certcollecion.net

DCUFD

Designing Cisco Data Center Unfied Fabric

Volume 1

Version 5.0



Student Guide

Text Part Number: 97-3184-01

https://t.me/learningnets

certcollecion.net



Americas Headquarters
Cisco Systems, Inc.
San Jose, CA

Asia Pacific Headquarters
Cisco Systems (USA) Pte. Ltd.
Singapore

Europe Headquarters
Cisco Systems International BV Amsterdam,
The Netherlands

Cisco has more than 200 offices worldwide. Addresses, phone numbers, and fax numbers are listed on the Cisco Website at **www.cisco.com/go/offices.**

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

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

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

https://t.me/learningnets

certcollecion.net



Students, this letter describes important course evaluation access information!

Welcome to Cisco Systems Learning. Through the Cisco Learning Partner Program, Cisco Systems is committed to bringing you the highest-quality training in the industry. Cisco learning products are designed to advance your professional goals and give you the expertise you need to build and maintain strategic networks.

Cisco relies on customer feedback to guide business decisions; therefore, your valuable input will help shape future Cisco course curricula, products, and training offerings. We would appreciate a few minutes of your time to complete a brief Cisco online course evaluation of your instructor and the course materials in this student kit. On the final day of class, your instructor will provide you with a URL directing you to a short post-course evaluation. If there is no Internet access in the classroom, please complete the evaluation within the next 48 hours or as soon as you can access the web.

On behalf of Cisco, thank you for choosing Cisco Learning Partners for your Internet technology training.

Sincerely,

Cisco Systems Learning

https://t.me/learningnets

certcollecion.net

https://t.me/learningnets

certcollecion.net

Table of Contents

Volume 1

Course Introduction ***1***

Overview 1
Learner Skills and Knowledge 2
Course Goal and Objectives 3
Course Flow 4
Additional References 5
Cisco Glossary of Terms 6
Your Training Curriculum 7
Additional Resources 10
Introductions 12

Cisco Data Center Solutions ***1-1***

Overview 1-1
Module Objectives 1-1

Defining the Data Center **1-3**

Overview 1-3
Objectives 1-3
Data Center Solution Components 1-4
Data Center Terminology 1-14
Data Center Challenges 1-18
Introduction to Cloud Computing 1-37
Data Center Virtualization 1-51
Summary 1-55

Identifying the Cisco Data Center Solution **1-57**

Overview 1-57
Objectives 1-57
Cisco Data Center Architecture Overview 1-58
Cisco Data Center Architecture Network 1-65
Cisco Data Center Architecture Storage 1-88
Summary 1-92

Designing the Cisco Data Center Solution **1-93**

Overview 1-93
Objectives 1-93
Design Process 1-94
Design Deliverables 1-108
Cisco Validated Designs 1-112
Summary 1-113
Module Summary 1-115
Module Self-Check 1-117
Module Self-Check Answer Key 1-119

Data Center Technologies ***2-1***

Overview 2-1
Module Objectives 2-1

Designing Layer 2 and Layer 3 Switching **2-3**

Overview 2-3
Objectives 2-3
Forwarding Architectures 2-4
IP Addressing and Routing 2-10
Summary 2-15

https://t.me/learningnets

certcollecion.net

Virtualizing Data Center Components 2-17
Overview 2-17
Objectives 2-17
Device Virtualization Mechanisms 2-18
Virtual Device Contexts 2-23
Virtualization with Contexts 2-32
Virtualization with Virtual Appliances 2-36
Summary 2-38
Designing Layer 2 Multipathing Technologies 2-39
Overview 2-39
Objectives 2-39
Network Scaling Technologies 2-40
vPC and MEC 2-43
Cisco FabricPath 2-58
Summary 2-79
References 2-79
Module Summary 2-81
Module Self-Check 2-83
Module Self-Check Answer Key 2-85
Data Center Topologies 3-1
Overview 3-1
Module Objectives 3-1
Designing the Data Center Core Layer Network 3-3
Overview 3-3
Objectives 3-3
Data Center Core Layer 3-4
Layer 3 Data Center Core Design 3-6
Layer 2 Data Center Core Design 3-8
Data Center Collapsed Core Design 3-13
Summary 3-15
Designing the Data Center Aggregation Layer 3-17
Overview 3-17
Objectives 3-17
Classic Aggregation Layer Design 3-18
Aggregation Layer with VDCs 3-21
Aggregation Layer with Unified Fabric 3-29
Aggregation Layer with IP-Based Storage 3-35
Summary 3-38
Designing the Data Center Access Layer 3-39
Overview 3-39
Objectives 3-39
Classic Access Layer Design 3-40
Access Layer with vPC and MEC 3-43
Access Layer with FEXs 3-44
Access Layer with Unified Fabric 3-52
Summary 3-56
Designing the Data Center Virtualized Access Layer 3-57
Overview 3-57
Objectives 3-57
Virtual Access Layer 3-58
Virtual Access Layer Solutions 3-60
Using Cisco Adapter FEX 3-62
Using Cisco VM-FEX 3-63
Solutions with the Cisco Nexus 1000V Switch 3-65
Summary 3-77

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

Designing High Availability 3-79
Overview 3-79
Objectives 3-79
High Availability for IP 3-80
High Availability Using vPC and VSS 3-85
High Availability Using IP Routing and FHRP 3-87
IP Routing Protocols Deployment Design 3-88
High Availability Using RHI 3-91
High Availability Using LISP 3-95
Summary 3-103
Designing Data Center Interconnects 3-105
Overview 3-105
Objectives 3-105
Reasons for Data Center Interconnects 3-106
Data Center Interconnect Technologies 3-111
Cisco OTV 3-113
Storage Replication Technologies and Interconnects 3-126
Summary 3-131
References 3-131
Module Summary 3-133
Module Self-Check 3-135
Module Self-Check Answer Key 3-138

https://t.me/learningnets

certcollecion.net

https://t.me/learningnets

certcollecion.net

DCUFD

Course Introduction

Overview

The *Designing Cisco Data Center Unified Fabric* (DCUFD) v5.0 is a five-day instructor-led course aimed at providing data center designers with the knowledge and skills needed to design scalable, reliable, and intelligent data center unified fabrics, and virtualization solutions based on fabric extenders (FEXs), Fibre Channel over Ethernet (FCoE), Cisco FabricPath, and equipment and link virtualization technologies.

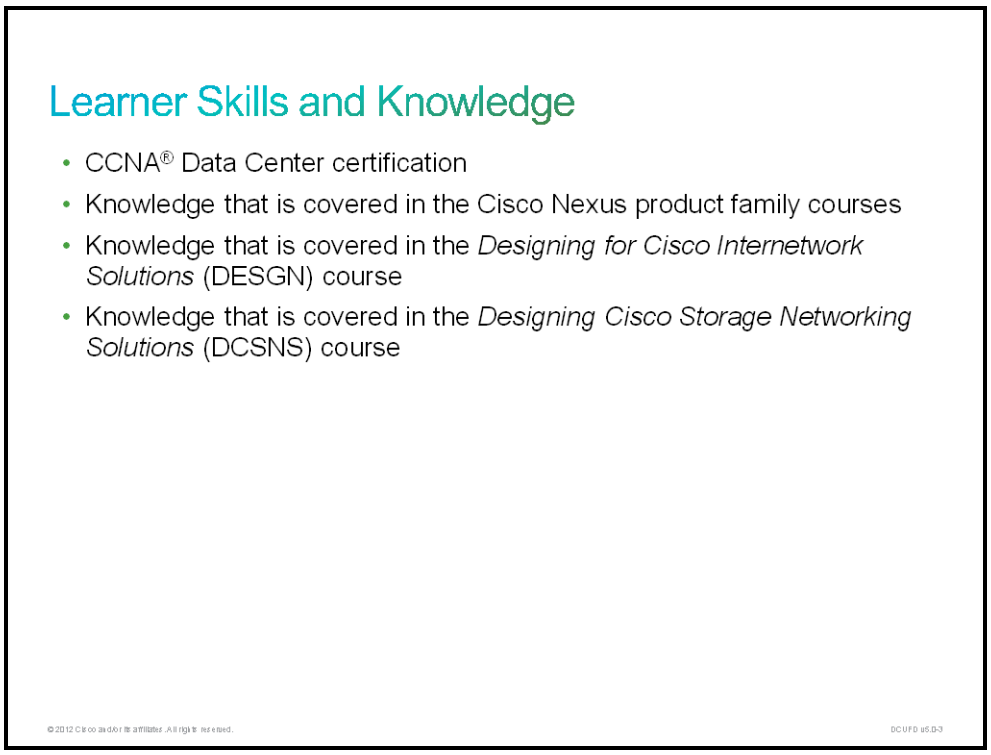
The course describes the Cisco data center unified fabric solutions, and explains how to evaluate existing data center infrastructure, determine the requirements, and design the Cisco data center unified fabric solution based on Cisco products and technologies.

https://t.me/learningnets

certcollecion.net

Learner Skills and Knowledge

This subtopic lists the skills and knowledge that learners must possess to benefit fully from the course. The subtopic also includes recommended Cisco learning offerings that learners should first complete to benefit fully from this course.

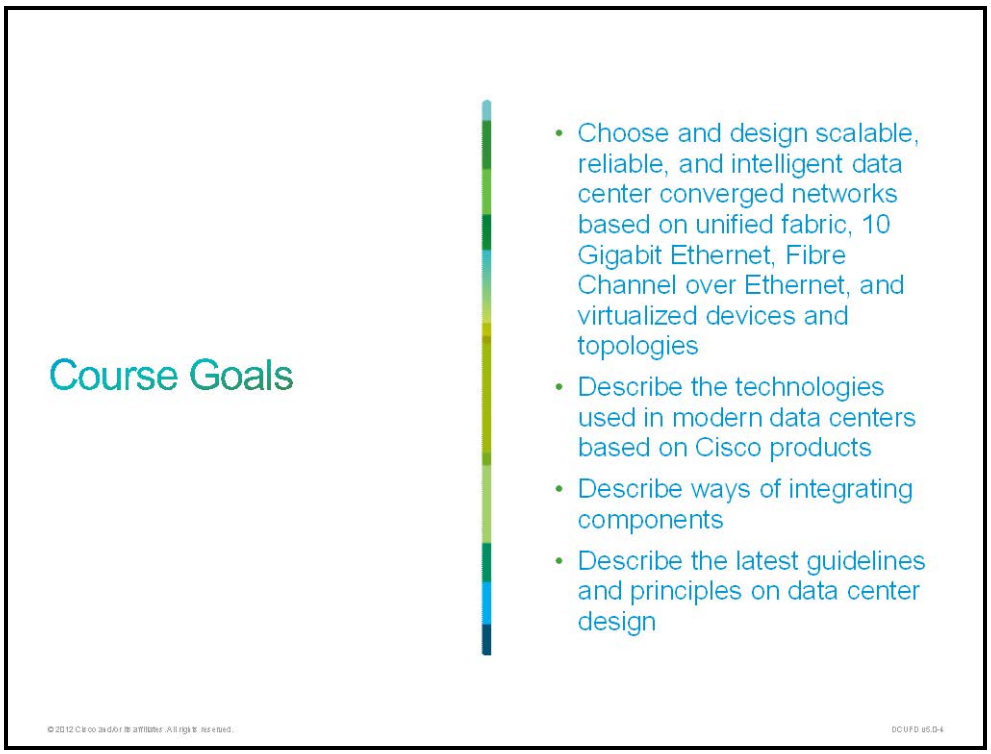


https://t.me/learningnets

certcollecion.net

Course Goal and Objectives

This topic describes the course goal and objectives.



Upon completing this course, you will be able to meet these objectives:

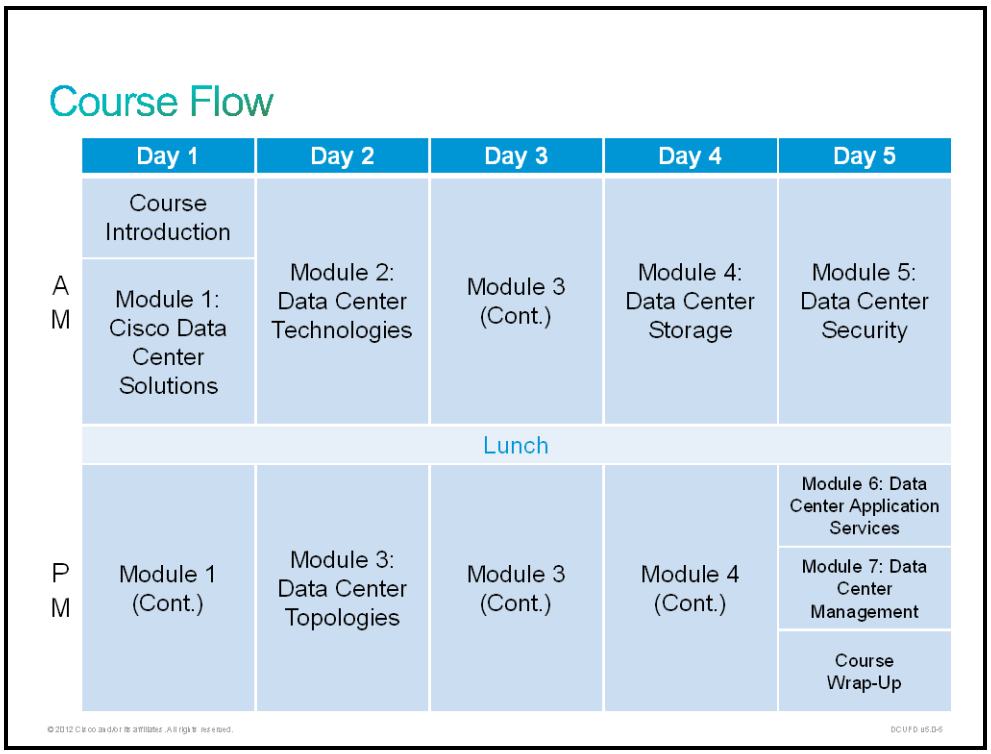
- Evaluate the data center solution design and design process regarding the contemporary data center challenges, Cisco Data Center Architecture solution, and components
- Provide a comprehensive and detailed overview of technologies used in data centers, and describe scalability implications and their possible use in cloud environments
- Design data center connections and topologies in the core layer
- Explain and design data center storage designs, solutions, and limitations of various storage technologies
- Design secure data centers that are protected from application-based threats, network security threats, and physical security threats
- Design data center infrastructure that is required to implement network-based application services
- Design data center management to facilitate monitoring, managing, and provisioning data center equipment and applications

https://t.me/learningnets

certcollecion.net

Course Flow

This topic presents the suggested flow of the course materials.



