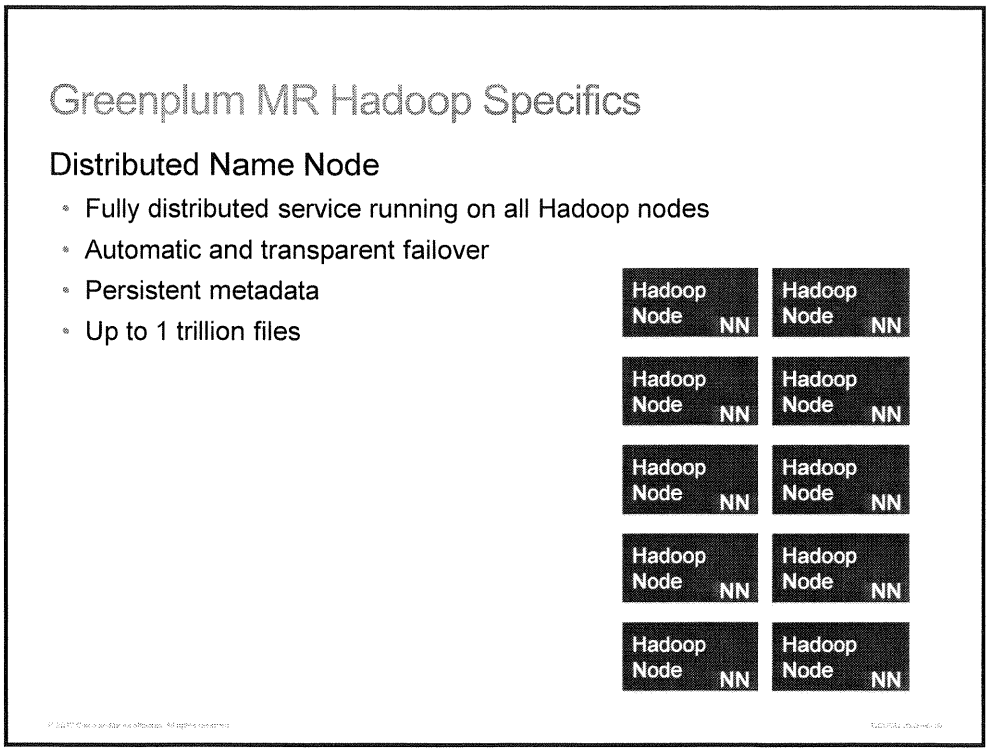


The Greenplum MR Hadoop solution is based on the MapR M5 distribution and is a completely reengineered implementation of the Apache Hadoop stack while preserving 100-percent compatibility. The solution enables near-real-time collection and organization of high volumes of structured, semi-structured, and unstructured data distributed across a cluster of servers while simultaneously delivering fast performance.

The solution enables you to take advantage of large data analytics without the overhead and complexity of a project built from the bottom up.. Available as software or in a preconfigured Data Computing Appliance Module, it provides a complete platform, including installation, training, global support, and value-add beyond simple packaging of the Apache Hadoop distribution. The Greenplum combines Hadoop and the Greenplum Database in a single, purpose-built Data Computing Appliance.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets



The architecture of the Greenplum Hadoop cluster incorporates distributed node name functionality. The software components include the following:

- **Lockless storage services** that represent replacement for the HDFS. The MapR lockless storage services provide multidimensional scalability and accelerate MapReduce performance. The services allow random read and write operations while automatically compressing data in real time.
- **MapReduce** part of the framework, which simplifies the creation of applications that process large amounts of unstructured and structured data in parallel. Underlying hardware failures are processed transparently for user applications, providing a reliable and fault-tolerant capability.
- **Hive** that serves as a data warehouse system for Hadoop that facilitates data summarization, queries, and analysis of large data sets. It has an SQL-like interface increase of the compression of stored data for improved storage resource utilization, but without affecting access speed.
- **Pig** that is a high-level procedural language for processing data sets in parallel using the Hadoop MapReduce platform. Its syntax simplifies the development of MapReduce jobs, providing an alternative programming language to Java.
- **HBase** is a distributed, versioned, column-oriented database that delivers random, real-time, read-write access to large data.
- **ZooKeeper** delivers a highly available system for coordinating distributed processes. Applications use ZooKeeper to store and mediate updates to important configuration information.
- **MapR Heatmap** provides visibility, access, and tools that give insight into the state of the cluster. Graphical and programmatic interfaces are designed to scale with the largest clusters.
- **MapR Control System** provides real-time monitoring of the cluster health, including alarms to notify you of conditions that need to be corrected. Alarms can also be configured to trigger email notifications.

https://t.me/learningnets

Greenplum MR Hadoop Specifics (Cont.)

Snapshots

Intelligent Snapshots

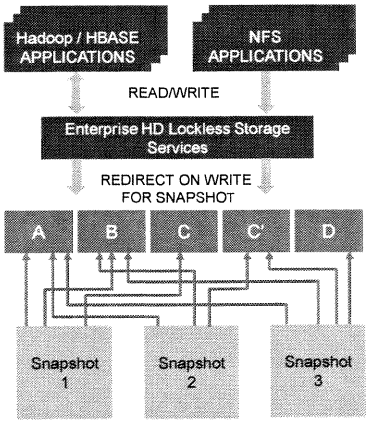
- Automatic data deduplication
- Block sharing for space savings

Fast and flexible

- Zero performance loss when writing to the original

Easy to manage

- Scheduled or on-demand
- Drag-and-drop recovery



Greenplum MR provides direct data input and output to the cluster with MapR Direct Access Network File System (NFS), real-time analytics, and is the first distribution to provide true high availability at all levels.

Greenplum MR introduces the concept of logical volumes to Hadoop: a means of grouping data and applying policy consistently across an entire data set.

Greenplum MR provides hardware status and control with the MapR Control System, a comprehensive user interface including a heatmap that displays the health of the entire cluster at a glance.

To scale and properly address performance, the intelligent snapshots are used within the solution.

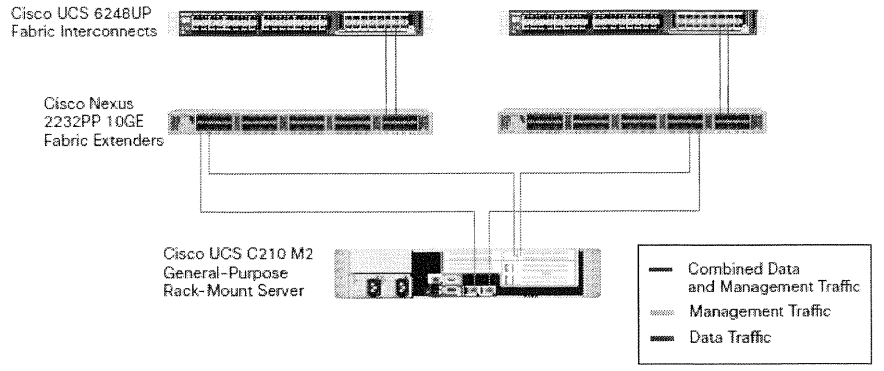
© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Node Connectivity and Management Architecture

Greenplum MR reference block

- Typical user capacity (compressed, three-way replicated) = 350 TB
- I/O bandwidth = 22.5 Gb/s



A fully integrated end-to-end hardware and software infrastructure for enterprise Hadoop deployments address performance, scalability, and advanced management capabilities.

The Greenplum Hadoop architecture consists of:

- Topology that defines how nodes are connected as defined in the figure
- Computing nodes

The Cisco UCS solution (a reference block) consists of:

- Two fully redundant Cisco UCS 6248UP 48-Port Fabric Interconnects and two Cisco Nexus 2232PP 10GE Fabric Extenders for node management and connectivity
- Integrated Cisco UCS C-Series rack mount servers for computing nodes
  - Per-node Cisco UCS P81E VIC for data traffic that connects to the unified fabric through two active-active 10-Gb/s links
  - Cisco Integrated Management Controller for node management
- Greenplum MR: Cisco UCS is the exclusive integrated platform for Greeplum MR, which can significantly reduce time-to-value and the operating expenses associated with Hadoop implementations.

Among other benefits, such a solution can connect across the same management plane to other Cisco UCS deployments running enterprise applications and thus simplify management and connectivity in the data center.

https://t.me/learningnets

Block Reference Architecture

Greenplum MR rack

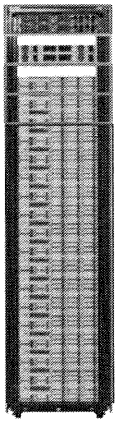
- Typical user capacity (compressed, three-way replicated) = 350 TB
- I/O bandwidth = 22.5 Gb/s

Nodes

Role	Component	Qty
GPMR Management	Name Node	1
	Secondary name node	1
GPMR Compute	Server nodes	16
	Processor cores	216
	Memory	1728 GB
	Unformatted storage	288 TB

Block management

Component	Qty
6248UP Fabric Interconnect	2/solution
Nexus 2232PP FEX	2/rack



6200 Series Fabric Interconnect

2232 Series Fabric Extender

Name Node (Master)

Name Node (Secondary)

Compute Nodes

Multirack architecture—up to 8 racks

The reference configuration starts in a single rack and can be expanded into multiple racks if required. A multirack configuration consists of components from a single rack and two additional Cisco Nexus 2232PP Fabric Extenders for every additional rack.

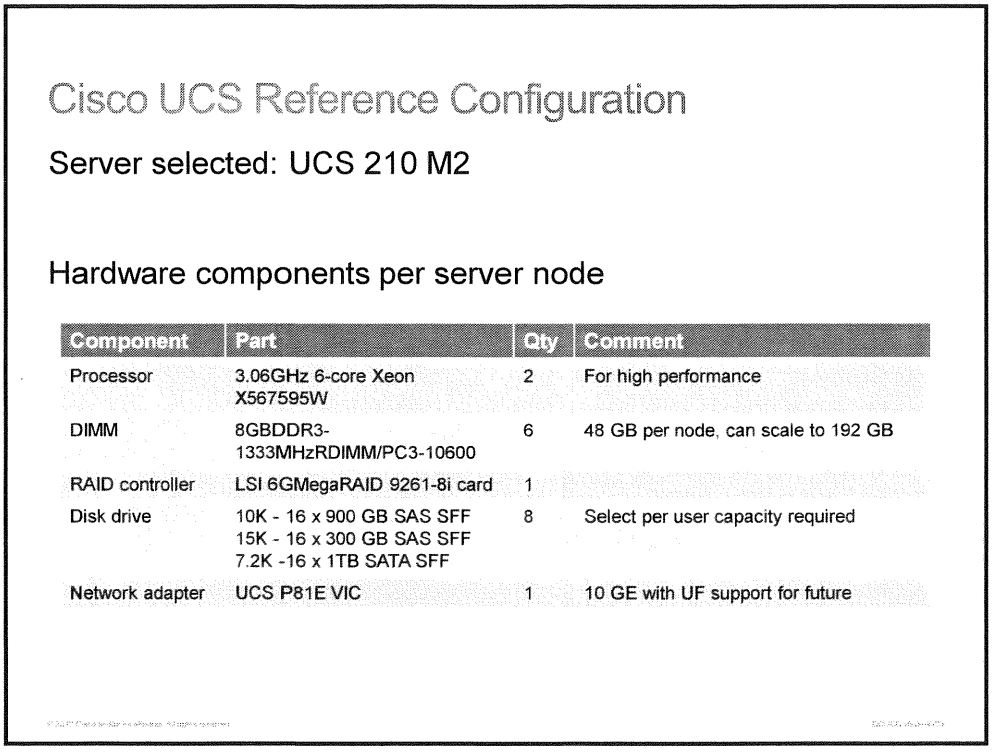
- **Single-rack configuration**: delivers 192 processor cores and a typical user data capacity of 320 TB.
- **Multirack configuration**: consists of components from a single rack with two additional Cisco Nexus 2232PP Fabric Extenders per rack added. Each rack adds 192 processor cores and 320 TB of capacity to the solution.

The complete block can fit different needs depending of the size and node types.

The benefits of reference block configuration are advanced performance, full data protection, no single point of failure, and improved data access features with expedited implementation of large data analytics environments.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets



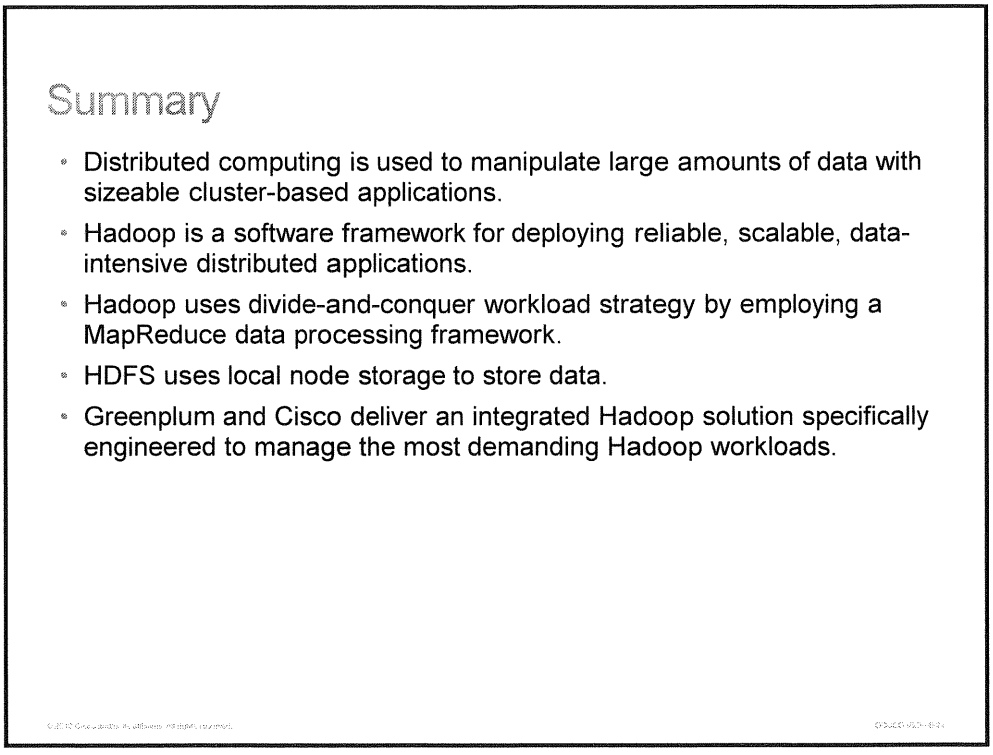
The reference architecture in the figure uses a Cisco UCS C210 M2 rack-mount server as a base cluster node in the following configuration:

- Two Intel Xeon X5670 processors (2.93 GHz),
- 96 GB of double data rate 3 (DDR3) memory
- Cisco UCS P81E VIC
- LSI 6G MegaRAID 9261-8i card
- 16 1-TB SATA SFF internal disk drives (for a total of 16 TB of storage)

https://t.me/learningnets

Summary

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



References

For additional information, refer to these resources:

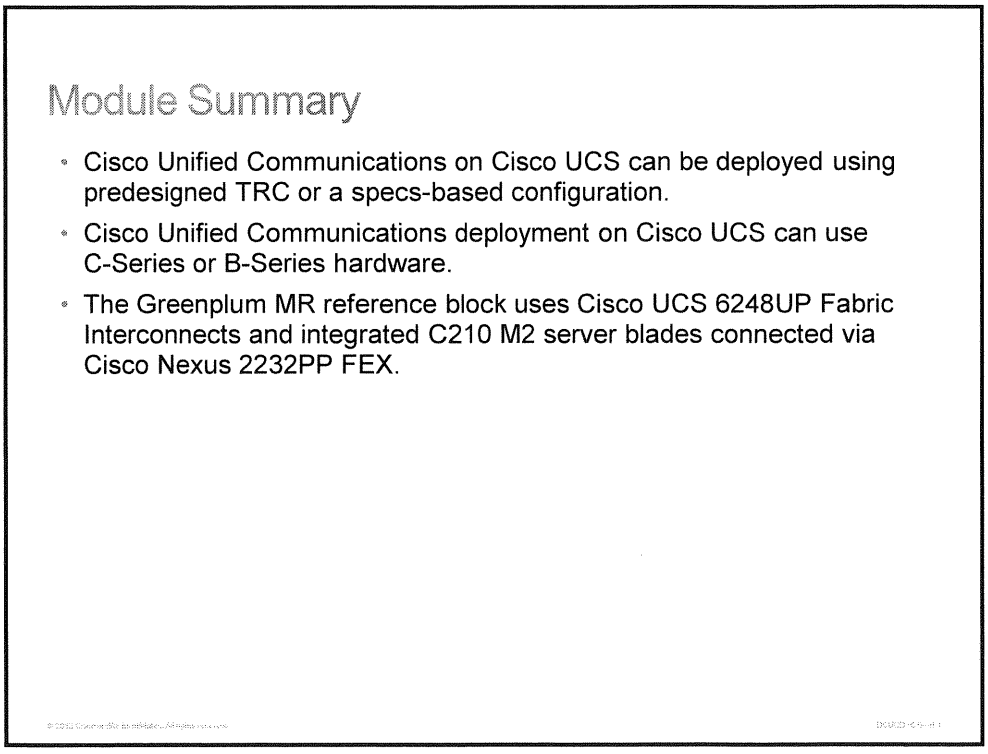
- http://hadoop.apache.org/
- http://www.greenplum.com/products/greenplum-mr
- http://www.cisco.com/en/US/solutions/collateral/ns340/ns517/ns224/ns944/wp_greenplum.pdf
- http://www.cloudera.com/
- http://en.wikipedia.org/wiki/BigTable
- http://en.wikipedia.org/wiki/NoSQL
- http://nosql-database.org/

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

Module Summary

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



https://t.me/learningnets

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Module Self-Check

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

Q1) Which Cisco UCS configuration is certified by Cisco? (Source: Designing Cisco Unified Communications on Cisco UCS)

A) specs-based
B) TRC
C) Cisco UCS C-Series rack mount server
D) Cisco UCS B-Series blade server

Q2) Which hardware component is used to connect the Greenplum MR Hadoop node data and management interfaces in a block? (Source: Designing Distributed Computing on Cisco UCS)

A) Cisco UCS 6248UP Fabric Interconnects
B) Cisco UCS C210 M2 LOM
C) Cisco Nexus 2232PP FEX
D) 10GE SR SFP+

https://t.me/learningnets

Module Self-Check Answer Key

Q1) B

Q2) C

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Appendix A

Creating Cisco UCS Solution BOM

Overview

This appendix describes the tools and process of creating the Bill of Materials (BOM) for the Cisco Unified Computing System (UCS) solution.

Objectives

Upon completing this appendix, you will be able to create the BOM for the Cisco UCS solution using the tools available. This ability includes being able to meet these objectives:

- Create a Cisco UCS C-Series server solution BOM with Online Configurator
- Create a Cisco UCS C-Series server solution BOM with Netformx DesignXpert
- Create a Cisco UCS B-Series server solution BOM with Netformx DesignXpert

https://t.me/learningnets

Create Cisco UCS C-Series Server BOM with Online Configurator

This topic describes how to create Cisco UCS C-Series server solution BOM with Online Configurator.