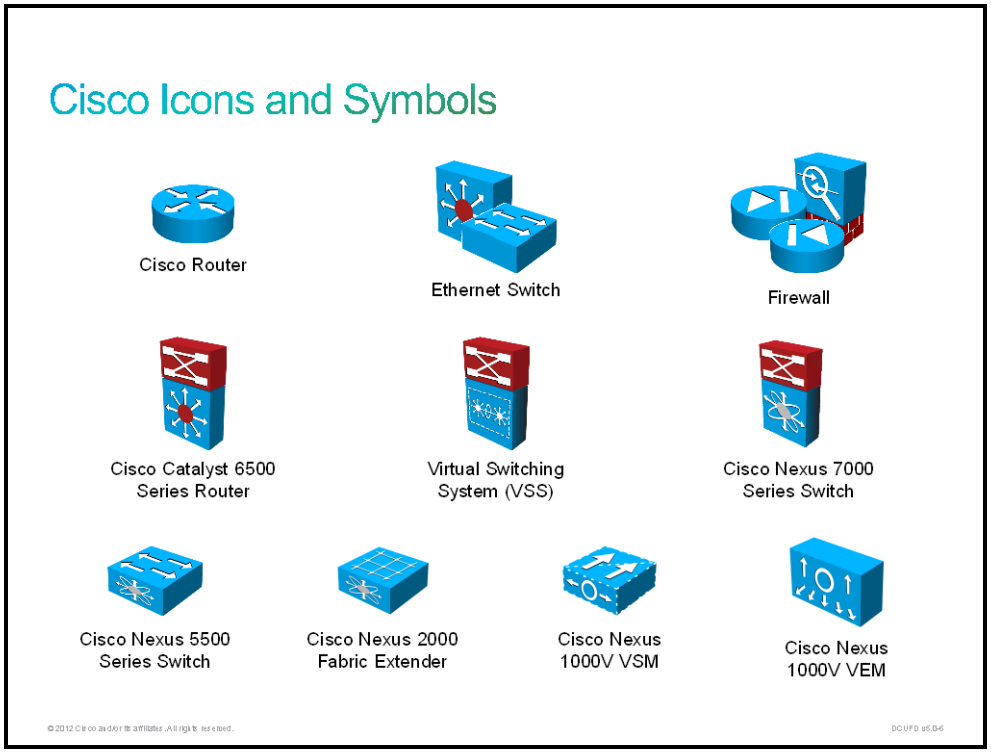
The schedule reflects the recommended structure for this course. This structure allows enough time for the instructor to present the course information and for you to work through the lab activities. The exact timing of the subject materials and labs depends on the pace of your specific class.

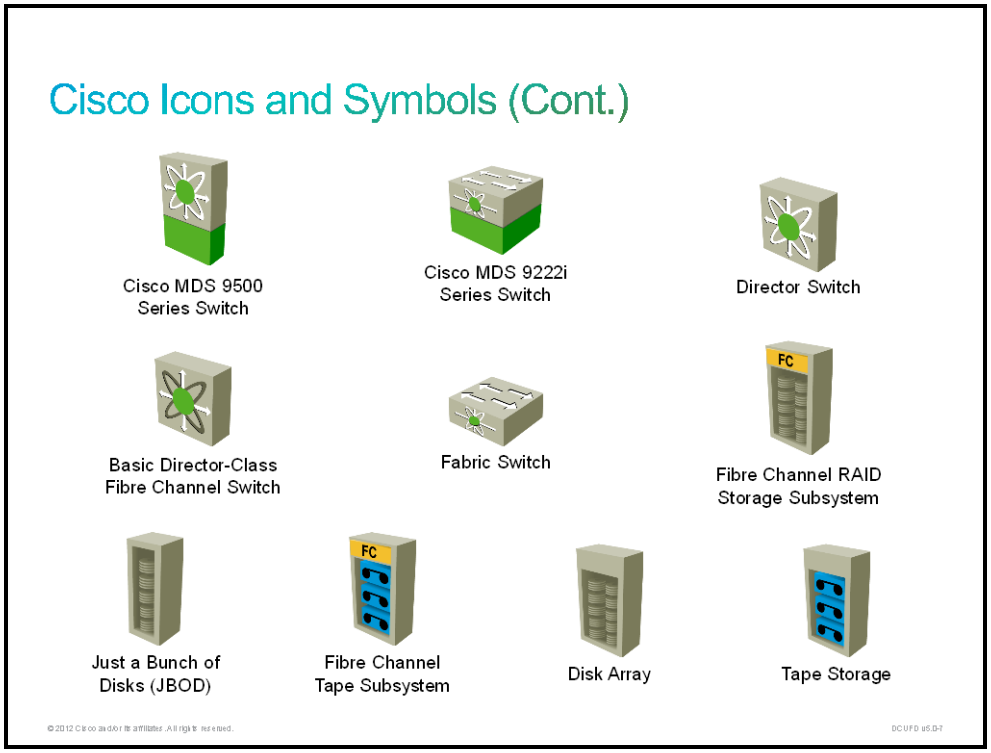
https://t.me/learningnets

certcollecion.net

Additional References

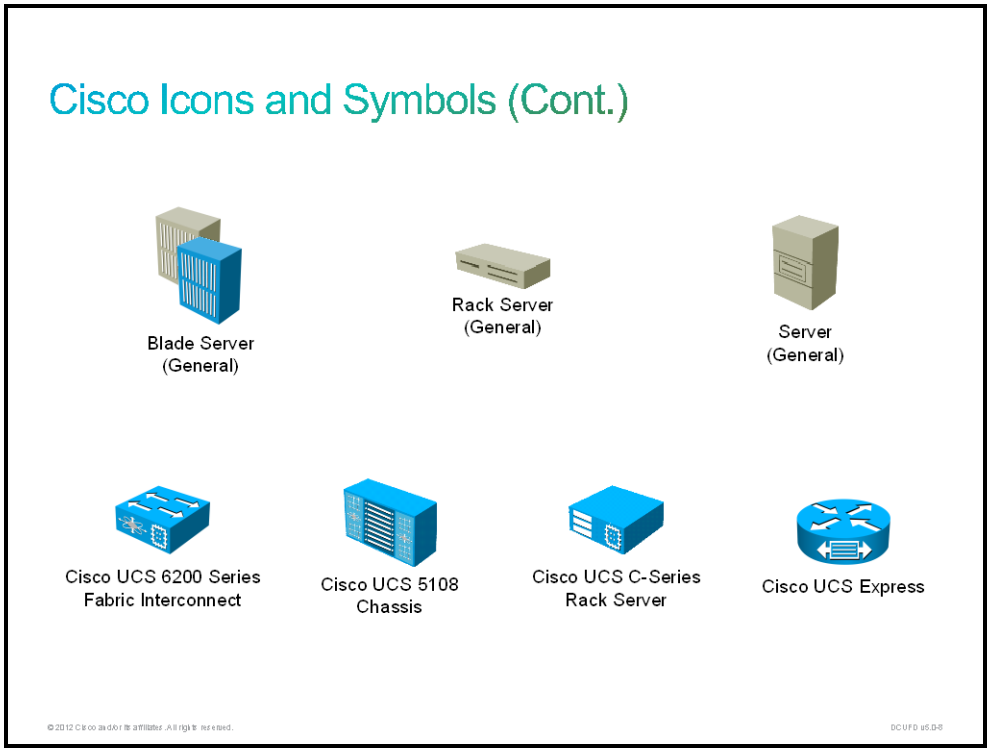
This topic presents the Cisco icons and symbols that are used in this course, as well as information on where to find additional technical references.





https://t.me/learningnets

certcollecion.net



Cisco Glossary of Terms

For additional information on Cisco terminology, refer to the *Cisco Internetworking Terms and Acronyms* glossary of terms at http://docwiki.cisco.com/wiki/Category:Internetworking_Terms_and_Acronyms_(ITA).

https://t.me/learningnets

certcollecion.net

Your Training Curriculum

This topic presents the training curriculum for this course.

Cisco Learning Network

Developing a world of talent through collaboration, social learning, online assessment, and mentoring

https://learningnetwork.cisco.com



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

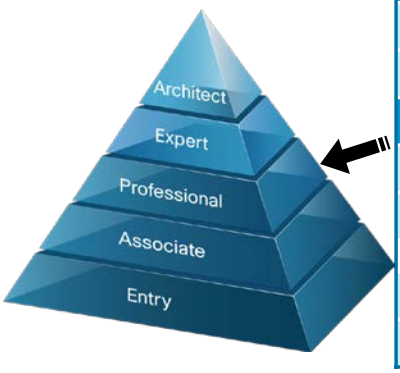
To prepare and learn more about IT certifications and technology tracks, visit the Cisco Learning Network, which is the home of Cisco Certifications.

https://t.me/learningnets

certcollecion.net

Cisco Career Certifications: Cisco CCNP Data Center

Expand Your Professional Options and Advance Your Career



Cisco CCNP Data Center
Implementing Cisco Data Center Unified Fabric (DCUFI)
Implementing Cisco Data Center Unified Computing (DCUCI)
Available Exams (pick a group of 2)
Designing Cisco Data Center Unified Computing (DCUCD)
Designing Cisco Data Center Unified Fabric (DCUFD)
or
Troubleshooting Cisco Data Center Unified Fabric (DCUFT)
Troubleshooting Cisco Data Center Unified Computing (DCUCT)

www.cisco.com/go/certifications

© 2012 Cisco and/or its affiliates. All rights reserved.
DCUFD v5.0-11

You are encouraged to join the Cisco Certification Community, a discussion forum open to anyone holding a valid Cisco Career Certification:

- Cisco CCDE®
- Cisco CCIE®
- Cisco CCDP®
- Cisco CCNP®
- Cisco CCNP® Data Center
- Cisco CCNP® Security
- Cisco CCNP® Service Provider
- Cisco CCNP® Service Provider Operations
- Cisco CCNP® Voice
- Cisco CCNP® Wireless
- Cisco CCDA®
- Cisco CCNA®
- Cisco CCNA® Data Center
- Cisco CCNA® Security
- Cisco CCNA® Service Provider
- Cisco CCNA® Service Provider Operations
- Cisco CCNA® Voice
- Cisco CCNA® Wireless

https://t.me/learningnets

certcollecion.net

It provides a gathering place for Cisco certified professionals to share questions, suggestions, and information about Cisco Career Certification programs and other certification-related topics. For more information, visit http://www.cisco.com/go/certifications.

https://t.me/learningnets

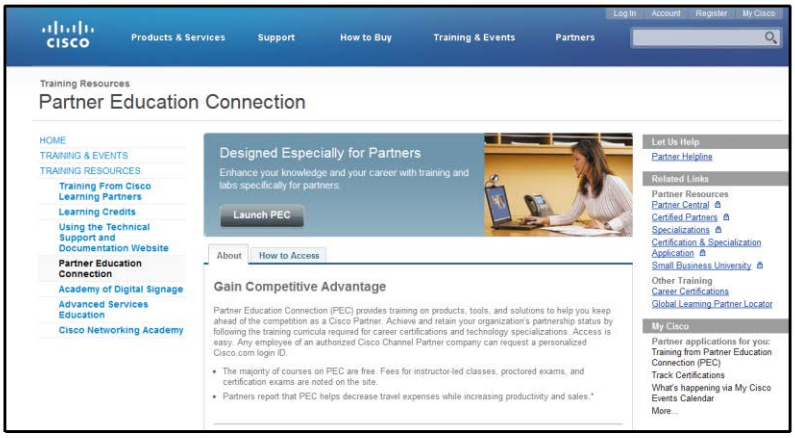
certcollecion.net

Additional Resources

For additional information about Cisco technologies, solutions, and products, refer to the information available at the following pages.

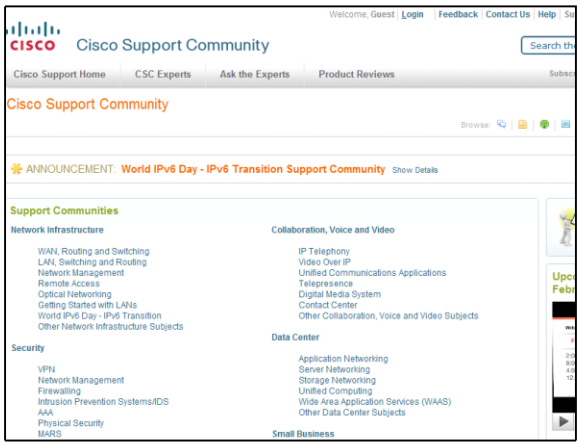
Cisco Partner Education Center

http://www.cisco.com/go/pec



Cisco Support Community

https://supportforums.cisco.com/index.jspa



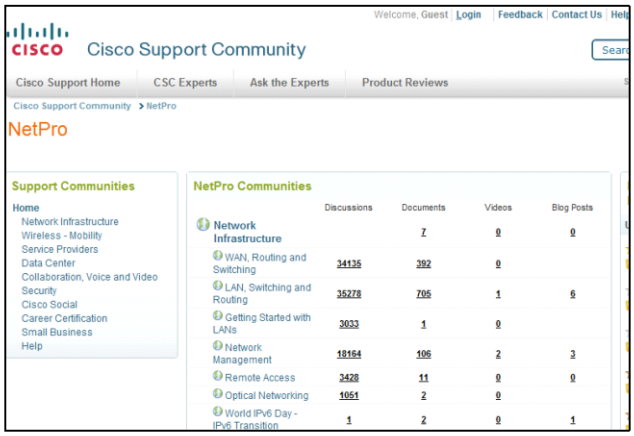
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

Cisco NetPro Forums

https://supportforums.cisco.com/community/netpro



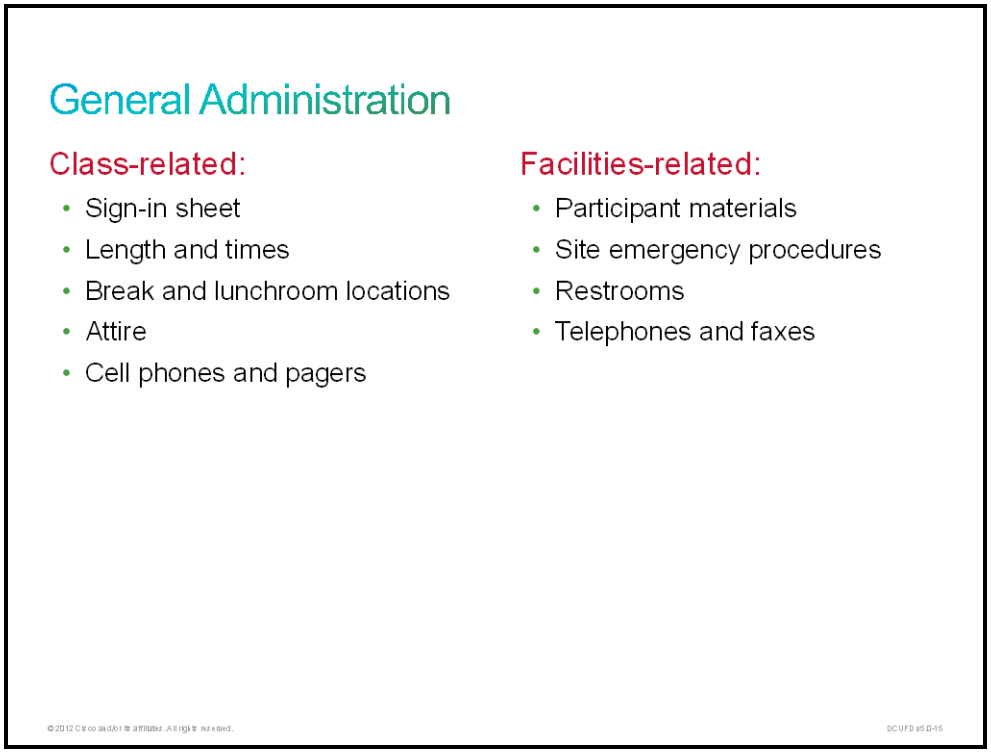
© 2012 Cisco and/or its affiliates. All rights reserved.