The BOM for the Cisco UCS C-Series can be created using one of the following two tools:

- Netformx DesignXpert
- Cisco Dynamic Configuration Tool

The Cisco Dynamic Configuration Tool is available at https://apps.cisco.com/qtc/config/jsp/configureHome.jsp and requires a Cisco Connection Online (CCO) account for access.

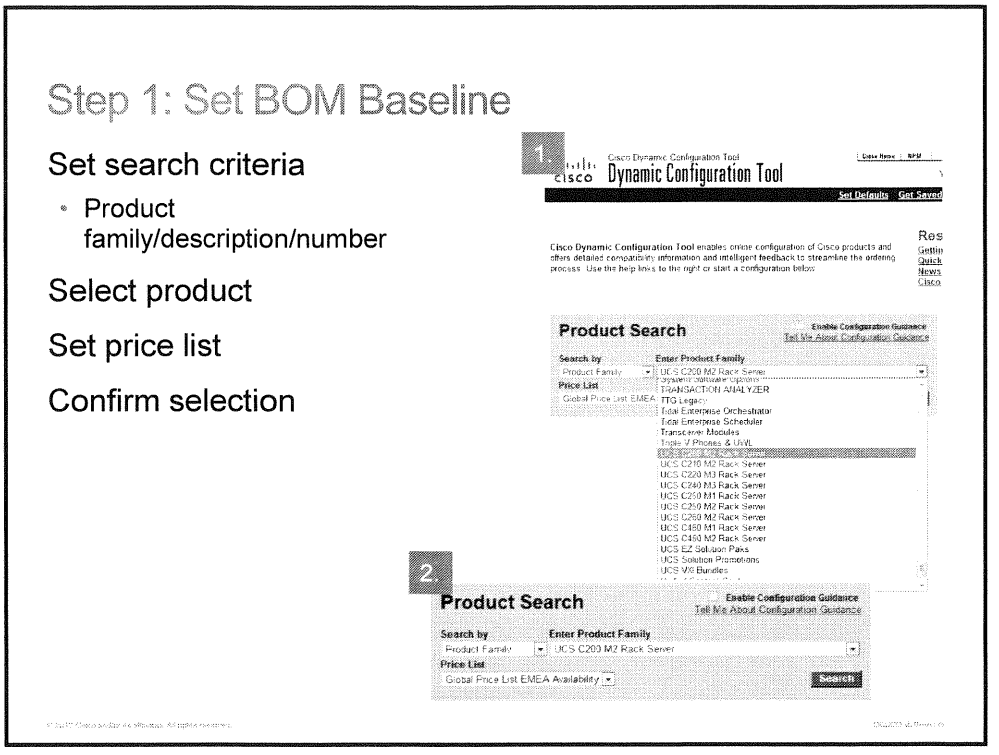
The Dynamic Configuration Tool can be used with a web browser and does not require special software to be installed. Next, you need to log in with your CCO account. The tool can also be used to create BOMs for other Cisco equipment.

The BOM creation process with the tool consists of the following steps:

- Selection of the baseline (that is, product type) on which the details are applied
- Selecting product details and setting quantities
- BOM review to verify parts and quantities
- Download BOM to proceed with order

© 2012 Cisco Systems, Inc.

https://t.me/learningnets



First, select the baseline, which means that a product type must be selected for which the detailed configuration is created.

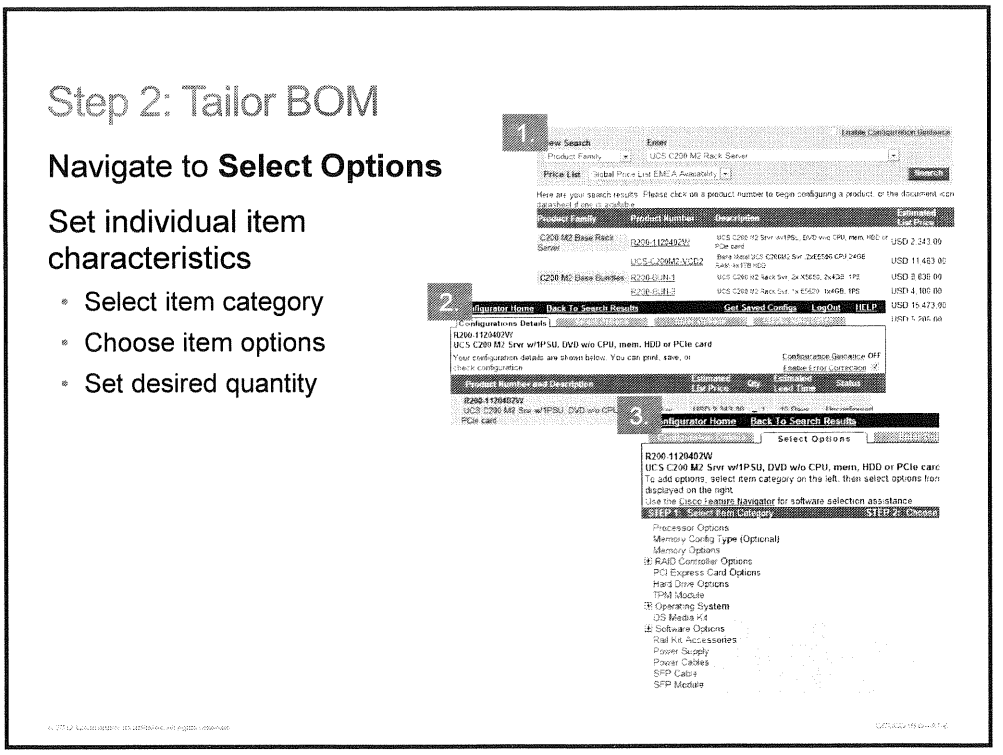
Selecting the baseline can be done in three options:

- **Product family:** The base product is selected from a drop-down list.
- **Product description:** The base product is found and selected based on the description—the information is entered in the field and then searched.
- **Product number:** The base product is selected based on the product ID, which can be acquired from the datasheet.

Before proceeding, select the price list to verify correct pricing.

Click **Search** to start the configuration project.

https://t.me/learningnets

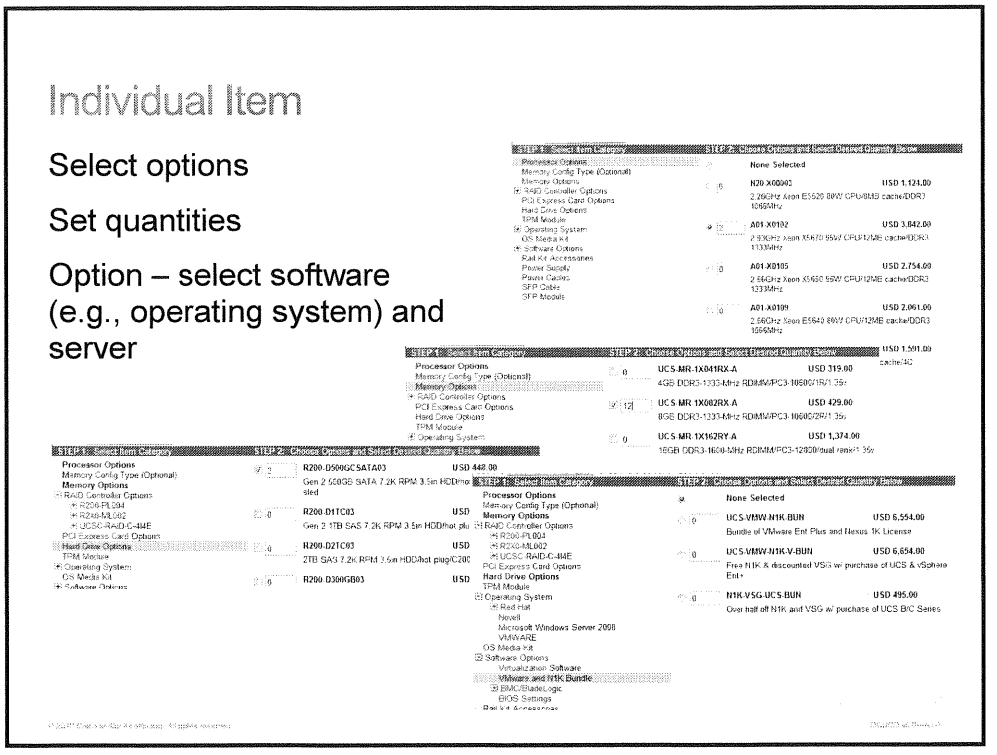


After you select the baseline product, select the details and apply them, and then set correct quantities.

For each server component, select the component type in the left pane, a part in the right pane, and then enter the quantity.

© 2012 Cisco Systems, Inc.

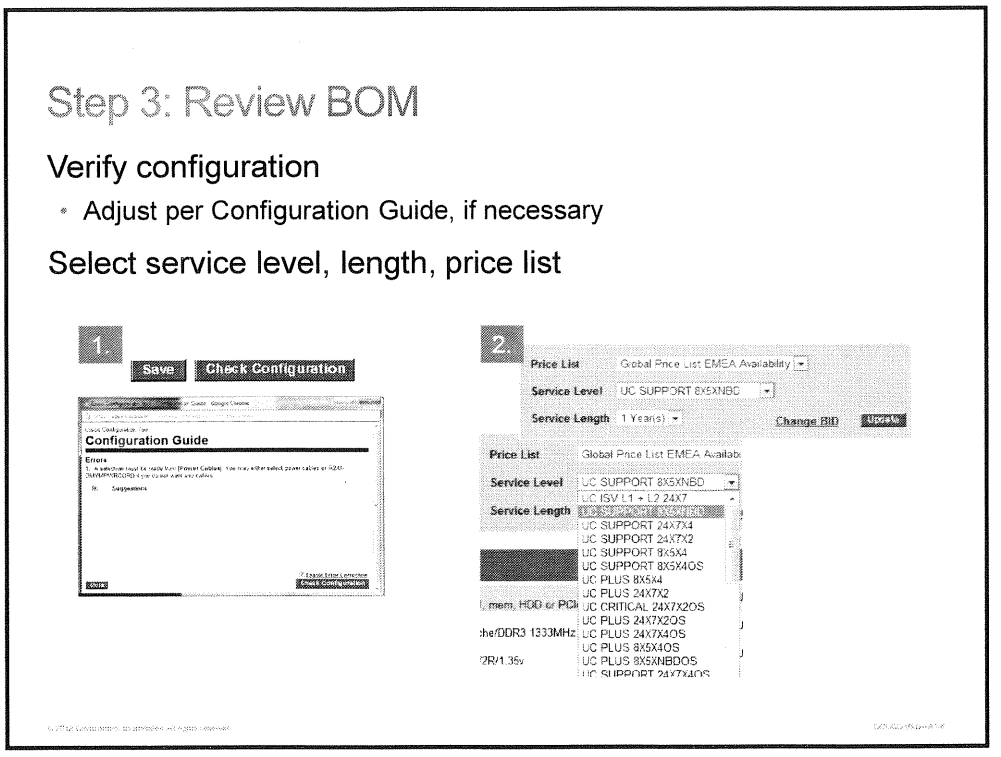
https://t.me/learningnets



For the Cisco UCS C-Series servers, the architect must select all the components as selected in the design process:

- **Processor:** Type and quantity
- **Memory config type:** Mirror or no
- **Memory options:** Proper DIMM size and quantity
- **RAID controller:** When it is required to deploy Redundant Array of Independent Disks (RAID) configurations other than 0, one RAID controller should be selected.
- **PCI Express Card:** Additional network interface cards (NICs), host bus adapters (HBAs), converged network adapters (CNAs)
- **Hard drives**
- **TPM module**
- **Operating system and additional software:** Operating system such as VMware vSphere, Microsoft Windows 2008 R2, BCM BladeLogic for management, Cisco Nexus 1000V for virtual network access layer, and so on
- **Rail kit**
- **Power supply and cables**
- **SFP modules and cables**

https://t.me/learningnets



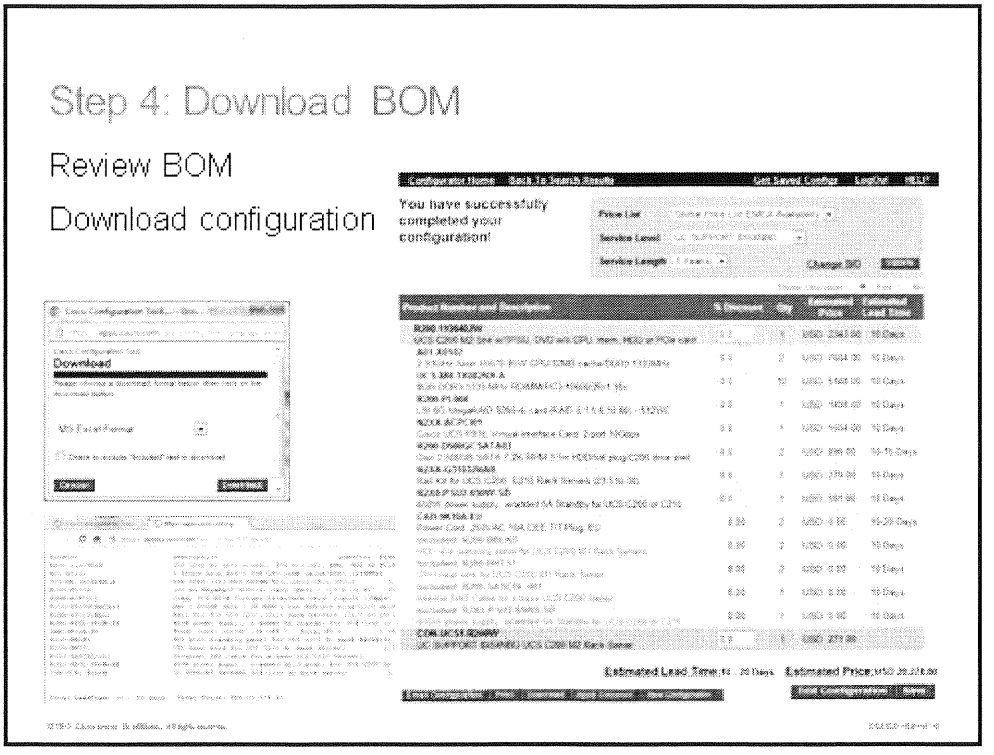
After selecting all the server parts, choose the created configuration. Click **Check Configuration**. The tool runs validation rules against the selected configuration and detects incompatibilities and errors.

You can correct the errors, re-run the configuration check, and, if the tool is correctly set, the final BOM is presented.

Then, choose a service level with length. You can also adjust the price list.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets



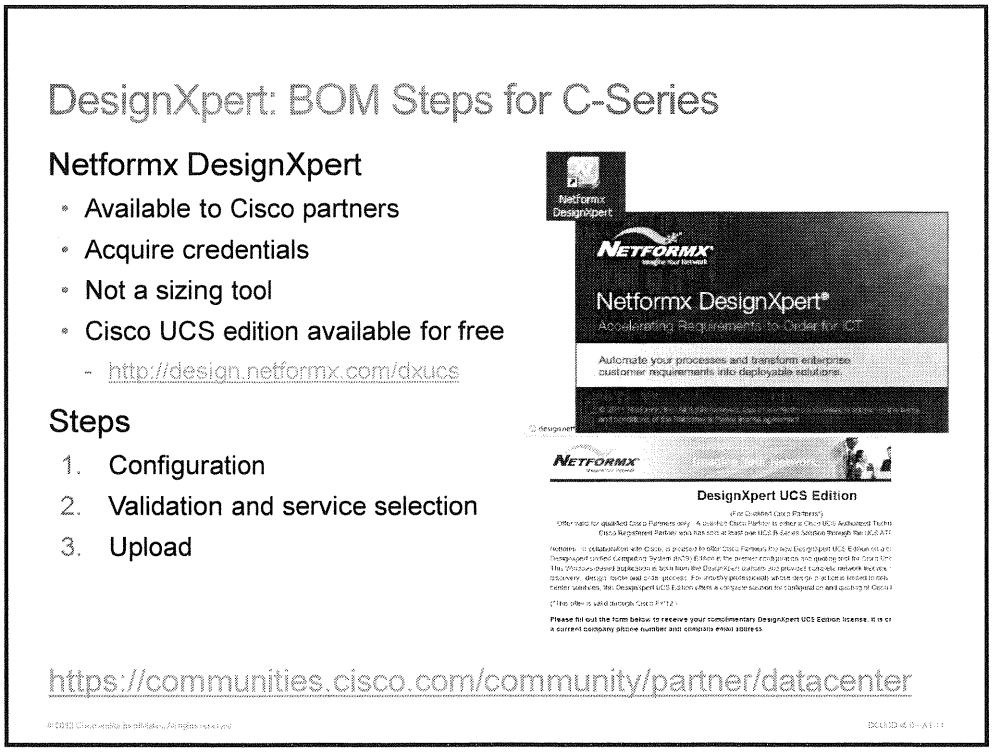
When the BOM is finalized, you can download it in different formats, such as Microsoft Office Excel, and use it to proceed with the order.

The finalized BOM also shows the estimated price and lead times. These are estimates for informational purposes only.

https://t.me/learningnets

Create Cisco UCS C-Series Server BOM with Netformx DesignXpert

This topic describes how to create Cisco UCS C-Series server solution BOM with Netformx DesignXpert.



The Netformx DesignXpert tool is available at http://design.netformx.com/cisco-navigate-to-accelerate/. In order to be able to use the tool to compile the BOM, you must have correct credentials, which are available to Cisco partners.

Netfromx DesignXpert must be installed on a Microsoft Windows machine, either on a physical workstation or server, or in a virtual machine. For correct operation, the tool requires internetwork connectivity, because the pricing is constantly updated and the BOM needs to be uploaded to Cisco to proceed with the order.

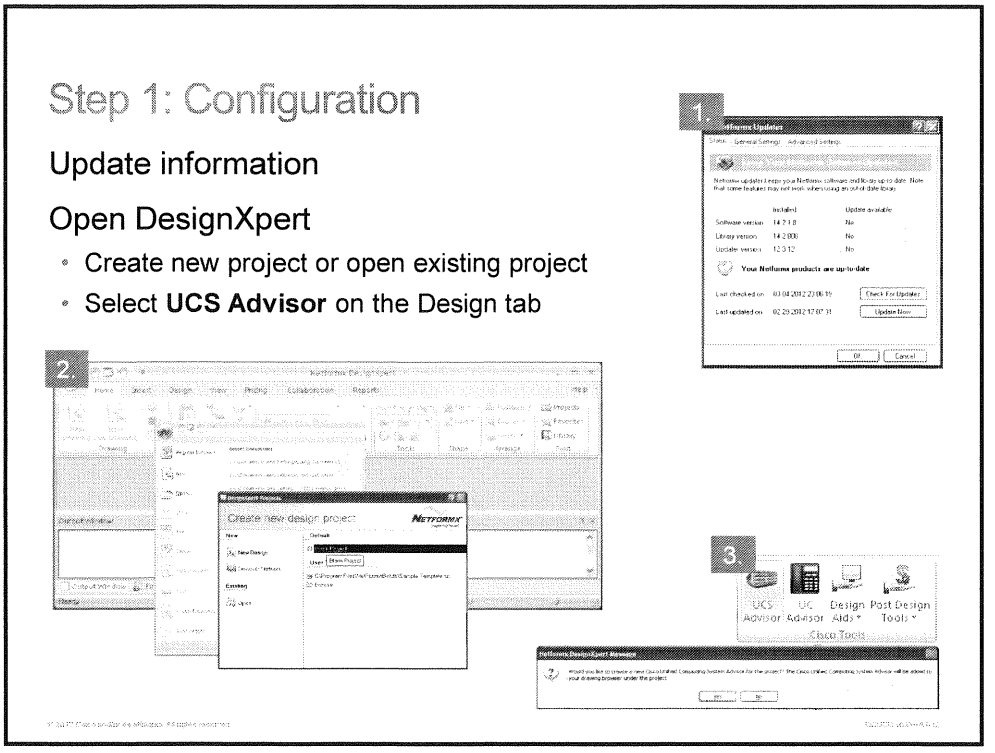
When the tool is installed with the correct credentials, the BOM can be created. This process has three steps:

- Configuration
- Validation and service selection
- Upload

More information about the Netformx DesignXpert tool is available at the Cisco Partner Resources Center at https://communities.cisco.com/community/partner/datacenter

© 2012 Cisco Systems, Inc.

https://t.me/learningnets



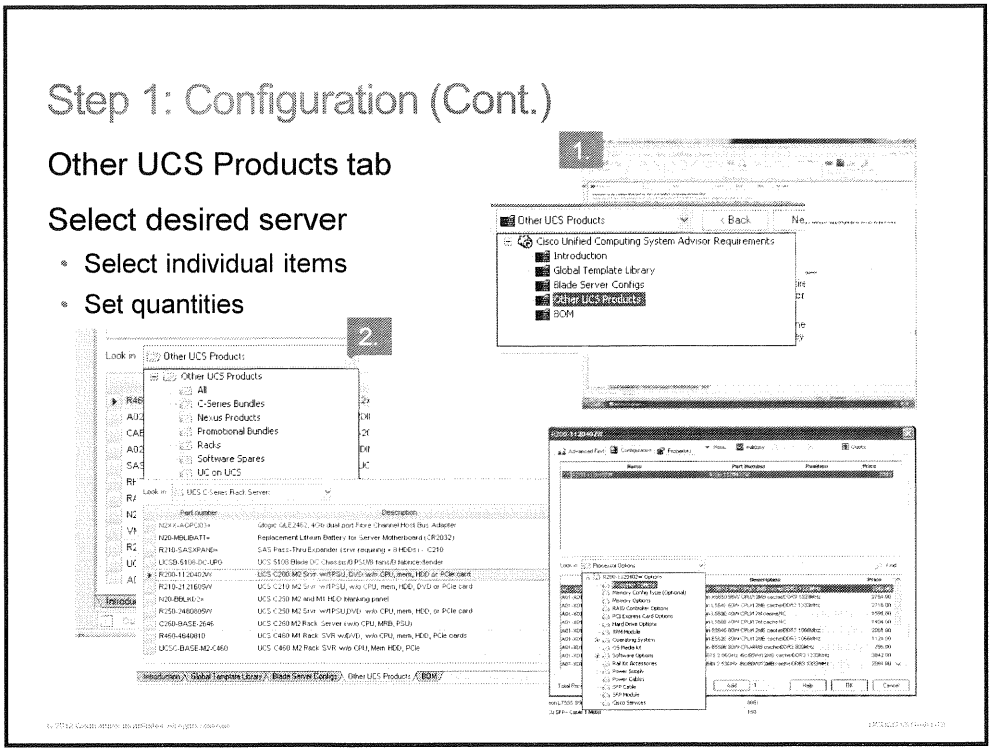
Although the pricing is constantly updated, it is a good practice to update the pricing before starting a new project or opening an existing one.

When starting a new project, start the DesignXpert tool.

When the tool is available, you can start a new project. Because the tool can be used not only for UCS BOMs, the project is general.

To proceed with the creation of a Cisco UCS BOM, navigate to the **Design** tab and click the **UCS Advisor** option, which starts the UCS BOM creation wizard.

https://t.me/learningnets



The configuration process employs the Cisco UCS Advisor. The last of the tabs available is "Other UCS Products."

On this tab, you can choose Cisco UCS C-Series server types and the quantity, and then configure them with the components such as CPU, memory DIMMs, PCIe adapters, RAID controller, disk drives, and so on.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

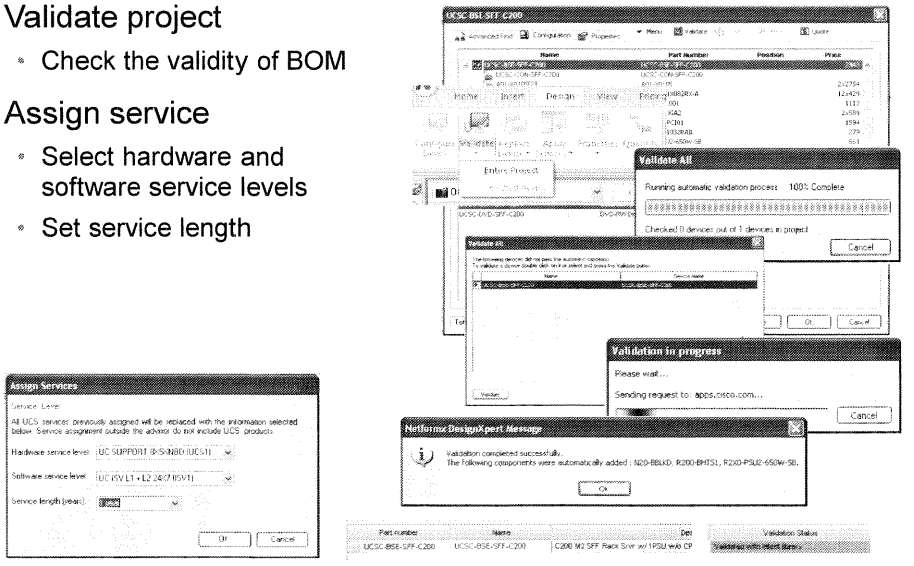
Step 2: Validation and Service Selection

Validate project

- Check the validity of BOM

Assign service

- Select hardware and software service levels
- Set service length

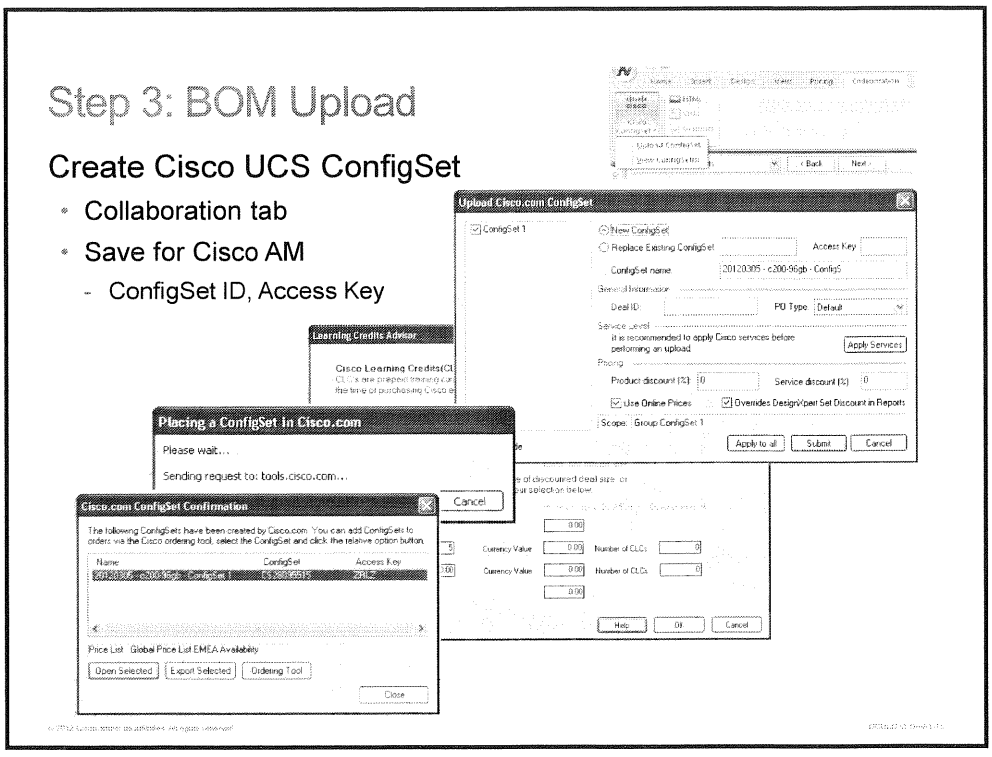


The service selection process is used to define the type of services that are to be applied to the selected Cisco UCS C-Series components. The services can be selected for the whole BOM or part of the BOM. The selection includes the definition of the hardware and software service level and service length.

Before proceeding to the final configuration step, which is uploading, the BOM should be reviewed to verify that it is correct.

The validation is accomplished with the **Validate** option under the **Design** tab. The process can show eventual incompatibilities and errors, which you can then correct. When validated, the status of the individual item in the table turns green, and says, "Validated with latest library."

https://t.me/learningnets



When validated with selected services, you can upload the BOM ConfigSet to Cisco.com, which makes it ready to be processed for order.

To create and upload the ConfigSet, select the **Cisco Configset => Upload ConfigSet** option under the **Collaboration** tab. The input window requires some additional information to be entered (such as product discount, and so on), which can be left at their defaults and can be later agreed upon with Cisco AM.

In the process of creating ConfigSet, you can apply Cisco Learning Credits, which can be also used for attending the courses.