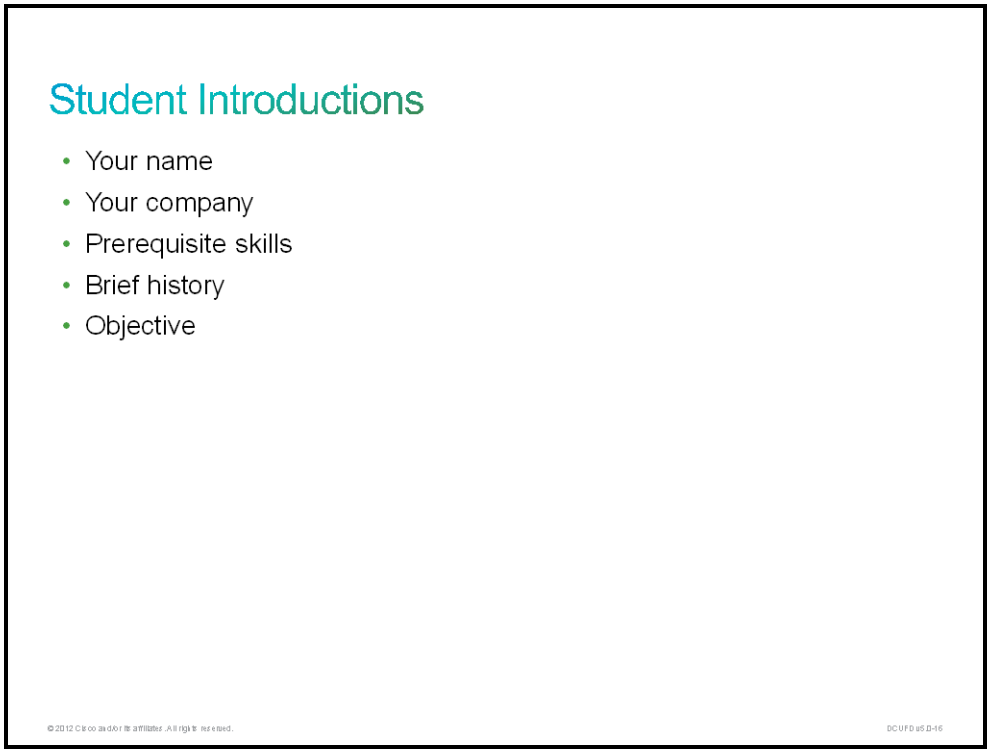
https://t.me/learningnets

certcollecion.net

Introductions

Please use this time to introduce yourself to your classmates so that you can better understand the colleagues with whom you will share your experience.





https://t.me/learningnets

certcollecion.net

Module 1

Cisco Data Center Solutions

Overview

Modern data centers operate in high availability and are the foundation for business processes. Additionally, the cloud computing model has been emerging and data centers provide the infrastructure that is needed to support various cloud computing deployments.

Cisco offers a comprehensive set of technologies and devices that are used to implement data centers. These include switches, servers, security appliances, virtual appliances, and so on.

In this module, you will learn how to define data centers, identify technologies, and design processes to successfully design a data center.

The data center design process needs to be well run and well documented. This module provides an overview of the design process and documentation.

Module Objectives

Upon completing this module, you will be able to evaluate data center solution designs and the design process regarding contemporary data center challenges, the Cisco Data Center Architecture solution, and components. This ability includes being able to meet these objectives:

- Analyze the relationship between the business, technical, and environmental challenges and goals for contemporary data center solutions
- Provide a high-level overview of the Cisco data center solution architectural framework and components within the solution
- Define the tasks and phases of the design process for the Cisco Unified Computing solution

https://t.me/learningnets

certcollecion.net

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

Lesson 1

Defining the Data Center

Overview

A modern data center is an essential component in any business, providing highly available services to its users and customers. This lesson outlines various categories of data centers, defines commonly used terms, and analyzes challenges and concepts with a focus on virtualization.

Objectives

Upon completing this lesson, you will be able to analyze the relationship between the business, technical, and environmental challenges and goals for contemporary data center solutions. This ability includes being able to meet these objectives:

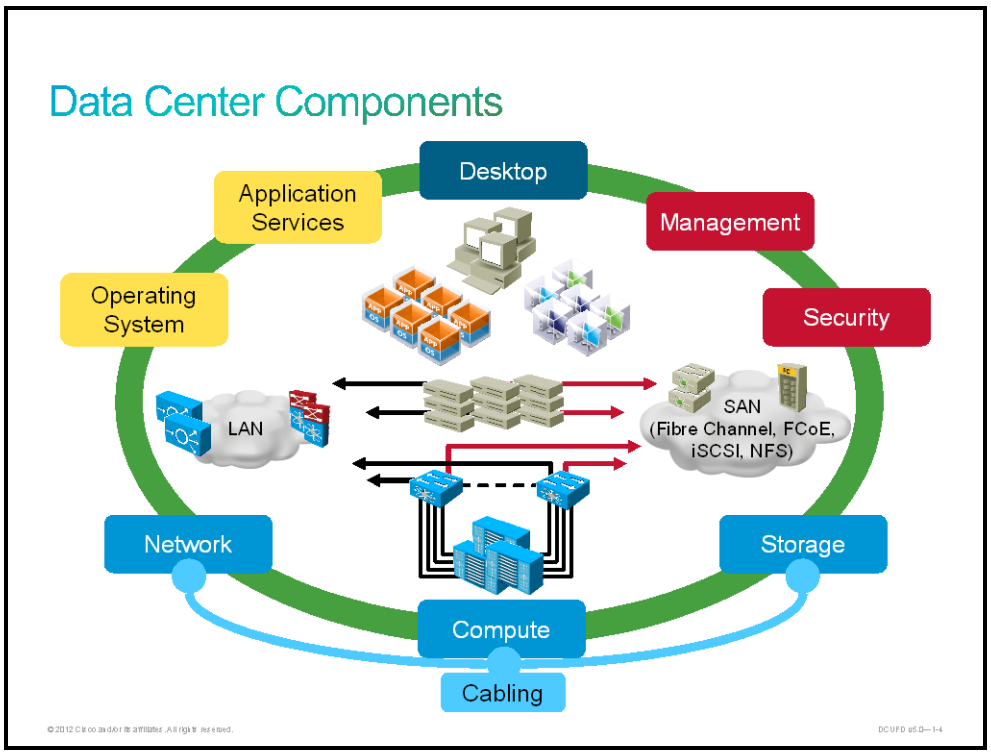
- Categorize general data center solution components
- Define the baseline technology and terminology used in data center solutions
- Analyze business, technical, and environmental challenges
- Recognize the cloud computing paradigm, terms, and concepts
- Recognize the importance of virtualization technologies and solutions for data center evolution

https://t.me/learningnets

certcollecion.net

Data Center Solution Components

This topic describes how to categorize general data center solution components.



Data Center Definition

A data center is a centralized or geographically distributed group of departments that house the computing systems and their related storage equipment or data libraries. A data center has controlled centralized management that enables an enterprise to operate according to business needs.

Data Center Solutions

A data center infrastructure is an essential component that supports Internet services, digital commerce, electronic communications, and other business services and solutions:

- Network technologies and equipment like intelligent switches, multilayer and converged devices, high-availability mechanisms, and Layer 2 and Layer 3 protocols
- Storage solutions and equipment that cover technologies ranging from Fibre Channel, Internet Small Computer Systems Interface (iSCSI), Network File System (NFS), Fibre Channel over Ethernet (FCoE), storage network equipment, and storage devices like disk arrays and tape libraries
- Computing technologies and equipment, including general purpose and specialized servers
- Operating system and server virtualization technologies
- Application services technologies and products like load balancers and session enhancement devices
- Management systems that are used to manage network, storage, and computing resources, operating systems, server virtualization, applications, and security aspects of the solution

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

- Security technologies and equipment that are employed to ensure confidentiality and security of sensitive data and systems
- Desktop virtualization solutions and access clients

Data Center Goals

A data center goal is to sustain the business functions and operations, along with flexibility for future data center changes. A data center network must be flexible and must support nondisruptive scalability of applications and computing resources to adapt the infrastructure for future business needs.

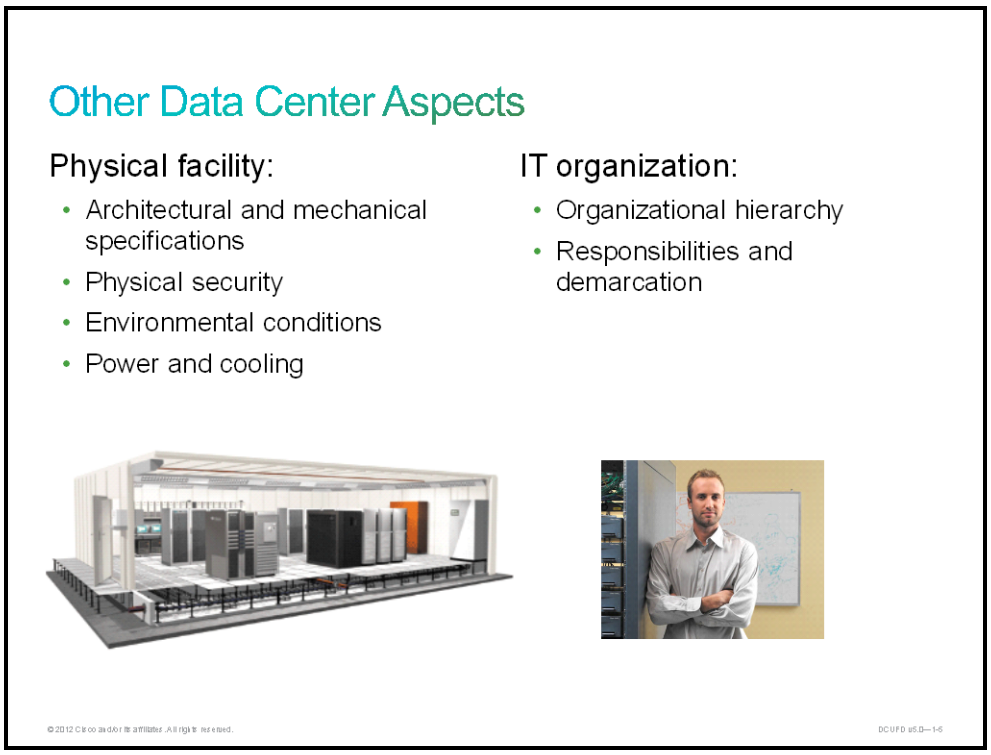
Business Continuance Definition

Business continuity is the ability to adapt and respond to risks as well as opportunities in order to maintain continuous business operations. There are four primary aspects of business continuity:

- High availability (disaster tolerance)
- Continuous operations
- Disaster recovery
- Disaster tolerance

https://t.me/learningnets

certcollecion.net



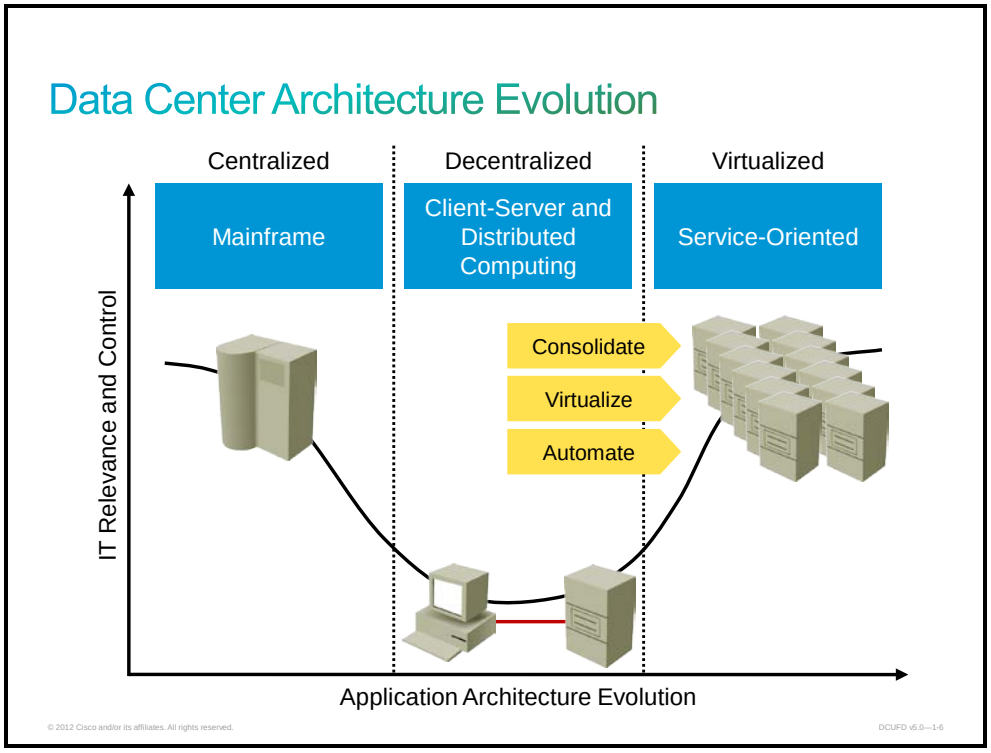
Apart from the already-mentioned aspects and components of the data center solution, there are two very important components that influence how the solution is used and scaled, and define the lifetime of the solution:

- **Physical facility:** The physical facility encompasses all the data center facility physical characteristics that affect the already-mentioned infrastructure. This includes available power, cooling capacity, physical space and racks, physical security and fire prevention systems, and so on.
- **IT organization:** The IT organization defines the IT departments and how they interact in order to offer IT services to business users. This organization can be in the form of a single department that takes care of all the IT aspects (typically with the help of external IT partners). Alternatively, in large companies, it can involve multiple departments, with each department taking care of a subset of the data center infrastructure.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net



Data centers have changed and evolved over time.

At first, data centers were monolithic and centralized, employing mainframes and terminals, which the users accessed to perform their work on the mainframe. Mainframes are still used in the finance sector, because they are an advantageous solution in terms of availability, resilience, and service level agreements (SLAs).

The second era of data center computing was characterized by pure client-server and distributed computing. Applications were designed in such a way that client software was used to access an application, and the services were distributed due to poor computing ability and high link costs. The mainframes were too expensive.

Today, the communication infrastructure is relatively cheaper and the computing capacities have increased. Consequently, data centers are being consolidated because the distributed approach is expensive in the long term. The new solution is equipment virtualization, which makes the utilization of servers much more common than in the distributed approach. This solution also provides significant gains in terms of return on investment (ROI) and the total cost of ownership (TCO).

Virtualization Is Changing the Data Center Architecture

The Cisco Data Center Business Advantage framework brings sequential and stepwise clarity to the data center. Cisco customers are at different stages of this journey.

Data center networking capabilities bring an open and extensible data center networking approach to the placement of your IT solution. This approach will support your business processes, whether the processes are conducted on a factory floor in a pod or in a Tier 3 data center 200 feet (61 m) below the ground in order to support precious metal mining. In short, Cisco data center networking delivers location freedom.

https://t.me/learningnets

certcollecion.net

Phase 1: Data Center Networking and Consolidation

Earlier, Cisco entered all the data center transport and network markets—server networking, storage networking, application networking, and security networking. By delivering these types of services as an integrated system, Cisco has improved reliability and validated best-practice architectures and has also improved customer success.

Phase 2: Unified Fabric

The introduction of a unified fabric is the first step toward removing the network barriers to deploying any workload, on any server, in seconds.

The new atomic unit, or building block, of the data center is not the physical server. It is the virtual machine (VM) that provides a consistent abstraction between physical hardware and logically defined software constructs.

Phase 3: Unified Computing and Automation

Bringing the network platform that is created by the unified fabric together with the virtualization platform and the computing and storage platform introduces a new paradigm in the market: Cisco Unified Computing. This solution is a simpler, more efficient architecture that extends the life cycle of capital assets, but it also enables the enterprise to execute business processes in the best places and in the most efficient ways, all with high availability. Cisco Unified Computing focuses on automation simplification for a predesigned virtualization platform. It is another choice that simplifies startup processes and ongoing operations within the virtualized environment, and it can deliver provisioning freedom. This subject will be discussed in more detail later on.

Phase 4: Enterprise-Class Clouds and Utility

With the integration of cloud technologies, principles, and architectures and the Cisco Unified Computing architecture, workloads can become increasingly portable. Bringing security, control, and interoperability to standalone cloud architectures can enable enterprise-class clouds. The freedom of choice about where business processes are executed is extended across the boundaries of an organization, to include the providers as well (with no compromise on service levels).

After the process is automated, enterprise internal IT resources will be seen as a utility that is able to automate and dynamically provision the infrastructure across the network, computing, and virtualization platforms, thus simplifying the line of business and services in the enterprise.

It is an evolution that is enabled by integrating cloud computing principles, architectures, and technologies.

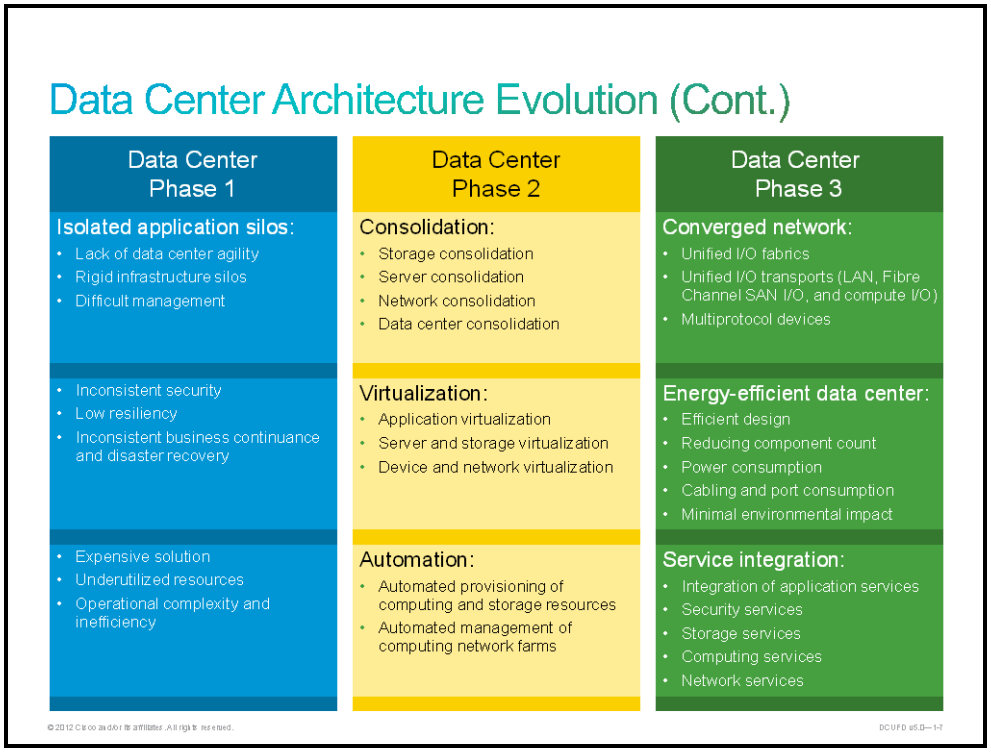
Phase 5: Intercloud Market

A goal of Cisco is to create a market as a new wave of innovation and investment similar to what the industry last saw with the Internet growth of the mid-1990s. However, this time the growth should be predicated not on addressing a federation across providers, but on portable workloads. This market extends from the enterprise to the provider and from the provider to another provider based on available capacity, power cost, proximity, and so on.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net



Data center trends that have affected the data center architecture and design can be summarized in significant phases and stages:

Phase 1

Isolated Application Silos

Data centers are about servers and applications. The first data centers were mostly mainframe, glass-house, raised-floor structures that housed the computer resources, as well as the intellectual capital (programmers and support staff) of the enterprise. Over the past decade, most data centers have evolved on an ad hoc basis. The goal was to provide the most appropriate server, storage, and networking infrastructure that supported specific applications. This strategy led to data centers with stovepipe architectures or technology islands that were difficult to manage or adapt to changing environments.

There are many server platforms in current data centers, all designed to deploy a series of applications:

- IBM mainframe applications
- Email applications on Microsoft Windows servers
- Business applications on IBM AS/400 servers
- Enterprise resource planning (ERP) applications on UNIX servers
- R&D applications on Linux servers

In addition, a broad collection of storage silos exists to support these disparate server environments. These storage silos can be in the form of integrated, direct-attached storage (DAS), network-attached storage (NAS), or small SAN islands.

This silo approach has led to underutilization of resources, difficulty in managing these disparate complex environments, and difficulty in applying uniform services such as security and application optimization. It is also difficult to implement strong, consistent disaster-recovery procedures and business continuance functions.

https://t.me/learningnets

certcollecion.net

Phase 2

Consolidation

Consolidation of storage, servers, and networks has enabled centralization of data center components, any-to-any access, simplified management, and technology convergence. Consolidation has also reduced the cost of data center deployment.

Virtualization

Virtualization is the creation of another abstraction layer that separates physical from logical characteristics and enables further automation of data center services. Almost any component of a data center can now be virtualized—storage, servers, networks, computing resources, file systems, file blocks, tape devices, and so on. Virtualization in a data center enables creation of huge resource pools, increased, efficient, and more flexible resource utilization, and automated resource provisioning, allocation, and assignment to applications. Virtualization represents the foundation for further automation of data center services.

Automation

Automation of data center services has been made possible by consolidating and virtualizing data center components. The advantages of data center automation are automated dynamic resource provisioning, automated Information Lifecycle Management (ILM), and automated data center management. Computing and networking resources can be automatically provisioned whenever needed. Other data center services can also be automated, such as data migration, mirroring, and volume management functions. Monitoring data center resource utilization is a necessary condition for an automated data center environment.

Phase 3

Converged Network

Converged networks promise the unification of various networks and single all-purpose communication applications. Converged networks potentially lead to reduced IT cost and increased user productivity. A unified data center fabric is based on a unified I/O transport protocol, which could potentially transport SAN, LAN and WAN, and clustering I/Os. Most protocols today tend to be transported across a common unified I/O channel and common hardware and software components of the data center architecture.

Energy-Efficient Data Center

A green data center is a data center in which the mechanical, lighting, electrical, and computer systems are designed for maximum energy efficiency and minimum environmental impact. The construction and operation of a green data center includes advanced technologies and strategies, such as minimizing the footprints of the buildings and using low-emission building materials, sustainable landscaping, and waste recycling. Installing catalytic converters on backup generators is also helpful. The use of alternative energy technologies, such as photovoltaic, heat pumps, and evaporative cooling, is also being considered.

Building and certifying a green data center or other facility can be expensive initially, but long-term cost savings can be realized in operations and maintenance. Another advantage is the fact that green facilities offer employees a healthy, comfortable work environment. There is growing pressure from environmentalists, and increasingly from the general public, for governments to offer green incentives. This pressure is in terms of monetary support for the creation and maintenance of ecologically responsible technologies.

Today, green data centers provide an 85 percent power reduction using virtualized, integrated modules. Rack space is saved with virtualized, integrated modules, and additional savings are derived from reduced cabling, port consumption, and support cost.

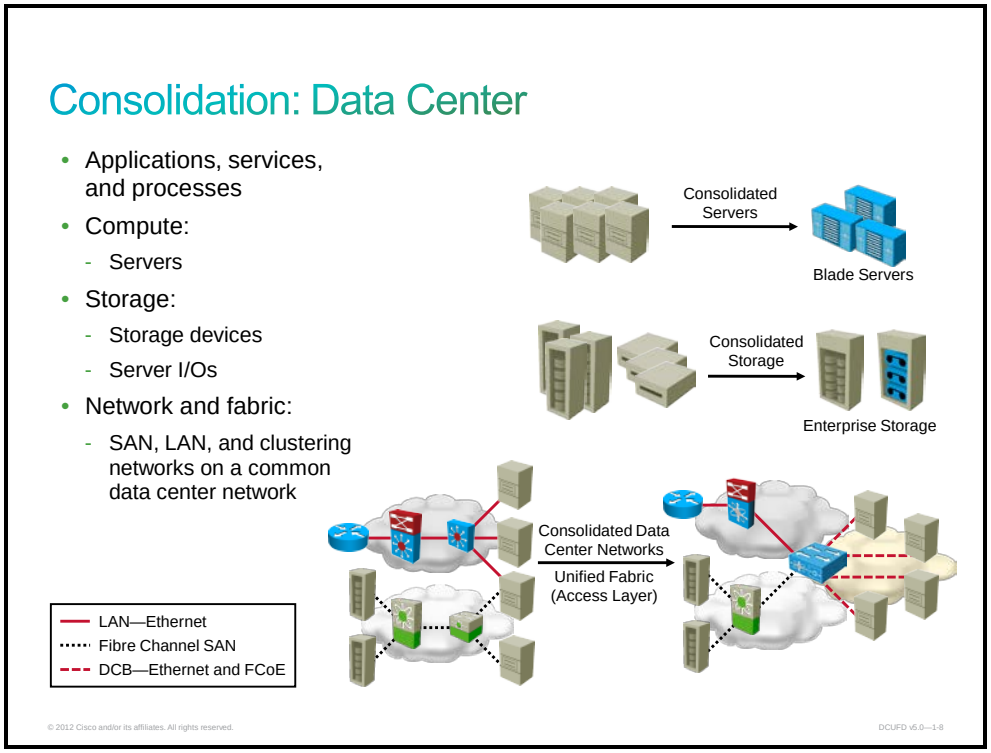
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

Service Integration

The data center network infrastructure is integrating network intelligence and is becoming application-agnostic, security-agnostic, computing-agnostic, and storage-agnostic by integrating application services, security services, computing services, and storage services, among others.



Consolidation is defined as the process of bringing together disconnected parts to make a single and complete whole. In the data center, it means replacing several small devices with a few highly capable pieces of equipment to provide simplicity.

The primary reason for consolidation is the sprawl of equipment and processes that is required to manage the equipment. It is crucial to understand the functions of each piece of equipment before consolidating it. There are various important reasons for server, storage, server I/O, network, application, and process consolidation:

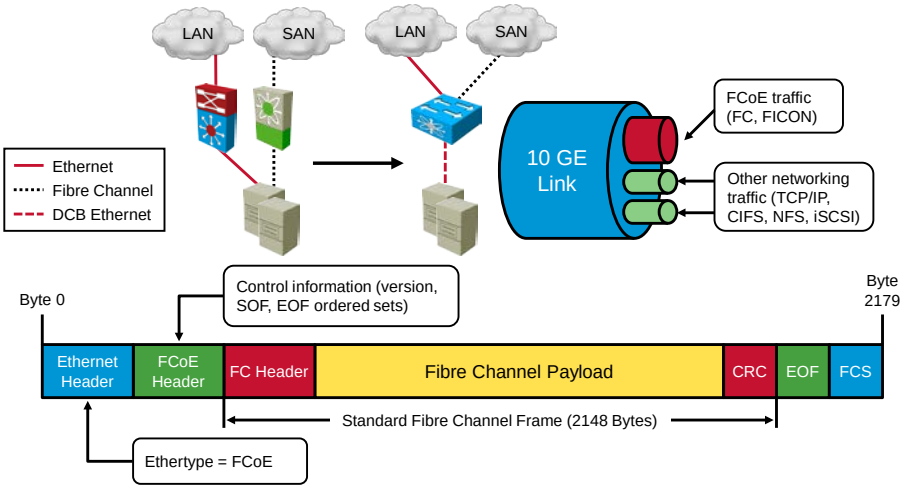
- Reduced number of servers, storage devices, networks, cables, and so on
- Increased usage of resources using resource pools (of storage and computing resources)
- Reduced centralized management
- Reduced expenses due to a smaller number of equipment pieces needed
- Increased service reliability

https://t.me/learningnets

certcollecion.net

Consolidation: Server I/O

- DCB enables deployment of converged unified data center fabrics:
  - Consolidates Ethernet and FCoE server I/O into common Ethernet



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

Server I/O consolidation has been attempted several times in the past with the introduction of Fibre Channel and iSCSI protocols that carry storage I/Os, data I/Os, and clustering I/Os across the same channel. All initial attempts have been unsuccessful.