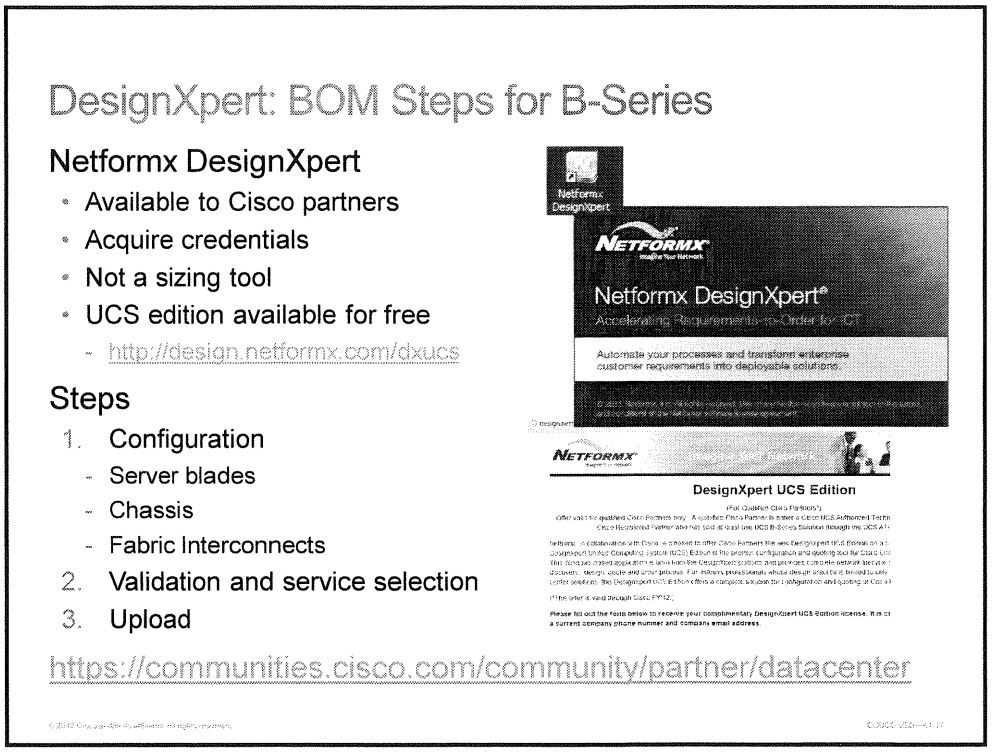
When the upload process is finished, the ConfigSet ID and Access Key are presented. Save them for later reference because they are required in the ordering process and when communicating with Cisco AM.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Create Cisco UCS B-Series Server BOM with Netformx DesignXpert

This topic describes how to create Cisco UCS B-Series server solution BOM with Netformx DesignXpert.

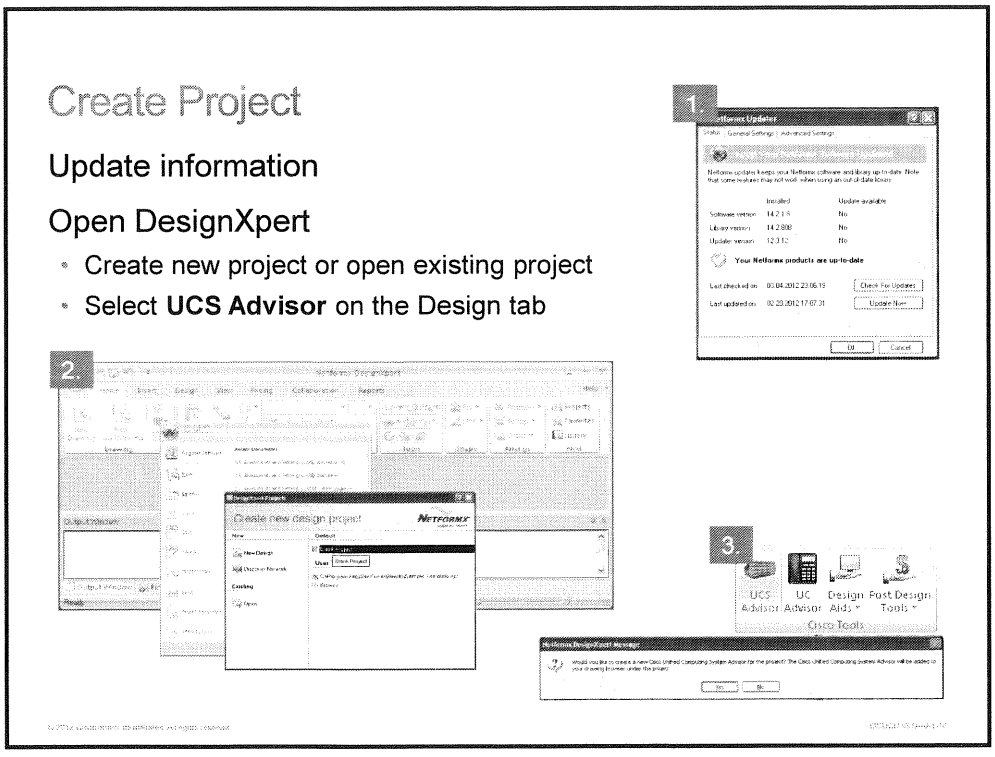


The Netformx DesignXpert used for creating the UCS B-Series BOMs is the same as for the C-Series servers.

The difference is that in the process of creating the BOM, there are more options to select with the B-Series because there are more components. In general the process is divided into the following steps:

- Configuration, which includes selecting components for server blades and setting their quantities, and configuring chassis and fabric interconnect characteristics and component quantities
- Validation and service selection
- Cisco ConfigSet upload

https://t.me/learningnets



To start creating a BOM for the Cisco UCS B-Series solution, you first must open a new project. If there is an existing BOM, you can also open and update that project per requirements.

You should update the pricing and items information prior to opening a new or existing project, which can be done with DesignXpert Updater.

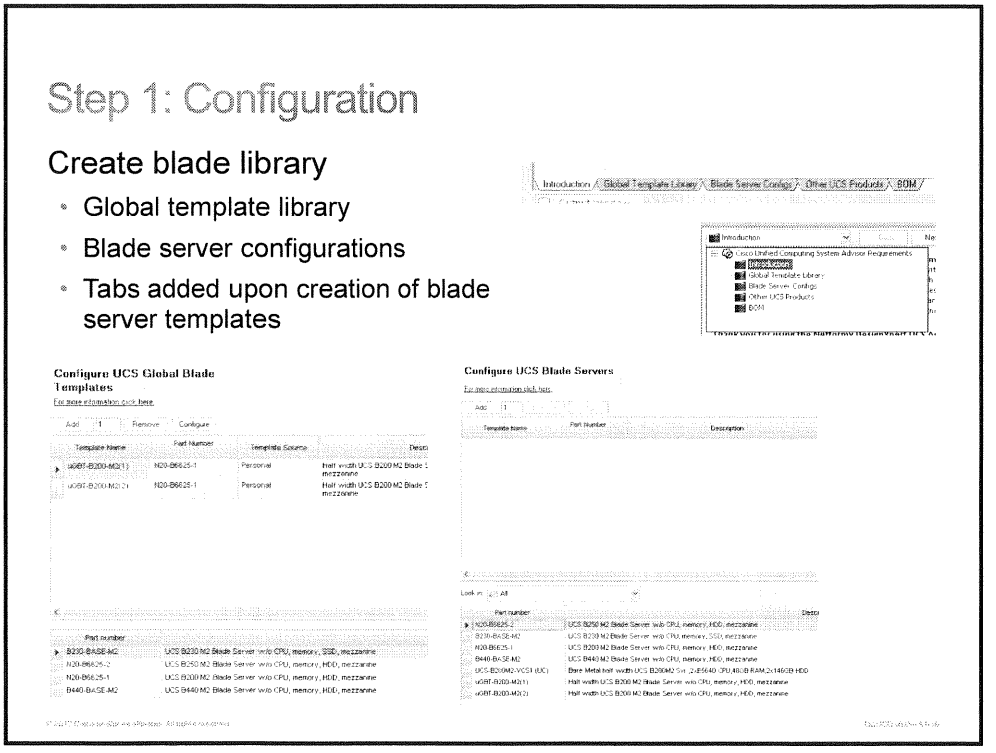
After a project is opened, navigate to the Design tab and select the **UCS Advisor** option to start the UCS BOM creation wizard.

The configuration process uses a bottom-up approach—first, the blades are configured, then the chassis is configured, and finally the fabric interconnects. The process consists of the following steps:

- Blade library
- Blade quantity
- Solution preferences
- Design summary

© 2012 Cisco Systems, Inc.

https://t.me/learningnets



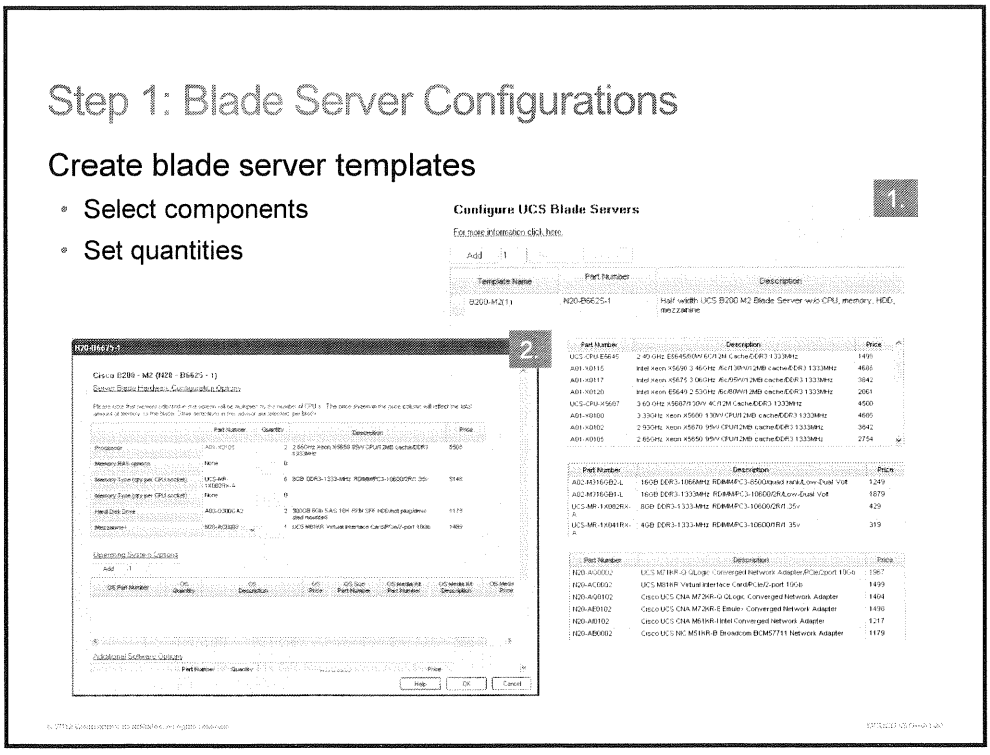
In the first step, to create the blade library, you can choose and configure blade templates or classes. Based on the requirements, blade types must be added to the library along with the quantity. An individual blade type can be configured per the requirements, which includes processors, memory, adapters, and so on. After the blade library configuration is finished and the library is populated with the desired blade types, the blade quantities can be entered.