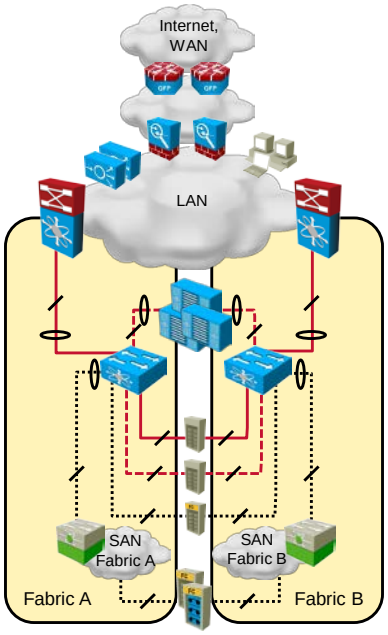
Enhanced Ethernet is a new converged network protocol that is designed to transport unified data and storage I/Os. Primary enabling technologies are Peripheral Component Interconnect Express (PCIe) and 10 Gigabit Ethernet. A growing demand for network storage is influencing network bandwidth demands. Server virtualization allows the consolidation of multiple applications on a server, therefore influencing the server bandwidth requirement of 10 Gb/s. 10-Gb/s Data Center Bridging (DCB) uses copper and twinax cables with short distances (32.8 feet [10 m]), but with lower cost, lower latency, and lower power requirements than 10BASE-T. FCoE and classical Ethernet can be multiplexed across the common physical DCB connection.

https://t.me/learningnets

certcollecion.net

Data Center Architecture

- Data center connects to other IT segments:
  - Campus LAN
  - Internet and WAN edge
- How components fit together:
  - Various connectivity options
  - Segments with different functionality



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

Data center architecture is the blueprint of how components and elements of the data center are connected. The components need to properly interact in order to deliver application services.

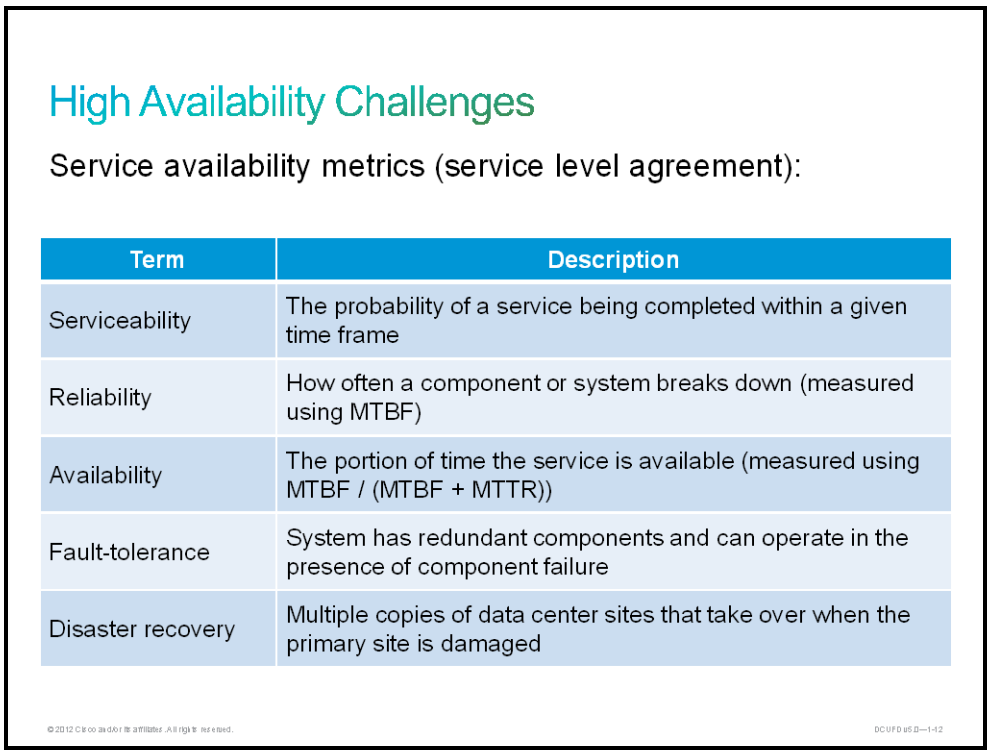
The data center is one of the components of the IT infrastructure. The data center needs to connect to other segments to deliver application services and enable users to access and use them. Such segments are Internet and WAN edge, campus LAN, and various demilitarized zone (DMZ) segments hosting public or semi-public services.

https://t.me/learningnets

certcollecion.net

Data Center Terminology

This topic describes the baseline technology and terminology used in data center solutions.



Protecting against failure is expensive, as is downtime. It is important to identify the most serious causes of service failure and to build cost-effective safeguards against them. A high degree of component reliability and data protection with redundant disks, adapters, servers, clusters, and disaster recovery decreases the chances of service outages. Data center architecture that provides service availability is a combination of several different levels of data center high-availability features, and depends on the following:

- **Serviceability:** Serviceability is the probability of a service being completed within a given time window. For example, if a system has serviceability of 0.98 for 3 hours, then there is a 98 percent probability that the service will be completed within 3 hours. In an ideal situation, a system can be serviced without any interruption of user support.
- **Reliability:** Reliability represents the probability of a component or system not encountering any failure over a time span. The focus of reliability is to make the data center system components unbreakable. Reliability is a component of high availability that measures how rarely a component or system breaks down, and is expressed as the mean time between failures (MTBF). For example, a battery may have a useful life of 10 hours, but its MTBF is 50,000 hours. In a population of 50,000 batteries, this translates into one battery failure every hour during its 10-hour life span. Mean time to repair (MTTR) is the average time that is required to complete a repair action. A server with 99 percent reliability will be down for 3.65 days every year.

  For a system with 10 components, where each component can fail independently but have 99 percent reliability, the reliability of the entire system is not 99 percent. The entire system reliability is 90.44 percent (0.99 to the 10th power). This translates to 34.9 days of downtime in one year. Hardware reliability problems cause only 10 percent of the downtime. Software, human, or process-related failures comprise the other 90 percent.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

- **Availability:** Availability is the portion of time that an application or service is available for productive work. A more resilient system results in higher availability. An important decision to consider when building a system is the required availability level, which is a compromise between the downtime cost and the cost of the high-availability configuration. Availability measures the ability of a system or group of systems to keep the application or service operating. Designing for availability assumes that the system will fail and that the system is configured to mask and recover from component-to-server failures with minimum application outage. Availability could be calculated as follows:

  Availability = MTBF / (MTBF + MTTR)

  or

  Availability = Uptime / (Uptime + Downtime)

  Achieving the property of availability requires either building very reliable components (high MTBF) or designing components and systems that can rapidly recover from failure (low MTTR). As downtime approaches zero, availability approaches 100 percent.

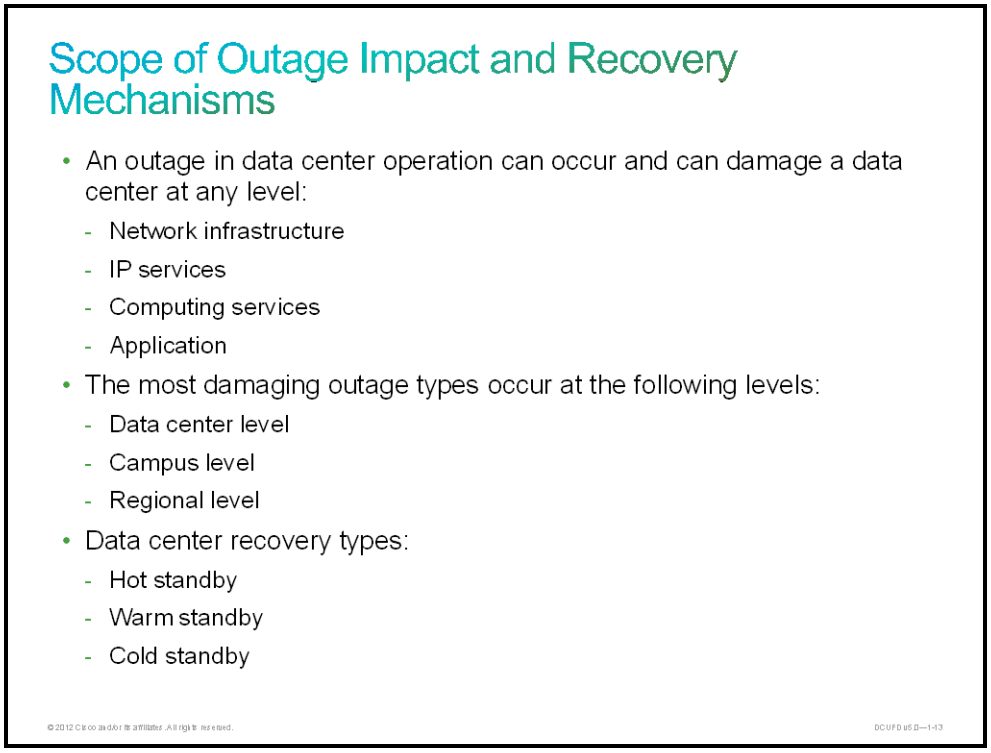
- **Fault-tolerant:** Fault-tolerant systems are systems that have redundant hardware components and can operate in the presence of individual component failure. Several components of the system have built-in component redundancy. Clusters are also examples of a fault-tolerant system. Clusters can provide uninterrupted service despite node failure. If a node that is running on one or more applications fails, one or more nodes in the cluster take over the applications from the failed server. A fault-tolerant server has a fully replicated hardware design that allows uninterrupted service in the event of component failure. The recovery time or performance loss that is caused by a component failure is close to zero, and information and disk content are preserved. The problem with fault-tolerant systems is that the system itself is a single point of failure.

- **Disaster recovery:** Disaster recovery is the ability to recover a data center at a different site if a disaster destroys the primary site or otherwise makes the primary site inoperable. Disaster, in the context of online applications, is an extended period of outage of mission-critical service or data that is caused by events such as fire or attacks that damage the entire facility. A disaster recovery solution requires a remote, mirrored (backup and secondary data center) site where business and mission-critical applications can be started within a reasonable period of time after the destruction of the primary site.

  Setting up a new, off-site facility with duplicate hardware, software, and real-time data synchronization enables organizations to quickly recover from a disaster at the primary site. The data center infrastructure must deliver the desired Recovery Point Objective (RPO) and Recovery Time Objective (RTO). RTO determines how long it takes for a certain application to recover, and RPO determines to which point (in backup and data) the application can recover. These objectives also outline the requirements for disaster recovery and business continuity. If these requirements are not met in a deterministic fashion, an enterprise carries significant risk in terms of its ability to deliver on the desired SLAs. SLAs are fundamental to business continuity. Ultimately, SLAs define your minimum levels of data center availability and often determine what actions will be taken in the event of a serious disruption. SLAs record and prescribe the levels of service availability, serviceability, performance support, and other attributes of the service, such as billing and even penalties, in the case of violation of the SLAs. For example, SLAs can prescribe different expectations in terms of guaranteed application response time, such as 1, 0.5, or 0.1 second. The SLA can also prescribe guaranteed application resource allocation time, such as 1 hour or automatic, and guaranteed data center availability, such as 99.999, 99.99, or 99.9 percent. Higher levels of guaranteed availability imply higher SLA charges.

https://t.me/learningnets

certcollecion.net



An outage in the data center operation can occur at any of the following levels:

- **Network infrastructure:** Data centers are built using sufficient redundancy so that a failure of a link or device does not affect the functioning of the data center. The infrastructure is transparent to the layers above it and aims to provide the continuous connectivity needed by the application and users.
- **IP services:** The IP protocol and services provide continuous reachability and autonomous path calculation so that traffic can reach the destination across multiple paths, if they are available. The most important components in this case are IP routing protocols.
- **Computing services:** Clusters of servers running in redundant operation are deployed to overcome a failure of a physical server. Clustered applications work in such a way that the servers are synchronized. Server virtualization clusters have the capability to start computing workloads on a different physical server if one is not available.
- **Application level:** Redundancy needs to be built into the application design.

There are different types of outages that might be expected and might affect the data center functions and operation. Typically, the types of outages are classified based on the scope of the outage impact:

- **Outage with an impact at the data center level:** An outage of this type is an outage of a system or a component such as hardware or software. These types of outages can be recovered using reliable, resilient, and redundant data center components, using fast routing and switching reconvergence, and stateful module and process failovers.
- **Outage with an impact at the campus level:** This type of outage affects a building or an entire campus. Fire or loss of electricity can cause damage at the campus level and can be recovered using redundant components such as power supplies and fans, or by using the secondary data center site or Power over Ethernet (PoE).

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

- **Outage with an impact at the regional level:** This type of outage affects a region, such as earthquakes, widespread power outage, flooding, or tornados. Such outages can be recovered using geographically dispersed, standby data centers that use global site selection and redirection protocols to seamlessly redirect user requests to the secondary site.

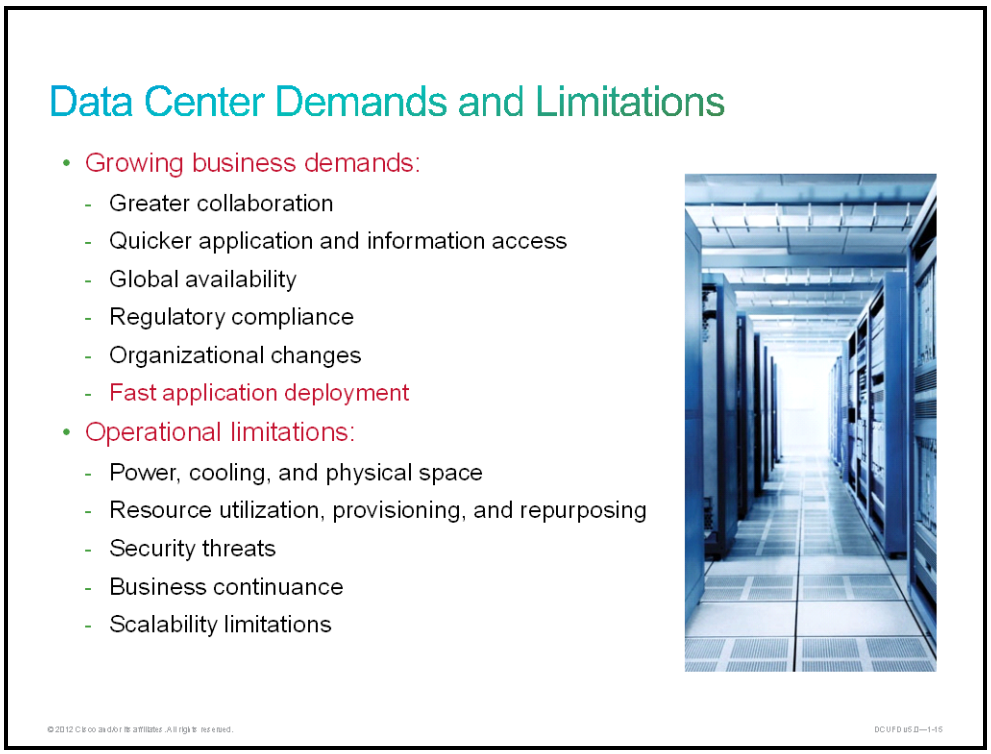
Data center recovery types: Different types of data center recovery provide different levels of service and data protection, such as cold standby, warm standby, hot standby, immediate recovery, continuous availability, continuous operation, gradual recovery, and a back-out plan.

https://t.me/learningnets

certcollecion.net

Data Center Challenges

This topic describes how to analyze business, technical, and environmental challenges.



The modern enterprise is being changed by shifting business pressures and operational limitations. While enterprises get ready to meet demands for greater collaboration, quicker access to applications and information, and ever-stricter regulatory compliance, they are also being pressured by issues relating to power and cooling, efficient asset utilization, escalating security and provisioning needs, and business continuance. All these concerns are central to data centers.

Modern data center technologies, such as multicore CPU servers and blade servers, require more power and generate more heat than older technologies, and moving to new technologies can significantly affect data center power and cooling budgets.

The importance of security has been rising as well, because more services are concentrated in a single data center. If an attack occurred in such a condensed environment, many people could be put out of work, resulting in lost time (and revenue). As a result, thorough traffic inspection is required for inbound data center traffic.

Security concerns and business continuance must be considered in any data center solution. A data center should be able to provide services if an outage occurs because of a cyber attack or because of physical conditions such as floods, fires, earthquakes, and hurricanes.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

Organizational Challenges

Organization	Challenges
Chief Officer	“I need to take a long-term view ... and have short-term wins. I want to see more business value out of IT.”
Applications Department	“Our applications are the face of our business.” “It is all about keeping the application available.”
Server Department	“As long as my servers are up, I am OK.” “We have too many underutilized servers.”
Security Department	“Our information is our business. We need to protect our data everywhere—in transit and at rest.”
Storage Department	“I cannot keep up with the amount of data that needs to be backed up, replicated, and archived.”
Network Department	“I need to provide lots of bandwidth and meet SLAs for application uptime and responsiveness.”



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

The data center is viewed from different perspectives, depending on the position of the person whose view is being expressed.

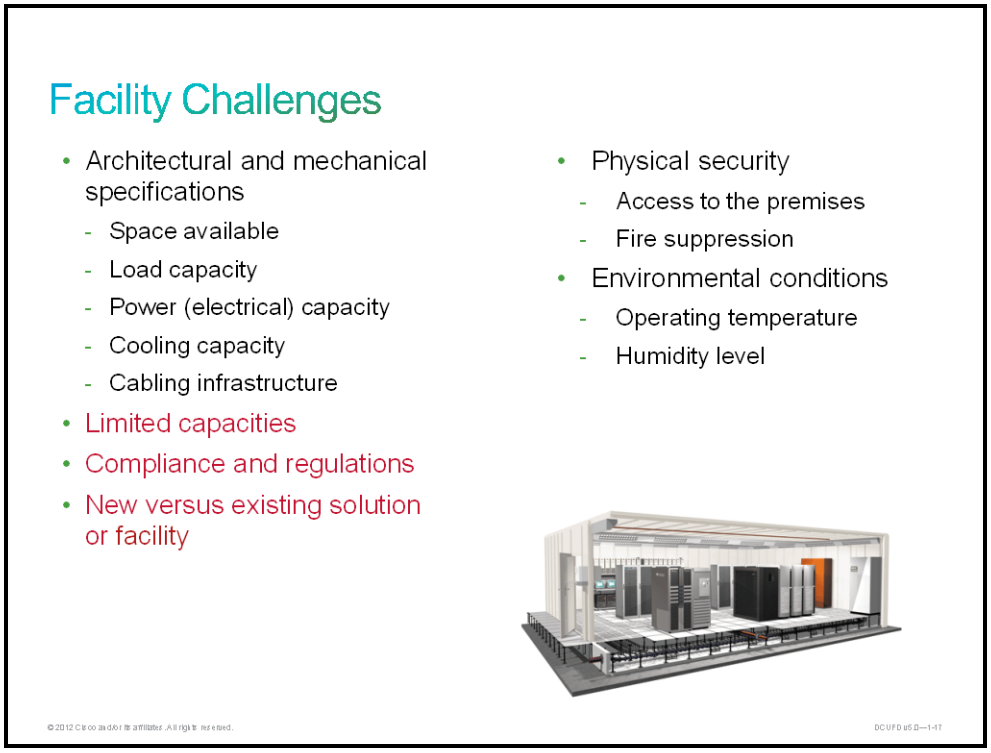
Depending on which IT team you are speaking with, you will find different requirements. You have the opportunity to talk on all levels because of your strategic position and the fact that you interact with all of the different components in the data center.

The data center involves multiple stakeholders, who all have different agendas and priorities. The traditional network contacts might not be the people who make the decisions that ultimately determine how the network evolves.

The organization might be run in silos, where each silo has its own budget and power base. Conversely, many next-generation solutions involve multiple groups.

https://t.me/learningnets

certcollecion.net



The data center facility has multiple aspects that need to be properly addressed (that is, taken into account when the facility is being planned, designed, and built).

Facility capacities are limited and need to be properly designed.

Companies must also address regulatory issues, enable business resilience, and comply with environmental requirements.

Data centers need infrastructures that can protect and recover applications, communications, and information, and that can provide uninterrupted access.

When it comes to building a reliable data center and maximizing an investment, the design must be considered early in the building development process. The design should include coordinated efforts that cut across several areas of expertise including telecommunications, power, architectural components, and heating, ventilation, and air conditioning (HVAC) systems.

Each of the components of the data center and its supporting systems must be planned, designed, and implemented to work together to ensure reliable access while supporting future requirements. Neglecting any aspect of the design can render the data center vulnerable to cost failures, early obsolescence, and intolerable levels of availability. There is no substitute for careful planning and following the guidelines for data center physical design.

Architectural and Mechanical Specifications

The architectural and mechanical facility specifications are defined as follows:

- How much space is available
- How much load that a floor can bear
- The power capacity that is available for data center deployment
- The cooling capacity that is available
- The cabling infrastructure type and management

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

In addition, the facility must meet certain environmental conditions. The types of data center devices define the operating temperatures and humidity levels that must be maintained.

Physical Security

Finally, physical security is vital because the data center typically houses data that should not be available to third parties, so access to the premises must be well controlled. Protection from third parties is important, as well as protection of the equipment and data from certain disasters. Fire suppression equipment and alarm systems to protect against fires should be in place.

Space

The space aspect involves the physical footprint of the data center. Space issues include how to size the data center, where to locate servers within a multipurpose building, how to make it adaptable for future needs and growth, and how to construct the data center to effectively protect the valuable equipment inside.

The data center space defines the number of racks that can be used and thus the equipment that can be installed. That is not the only important parameter. The floor-loading capability is equally important, and determines which and how much equipment can be installed in a certain rack, and thus what the rack weight should be. The placement of current and future equipment must be very carefully considered so that the data center physical infrastructure and support is deployed optimally.

Although sometimes neglected, the size of the data center has a great influence on cost, life span, and flexibility. Determining the proper size of the data center is a challenging and essential task that should be done correctly and must take into account several variables:

- The number of people supporting the data center
- The number and type of servers and the storage and networking equipment that is used
- The sizes of the non-server, storage, or network areas, which depend on how the passive infrastructure is deployed

A data center that is too small will not adequately meet server, storage, and network requirements and will thus inhibit the productivity and will incur additional costs for upgrades or expansions.

Alternatively, a data center that is too spacious is a waste of money, not only from the initial construction cost but also from the perspective of ongoing operational expenses.

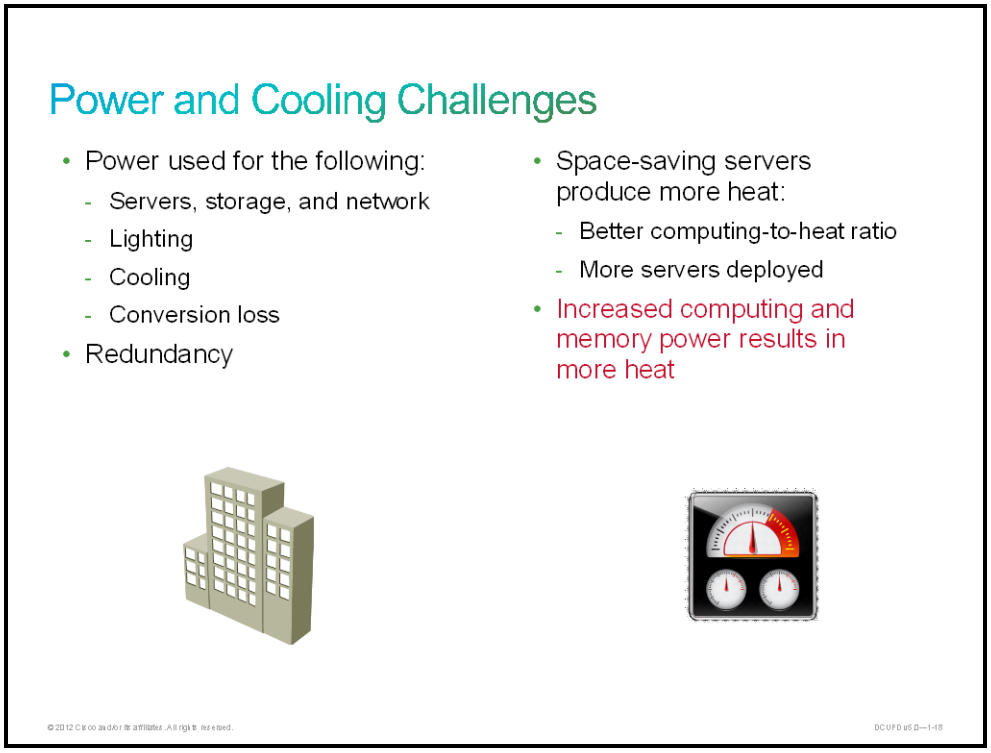
Properly sized data center facilities also take into account the placement of equipment. If properly selected, the data center facility can grow when needed. Otherwise, costly upgrades or relocations must be performed.

Cabinets and racks are part of the space requirements and other aspects:

- Loading, which determines what and how many devices can be installed
- The weight of the rack and equipment that is installed
- Heat that is produced by the equipment that is installed
- Power that is consumed by the equipment that is installed

https://t.me/learningnets

certcollecion.net



Power

The power in the data center facility is used to power servers and also storage, network equipment, lighting, and cooling devices (which take up most of the energy), and some power is “lost” upon conversion.

The variability of usage is difficult to predict when determining power requirements for the equipment in the data center. For the server environment, the power usage depends on the computing load. If the computing power of the server must work harder, more power that has to be drawn from the AC supply and more heat output needs to be dissipated.

Power requirements are based on the desired reliability and may include two or more power feeds from the utility, an uninterruptible power supply, multiple circuits to systems and equipment, and on-site generators. Determining power requirements requires careful planning.

Estimating power needs involves determining the power that is required for all existing devices and for devices that are anticipated in the future. Power requirements must also be estimated for all support equipment such as the uninterruptible power supply, generators, conditioning electronics, HVAC system, lighting, and so on. The power estimation must be made to accommodate required redundancy and future growth.

The facility electrical system must not only power data center equipment (servers, storage, network equipment, and so on) but must also insulate the equipment against surges, utility power failures, and other potential electrical problems (thus addressing the redundancy requirements).

The power system must physically accommodate electrical infrastructure elements such as power distribution units (PDUs), circuit breaker panels, electrical conduits, wiring, and so on.

Cooling

The temperature and humidity conditions must be controlled and considered by deploying probes to measure temperature fluctuations, data center hotspots, and relative humidity, and by using smoke detectors.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

Overheating is an equipment issue with high-density computing:

- More heat overall
- Hotspots
- High heat and humidity, which threaten equipment life spans
- Computing power and memory requirements, which demand more power and generate more heat
- Data center demand for space-saving servers: density = heat; 3 kilowatt (kW) per chassis is not a problem for one chassis, but five or six chassis per rack = 20 kW.
- Humidity levels that affect static electricity and condensation must be considered. Maintaining a 40 to 55 percent relative humidity level is recommended

The cooling aspect of the facilities must have proper airflow to reduce the amount of heat that is generated by concentrated equipment. Adequate cooling equipment must be used for more flexible cooling. Additionally, the cabinets and racks should be arranged in an alternating pattern to create “hot” and “cold” aisles. In the cold aisle, equipment racks are arranged face to face. In the hot aisle, the equipment racks are arranged back to back. Perforated tiles in the raised floor of the cold aisles allow cold air to be drawn into the face of the equipment. This cold air washes over the equipment and is expelled out of the back into the hot aisle. In the hot aisle, there are no perforated tiles. This fact keeps the hot air from mingling with the cold air.

Because not every active piece of equipment exhausts heat out of the back, other considerations for cooling include the following:

- Increasing airflow by blocking unnecessary air escapes or by increasing the height of the raised floor
- Spreading equipment out over unused portions of the raised floor if space permits
- Using open racks instead of cabinets when security ID is not a concern, or using cabinets with mesh fronts and backs
- Using perforated tiles with larger openings

Helpful Conversions

One watt is equal to 3.41214 British thermal units (BTUs). This value is a generally used value for converting electrical values to BTUs and vice versa. Many manufacturers publish kilowatt (kW), kilovolt-ampere (kVA), and BTU measurements in their equipment specifications. Sometimes, dividing the BTU value by 3.412 does not equal the published wattage. Where the information is provided by the manufacturer, use it. Where it is not provided, this formula can be helpful.

Increasing Heat Production

Although the blade server deployment optimizes the computing-to-heat ratio, the heat that is produced actually increases because the blade servers are space optimized and allow more servers to be deployed in a single rack. High-density equipment produces much heat.

In addition, the increasing computing and memory power of a single server results in higher heat production.

Thus, the blade server deployment results in more heat being produced, which requires proper cooling capacity and proper data center design. The solutions that address the increasing heat requirements must be considered when blade servers are deployed within the data center. The design must also take into consideration the cooling that is required for the current sizing of the data center servers, but the design must anticipate future growth, thus also taking into account future heat production.

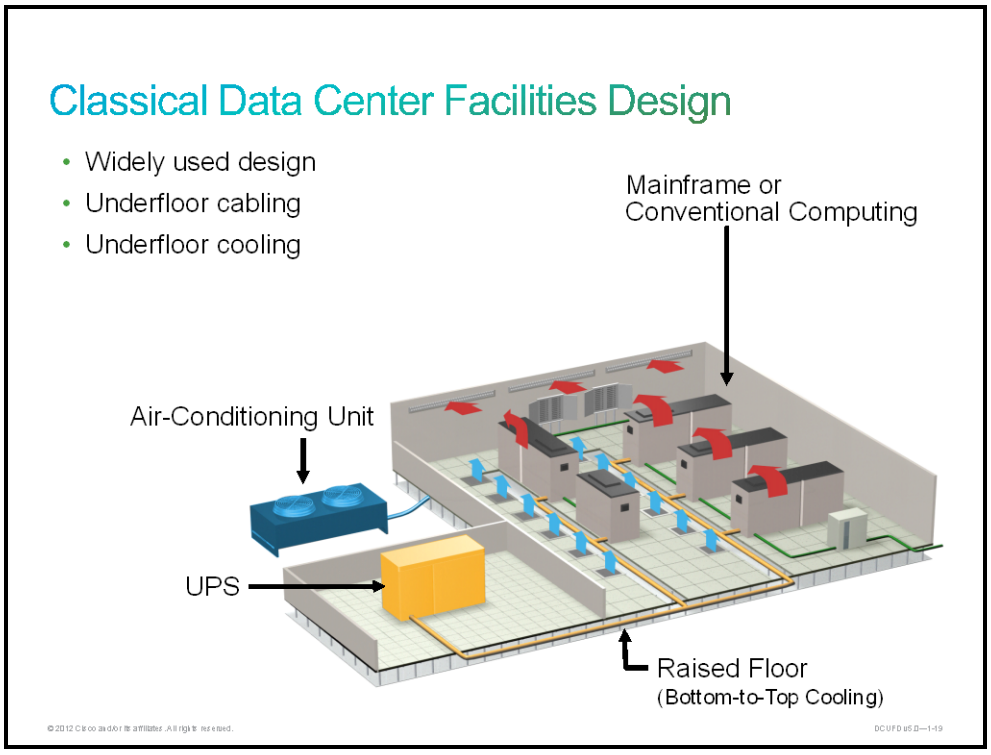
https://t.me/learningnets

certcollecion.net

If cooling is not properly addressed and designed, the result is a shortened equipment life span.