There are also two additional options:

- Selecting the predefined blade templates from the global blade template library and adjusting the characteristics per requirements
- Using special bundle promotions, where blade configurations are predefined and cannot be changed or updated

When the blade library is defined, additional tabs appear, enabling you to configure the whole UCS B-Series solution (that is, chassis and fabric interconnects).

https://t.me/learningnets

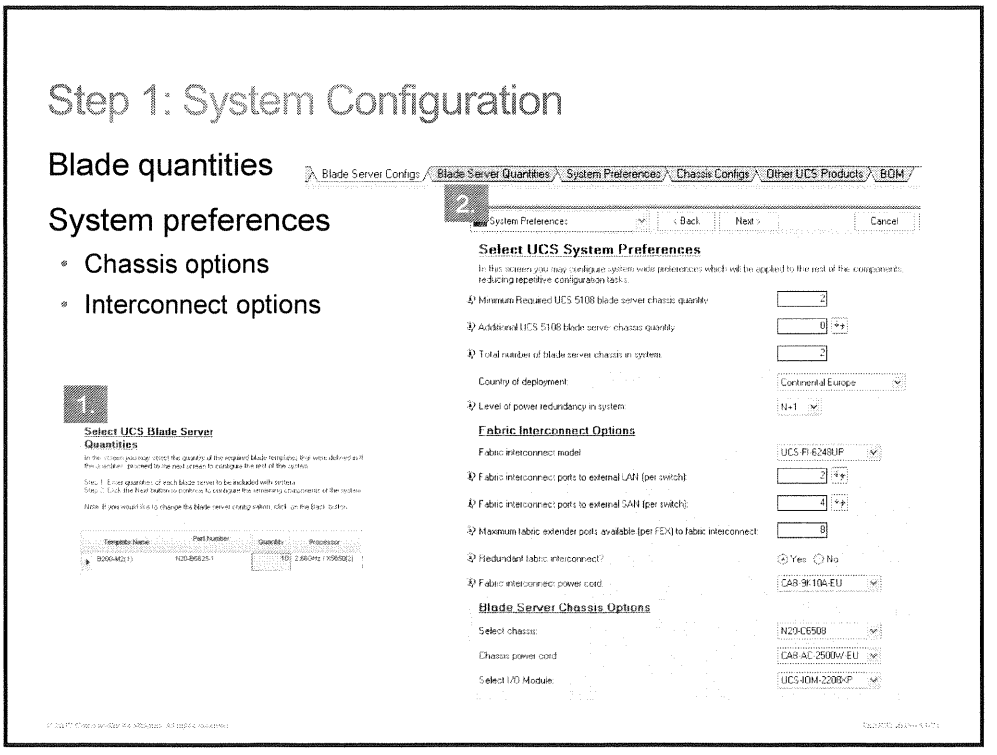


When defining the blade server template, you must define the server characteristics:

- Processor type and quantities
- Select or deselect memory mirroring option
- Select DIMM size, type, and quantity per processor
- Hard drives
- Mezzanine type and quantity
- Optionally, select operating system type and quantities
- Optionally, select additional software with the blade (for example, VMware vCenter Server or Cisco Nexus 1000V)

© 2012 Cisco Systems, Inc.

https://t.me/learningnets



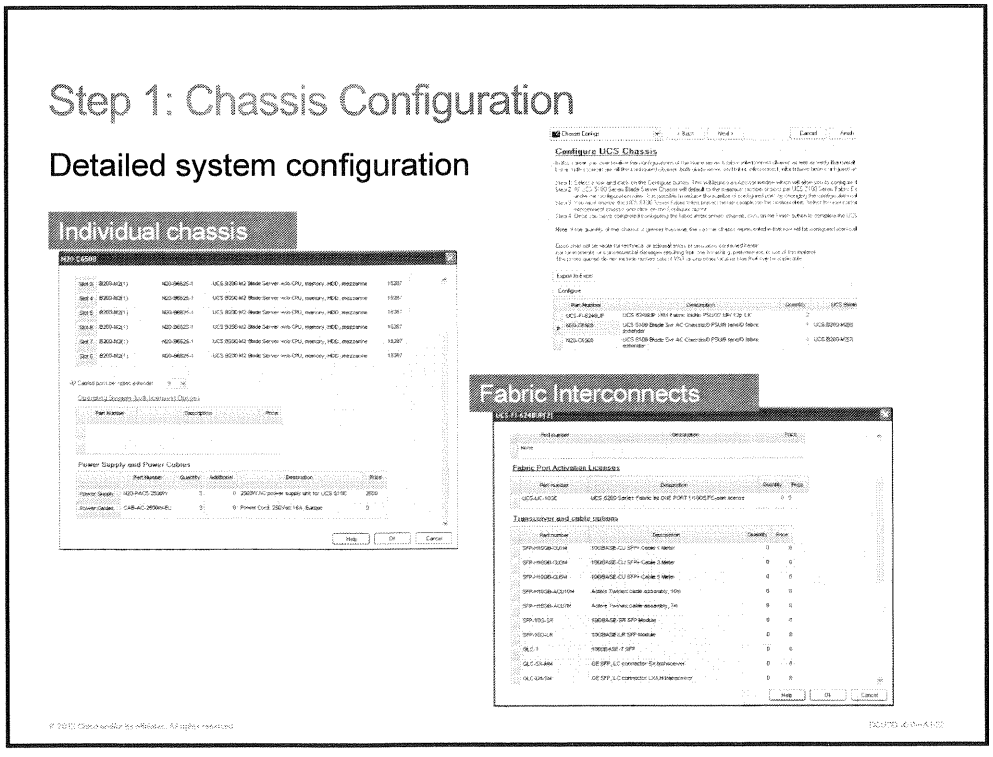
With blade templates defined, you must select the desired blade templates and enter the required quantities. This enables you to set the characteristics of the system.

The following parameters are set on the System Preferences tab:

- Chassis general preferences:
  - The tool suggests the minimum chassis quantity. The number is calculated by populating the maximum number of blades into each chassis (that is, eight per chassis) up to the configured quantity of blades.
  - Additional UCS 5108 chassis—you can add additional chassis if necessary or required by the design
  - Country of deployment
  - Level of power redundancy in system, where you set the selected power scheme per design (that is, nonredundant, N+1 redundancy, grid redundancy)
- Fabric interconnect options:
  - Fabric interconnect model – UCS 6120XP, UCS 6140XP, or the newest 6248UP and 6296UP fabric interconnect
  - External LAN ports per switch, which defines the number of Ethernet uplinks per fabric interconnect
  - External SAN ports per switch, which defines the number of SAN uplinks per fabric interconnect
  - Redundant fabric interconnect options, which defines whether the BOM will include one or two fabric interconnects
  - Power cord, which defines the type of power cable per outlet used in the country where fabric interconnects will be deployed

https://t.me/learningnets

- Blade server chassis options:
  - Chassis, which defines the type of chassis—currently, you can select between AC or DC powered chassis
  - Power cord, which defines the type of power cable per outlet used in the country where the chassis will be used
  - I/O module, which defines the type of I/O module used in the chassis, where you can select between 4-port first generation 2104XP or 4/8-port second generation 2204/2208XP IOM.



The following is a more detailed configuration of system chassis and fabric interconnects.