The cooling solution can resolve the increasing heat in these ways:

- Increasing the space between the racks and rows
- Increasing the number of HVAC units
- Increasing the airflow through the devices
- Using new technologies like water-cooled racks



Traditionally, the data center evolved primarily to support mainframe computing. A feature of this environment is that change is the exception, rather than the rule. This situation led to a Layer 0 infrastructure in which the following occurred:

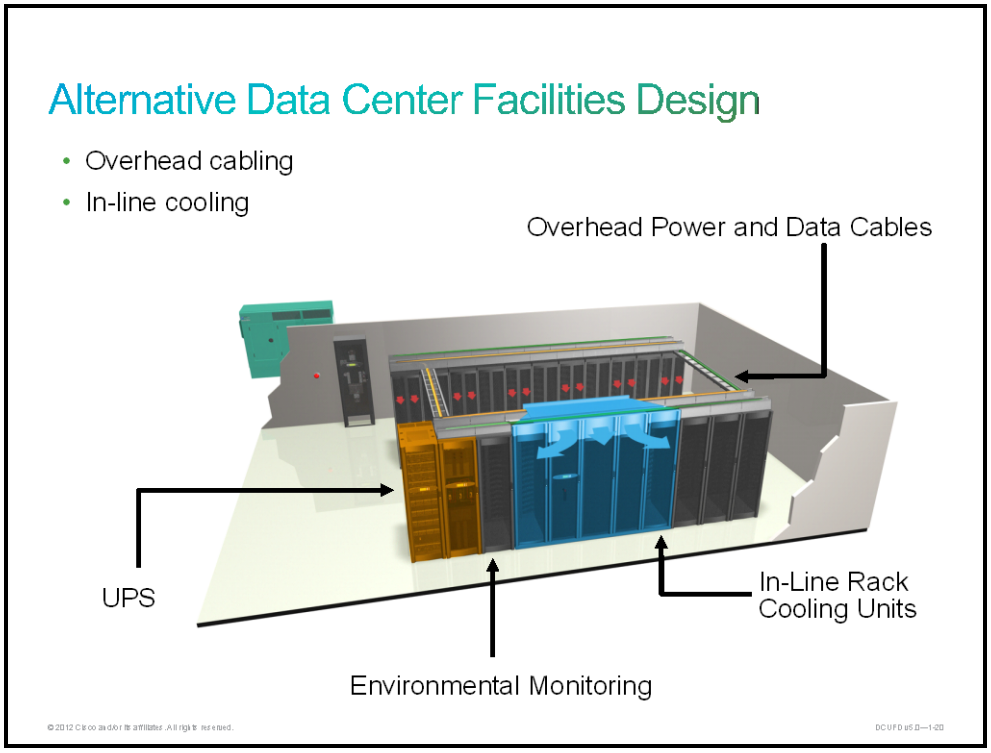
- Power and cooling were overprovisioned from the onset.
- The floor was raised for bottom-to-top cooling.
- Racks and cabling did not need to be reconfigured.
- Floor loading was predetermined.
- Raised floors were the norm.

The modern server-based environment is one where change is a constant, but the environmental infrastructure support is often relatively inflexible.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net



In the figure, the uninterruptible power supply (UPS in the figure) and air conditioning are no longer segregated from the computer equipment. There is no longer a raised floor, the power and data cabling is being run over the tops of the racks, and the power distribution panels are now within the racks.

Incorporating the physical layer into an on-demand framework achieves significant cost savings, and the modular nature of the framework results in much higher levels of availability versus investment. There are some additional benefits:

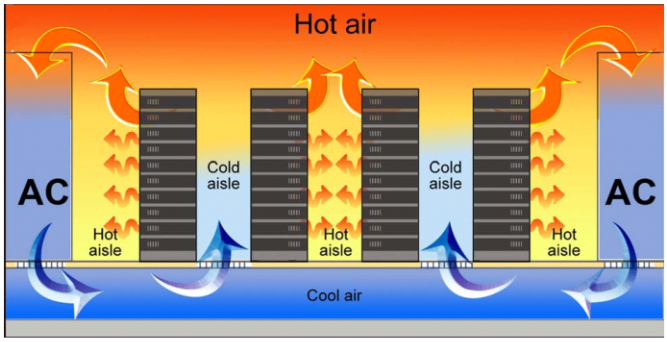
- Significantly lower TCO
- Ability to add power and cooling as needed
- Shorter deployment cycles
- Ability to provision facilities in the same manner as IT infrastructure
- Increased uptime

https://t.me/learningnets

certcollecion.net

Data Center Thermal Control Model

- Cold air is pumped from the air-conditioning units through the floor into the cold aisles between equipment racks.
- Cold air is pulled from the cold aisle through the racks and exits at the back of the equipment.
- The heat from the equipment racks exhausts into the hot aisles where it is returned to the air-conditioning units to be chilled and recycled.



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

A traditional data center thermal control model is based on controlled air flows. Before trying to solve the heat problem, it is important to understand what may be causing additional heat in the data center environment. To help dissipate heat from electrically active IT hardware, air that is treated by air conditioning is pumped into the floor space below the rows of server racks. The chilled air enters the data center through vented tiles in the floor and fills in the cold aisle front of the rows of racks. For data centers without raised floors, the chilled air enters through diffusers above the cold aisle. Although construction of data centers differs, the heat problem that they face is similar, and the technology described here applies to both. The built-in fans of the stored hardware pull the cool air through each rack, chilling the warm components (processors, power supplies, and so on) of the hardware. Heat is exchanged and the air becomes warm. The heated air then exits out the back of the rack into the hot aisle. The heated air is pulled back into the air-conditioning units and chilled. The cycle repeats, and the data center environment should be kept at a safe, cool temperature, as long as the BTU capacity of the air-conditioning units is sufficient to cool all the equipment that is installed.

Summary of the air flow in a data center facility:

- Cold air is pumped from the air-conditioning units through the raised floor of the data center and into the cold aisles between facing server racks.
- Air-conditioned air is pulled from the cold aisle through the racks and exits the back of the servers.
- The heat from the server racks exhausts into the hot aisles, where it is returned to the air-conditioning units to be chilled.

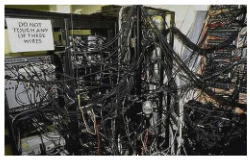
https://t.me/learningnets

certcollecion.net

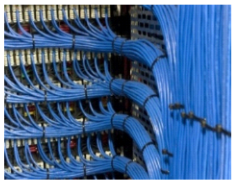
Cabling Challenges

- Passive networking infrastructure:
  - LAN and SAN physical network infrastructure
- Cabling characteristics:
  - Type (copper versus fiber optics)
  - Length
- Cabling should not:
  - Restrict air flow
  - Hinder troubleshooting
  - Create unplanned dependencies
  - Incur accidental downtime

Incorrect



Correct



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

DCUFD v5.0—1-22

The data center cabling (the passive infrastructure) is equally important for proper data center operation. The infrastructure needs to be a well-organized physical hierarchy that aids the data center operation. The electrical infrastructure is crucial for keeping server, storage, and network devices operating. The physical network, which is the cabling that runs and terminates between devices, dictates if and how these devices communicate with one another and the outside world.

The cabling infrastructure also governs the physical connector and the media type of the connector. Two options are widely used today: copper-based cabling and fiber optics-based cabling.

Fiber optics-based cabling is less susceptible to external interferences and offers greater distances, while copper-based cabling is ubiquitous and less costly. The cabling must be abundant to provide ample connectivity and must employ various media types to accommodate different connectivity requirements. However, but it must remain well organized for the passive infrastructure to be simple to manage and easy to maintain. (No one wants a data center where the cables are on the floor, creating a health and safety hazard.) Typically, the cabling needs to be deployed in tight spaces, terminating at various devices.

Cabling usability and simplicity are affected by the following:

- Media selection
- Number of connections provided
- Type of cabling termination organizers

These parameters must be addressed during the initial facility design, and the server, storage, and network components and all the technologies to be used must be considered.

The cabling infrastructure must not incur the following:

- Improper cooling due to restricted airflow
- Difficult-to-implement troubleshooting
- Unplanned dependencies that result in more downtime for single component replacement

https://t.me/learningnets

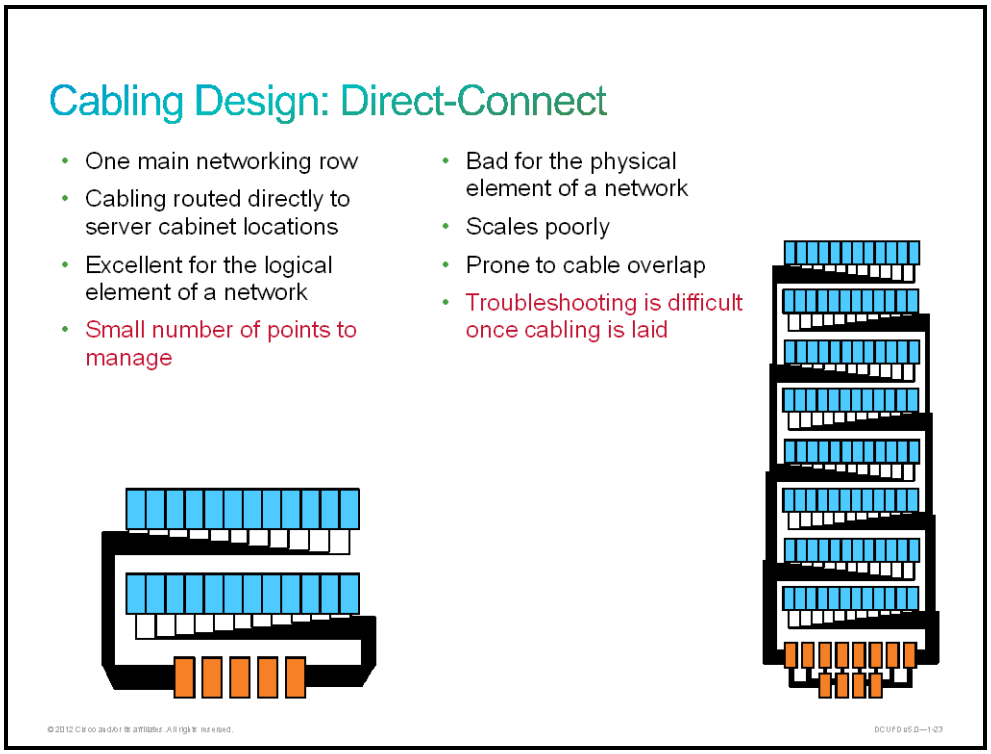
certcollecion.net

- Downtimes due to accidental disconnects

For example, with underfloor cabling, airflow is restricted by the power and data cables. Raised flooring is a difficult environment in which to manage cables because cable changes mean lifting floor panels and potentially having to move equipment racks.

The solution is a cable management system that consists of integrated channels that are located above the rack for connectivity. Cables should be located in the front or rear of the rack for easy access. Typically, cabling is located in the front of the rack in service provider environments.

When data center cabling is deployed, the space constraints and presence of operating devices (namely servers, storage, and networking equipment) make the cabling infrastructure reconfiguration very difficult. Thus, scalable cabling is crucial for proper data center operation and life span. Conversely, poorly designed cabling will incur downtime due to reconfiguration or expansion requirements that were not considered by the original cabling infrastructure shortcomings. Cable management is a major topic in its own right. The designer of a data center should work with the facilities team that installs and maintains the data center cabling in order to understand the implications of any new or reconfigured environment in the data center.



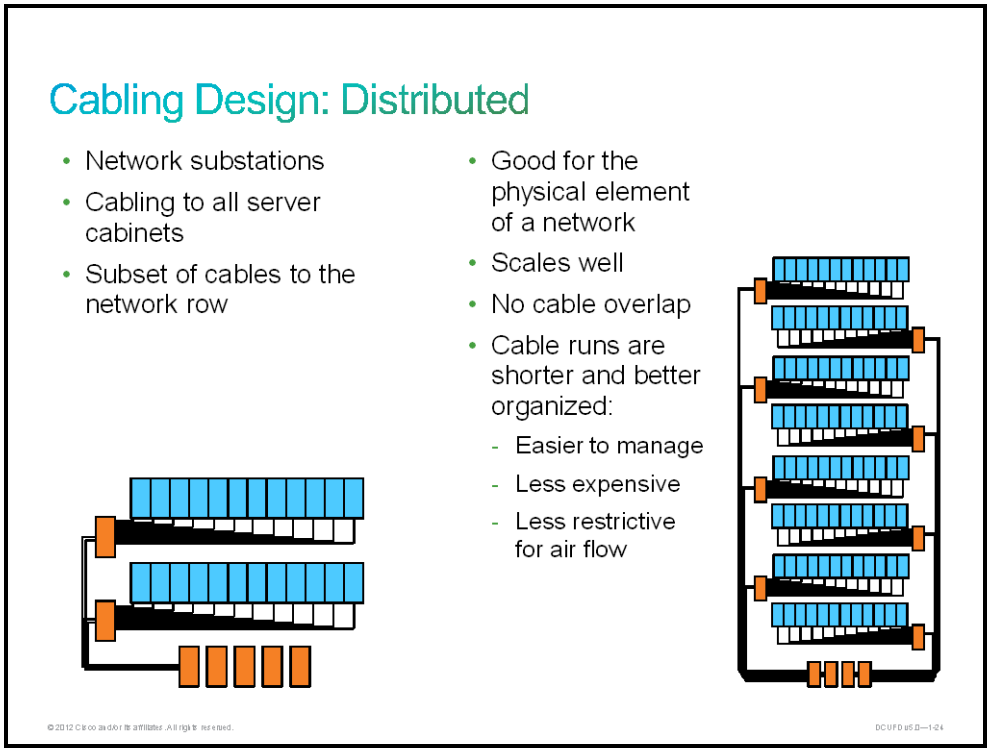
Smaller networks often use a direct-connect cabling design in which there is one main networking row. Cabling is routed directly from the networking row to server cabinet locations.