For the chassis, you can select the following parameters on a chassis-by-chassis basis:

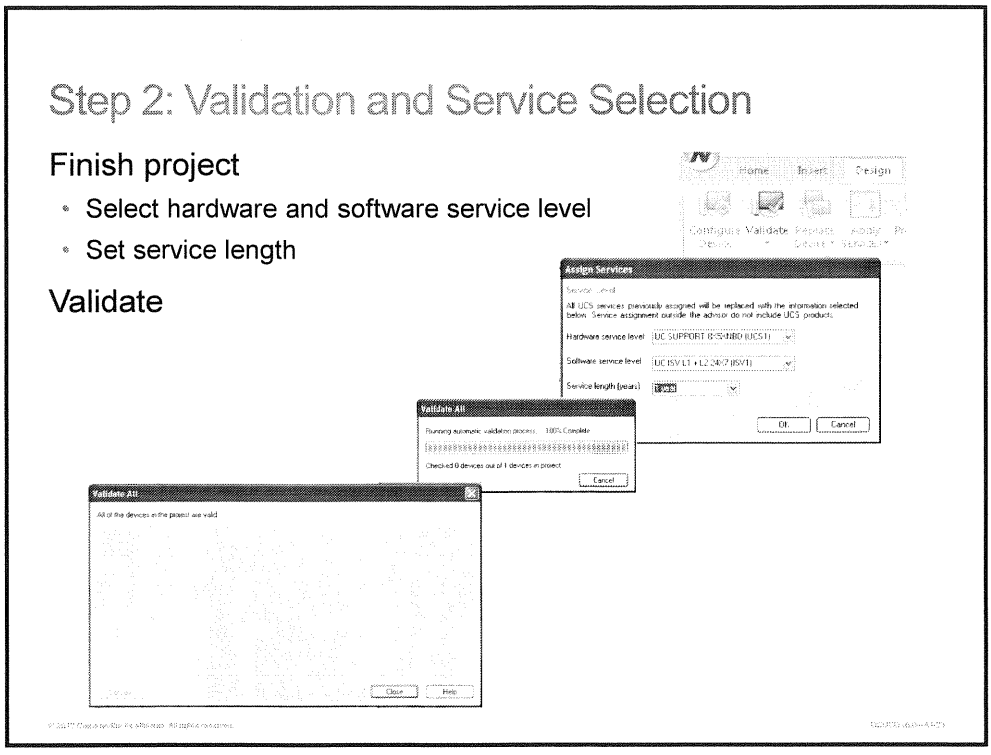
- Number of cabled ports per I/O module (1,2,4, or 8 in the case of 2208XP IOM)
- Operating system licenses (optional)
- Power supplies and cables

For the fabric interconnects, you can define the following parameters:

- Expansion slot – depending on the interconnect type
- Fabric port activation licenses
- Transceivers and cables for Ethernet and Fibre Channel interfaces
- Power supplies and cables

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

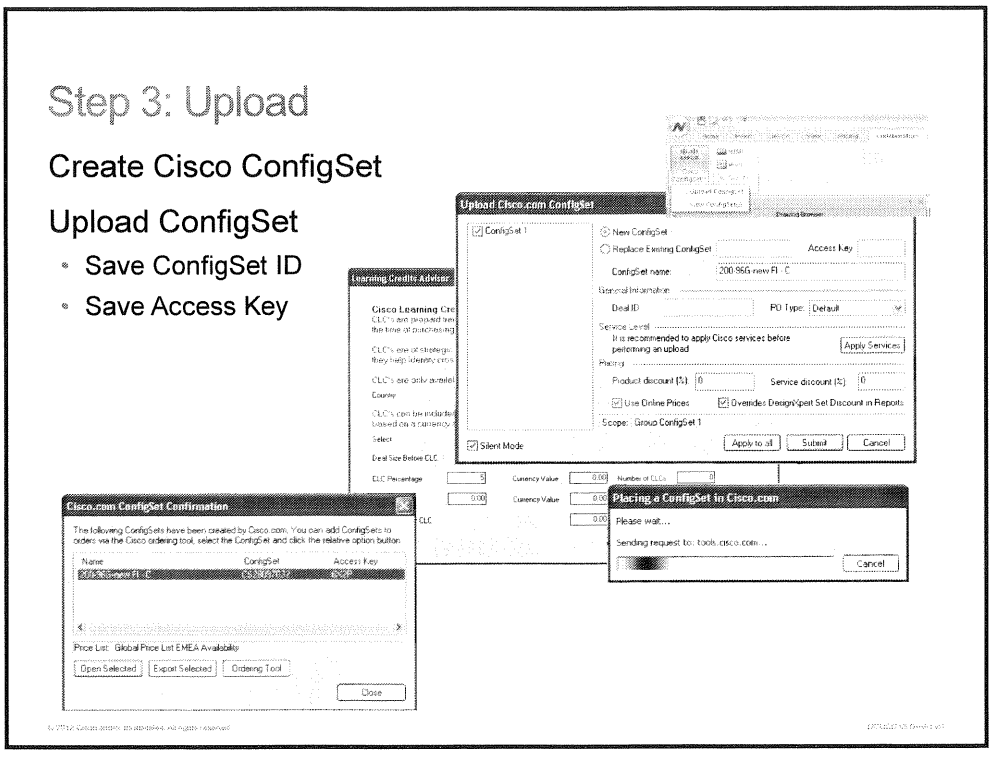


In the final step, you distribute blades into the chassis, apply the chassis and fabric interconnect configuration, and ensure that the BOM is complete.

The service selection process is used to define the type of services that are to be applied to the selected Cisco UCS components. The services can be selected for the whole BOM or part of the BOM. The selection includes definition of the hardware and software service level, length of the service, and so on.

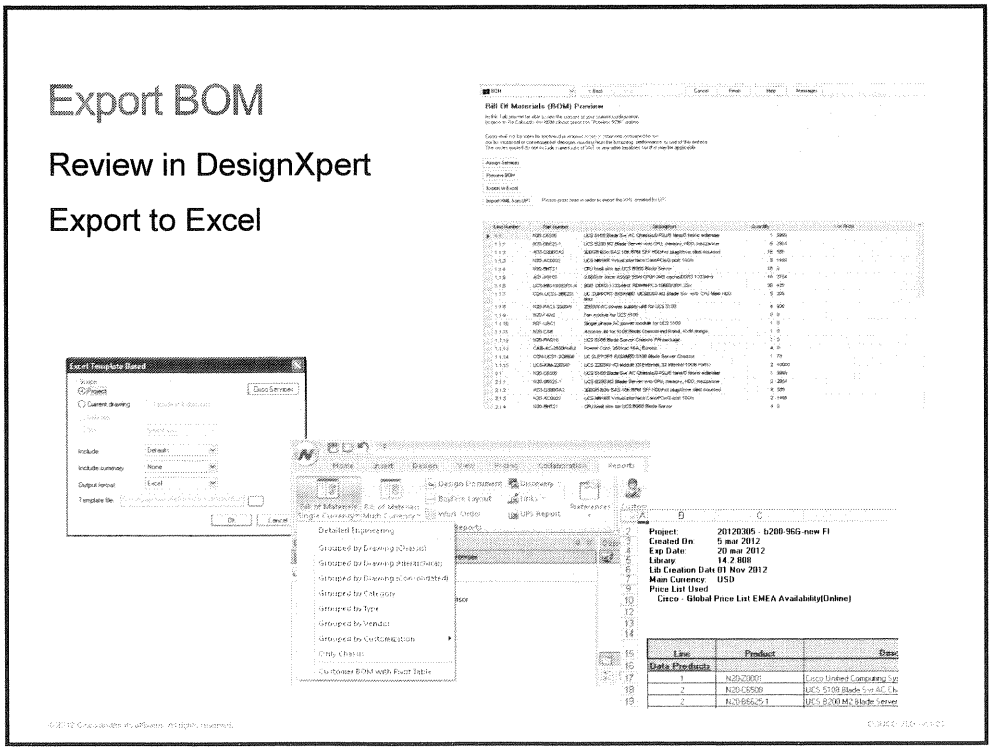
Before proceeding to the final configuration step—uploading—the BOM should be reviewed to verify that it is correct.

https://t.me/learningnets



Once the hardware and services have been selected, the BOM ConfigSet can be created and BOM uploaded to Cisco.com.

The ConfigSet ID and Access key are generated and they should be saved for later reference.



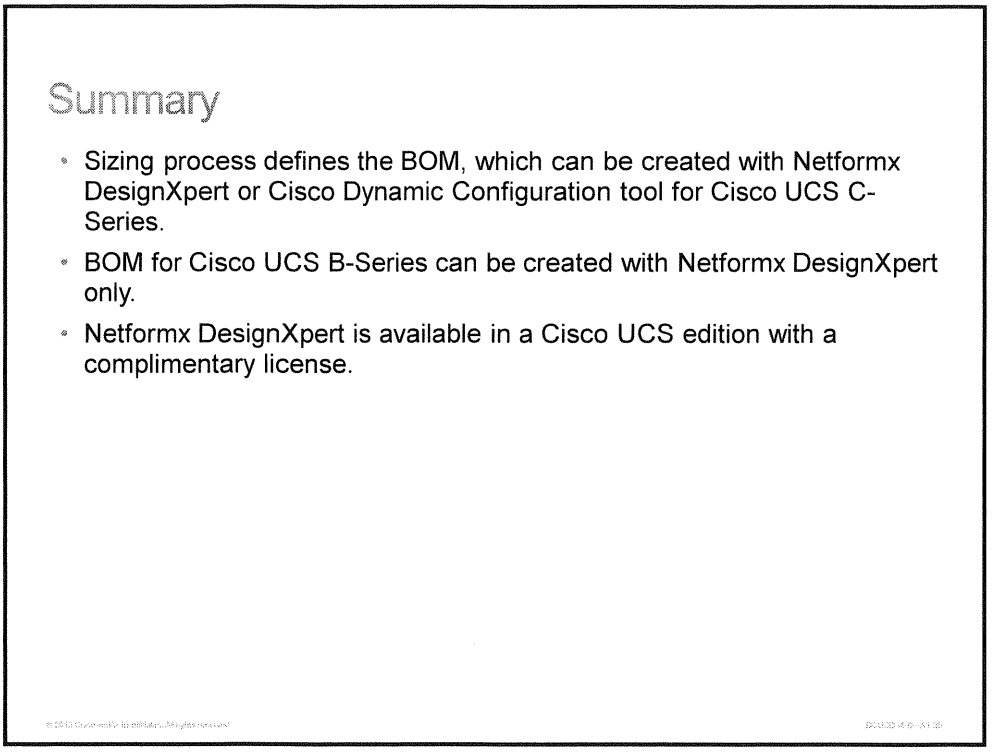
The BOM can be reviewed in the DesignXpert or exported in different formats (for example, Microsoft Excel) if required.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

Summary

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



References

For additional information, refer to these resources:

- http://design.netformx.com/dxucs
- https://communities.cisco.com/community/partner/datacenter

https://t.me/learningnets

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