A direct-connect design is excellent for the logical element of a network, in terms of Layer 2 configuration, but a direct-connect design is not optimal for the physical element of a network (for example, a switch). A direct-connect design scales poorly and is prone to cable overlap. Enough redundancy must be provided within the physical network device.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

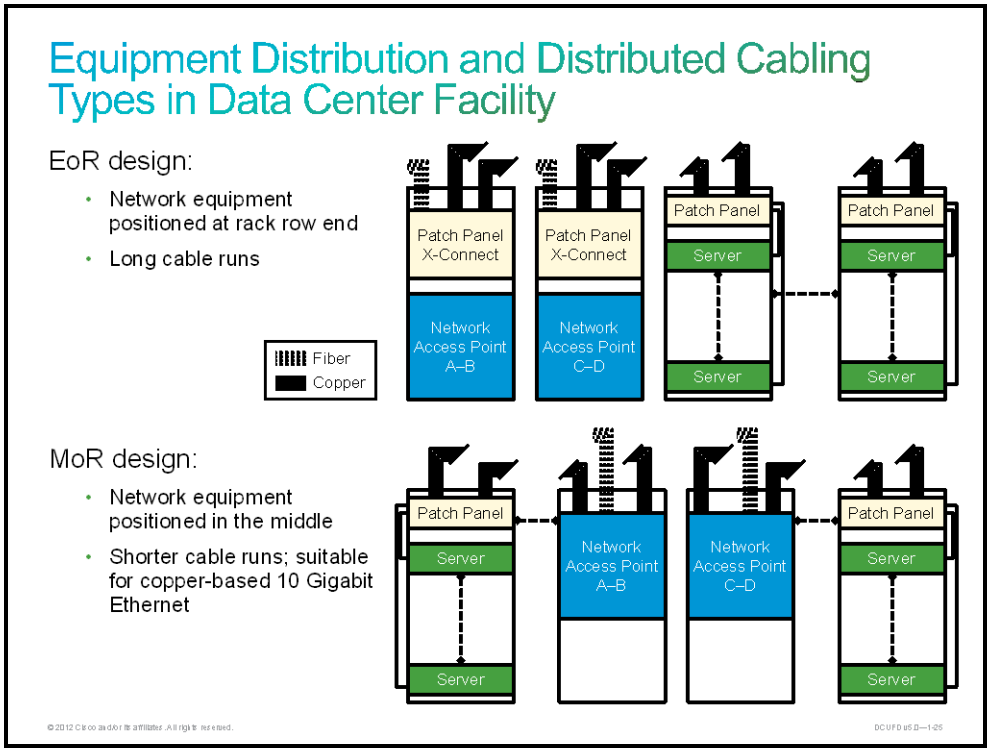
certcollecion.net



A distributed cabling design uses network substations within each server row. Server cabinets are cabled to the local substation, and from there to the main network row. This design is superior for the physical element of a network. This design scales well, and makes it possible to avoid cable overlap. Cable runs are shorter and better organized, which makes the physical network easier to manage, less expensive, and less restrictive for air flow.

https://t.me/learningnets

certcollecion.net



There are different types of network equipment distribution of the data center racks. The most typically used distribution type is end-of-row (EoR) distribution, with servers and network access layer switches connected using copper cables. The middle-of-row (MoR) and EoR design distributions are typically used for modular access, and cabling is done at data center build-out. These two types of rack distributions enable lower cabling distances (and lower cost) and allow a dense access layer. On average, these are 6 to 12 multi-RU servers per rack, requiring 4 to 6 kW per server rack and 10 to 20 kW per network rack. Every switch supports one or many medium or large subnets and VLANs.

EoR design characteristics are as follows:

- Copper from server to access switches
- Poses challenges on highly dense server farms:
  - Distance from farthest rack to access point
  - Row length may not lend itself well to switch port density

MoR design characteristics are as follows:

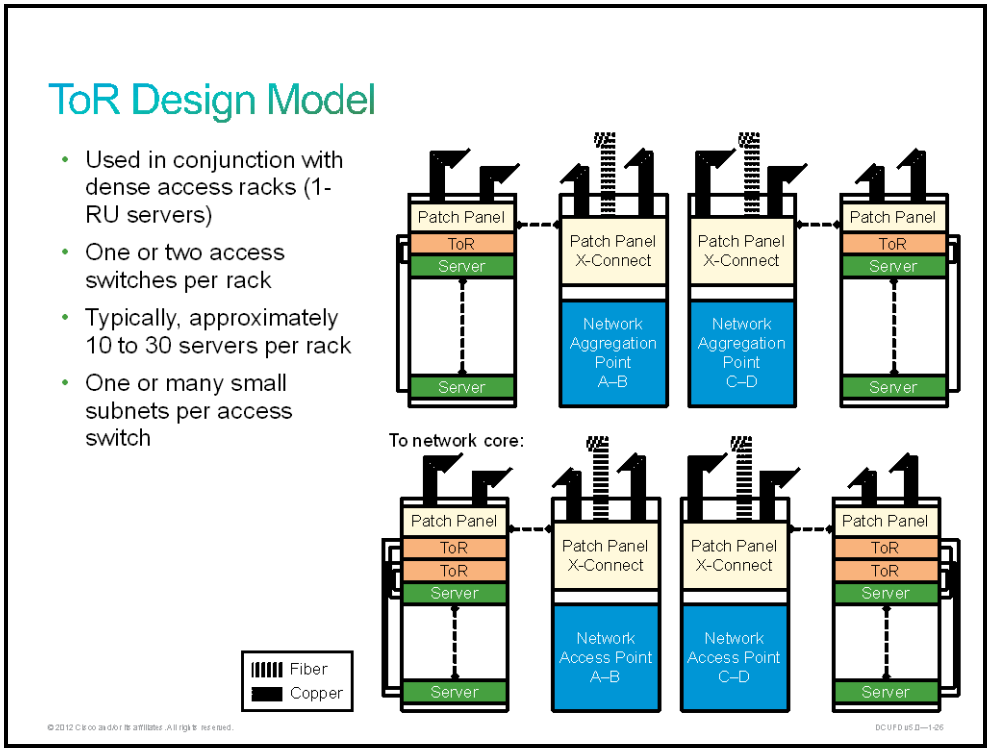
- Use is starting to increase, given EoR challenges
- Copper cabling from servers to access switches
- Fiber cabling may be used to aggregate top-of-rack (ToR) servers
- Addresses aggregation requirements for ToR access environments

Note The MoR approach is especially suitable for 10 Gigabit Ethernet environments. You can take advantage of twinax cables, which are economical, but extend 32.8 feet (10 m) in length at most. Suitable high-density 10 Gigabit Ethernet switches are the Cisco Nexus 7009 or Cisco Nexus 5596.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net



The ToR design model is appropriate for many 1- rack unit (RU) servers, making it much easier for cabling, compared to an EoR design. It is easier because the access switch is located on the top of every server rack and cabling occurs within the rack. Outside the server access rack, copper or fiber uplinks are utilized to the aggregation layer MoR switching rack. Every access layer switch is configured with one or more small subnets.

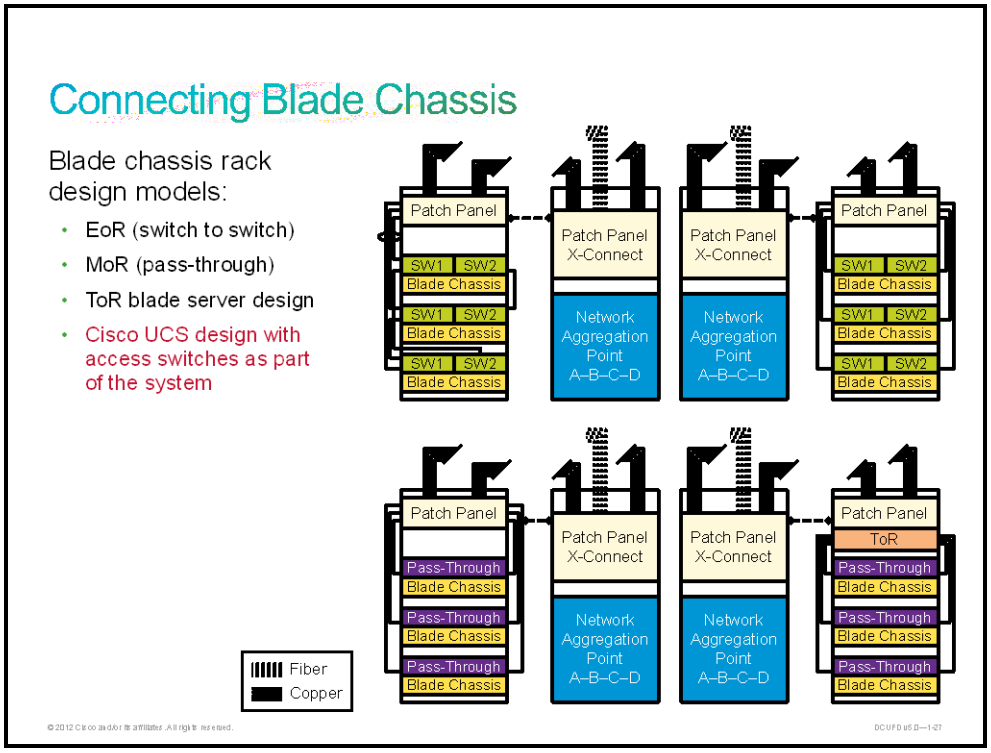
ToR design characteristics are as follows:

- Used with dense access racks (1-RU servers)
- Typically one access switch per rack (although some customers are considering two plus a cluster)
- Typically approximately 10 to 15 servers per rack (enterprises)
- Typically approximately 15 to 30 servers per rack (service providers)
- Use of either side of rack is becoming popular
- Cabling within rack: Copper for server to access switch
- Cabling outside rack (uplink):
  - **Copper (Gigabit Ethernet):** Needs an MoR model for fiber aggregation
  - **Fiber (Gigabit Ethernet or 10 Gigabit Ethernet):** More flexible and also requires aggregation model (MoR)
- Subnets and VLANs: One or many subnets per access switch

Note A suitable selection for ToR cabling is a fabric extender (FEX), such as the Cisco Nexus 2248, 2224, or 2232 FEX. From the perspective of the physical topology, they provide ToR connectivity. From the management perspective, they are managed from a single device, resembling the EoR design.

https://t.me/learningnets

certcollecion.net



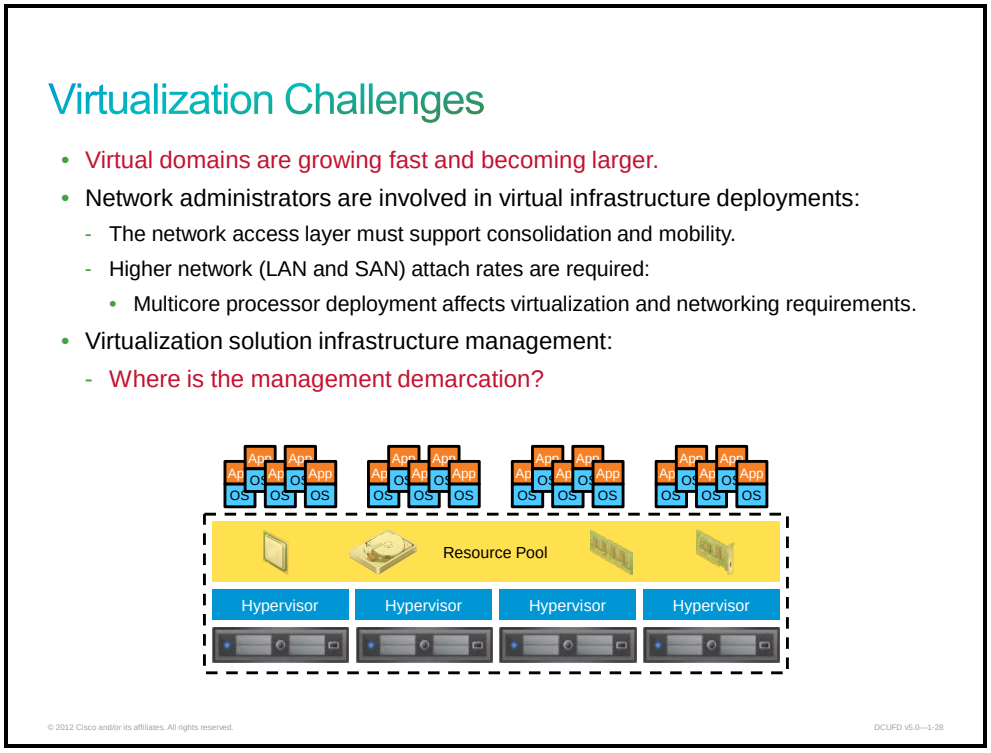
For blade chassis, the following four design models can be used:

- EoR design (switch-to-switch) characteristics:
  - Scales well for blade server racks (approximately three blade chassis per rack)
  - Most current uplinks are copper, but the new switches will offer fiber
- MoR (pass-through) design characteristics:
  - Scales well for pass-through blade racks
  - Copper from servers to access switches
- ToR blade server design characteristics:
  - Has not been used with blade switches
  - May be a viable option on pass-through environments if the access port count is right
- Cisco Unified Computing System:
  - The Cisco Unified Computing System features the fabric interconnect switches, which act as access switches to the servers
  - Connects directly to the aggregation layer

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net



Virtualization, although being the “promised solution” for server-, network-, and space-related problems, brings a few challenges.

- **The complexity factor:** The leveraging of high-density technologies, such as multicore processors, unified fabrics, higher-density memory formats, and so on, brings increased equipment and network complexity.
- **Support efficiency:** Trained personnel are required to support such networks, and the support burden is heavier. However, new-generation management tools ease these tasks.
- **The challenges of virtualized infrastructure:** These challenges involve management, common policies, security aspects, and adaptation of organizational processes and structures.

All these aspects require higher integration and collaboration from the personnel of the various service teams.

LAN and SAN Implications

The important challenge that server virtualization brings to the network is the loss of administrative control on the network access layer. By moving the access layer into the hosts, the network administrators have no insight into configuration or troubleshooting of the network access layer. On the other hand, when obtaining access, network administrators are faced with virtual interface deployments.

Second, by enabling mobility for VMs, the information about the VM connection point gets lost. If the information is lost, the configuration of the VM access port does not move with the machine.

Third, by using virtualization, the servers, network, and storage facilities are under increased loads.

Server virtualization results in multiple VMs being deployed on a single physical server. Though the resource utilization is increased, which is desired, this increase can result in more

https://t.me/learningnets

certcollecion.net

I/O throughput. When there is more I/O throughput, more bandwidth is required per physical server.

To solve this challenge, multiple interfaces are used to provide server connectivity.

- Multiple Gigabit Ethernet interfaces provide LAN connectivity for data traffic to flow to and from the clients or to other servers. Using multiple interfaces also ensures that the redundancy requirement is properly addressed.
- Multiple Fibre Channel interfaces provide SAN connectivity for storage traffic to allow servers, and therefore VMs, to access storage on a disk array.
- Typically, a dedicated management interface is also provided to allow server management.

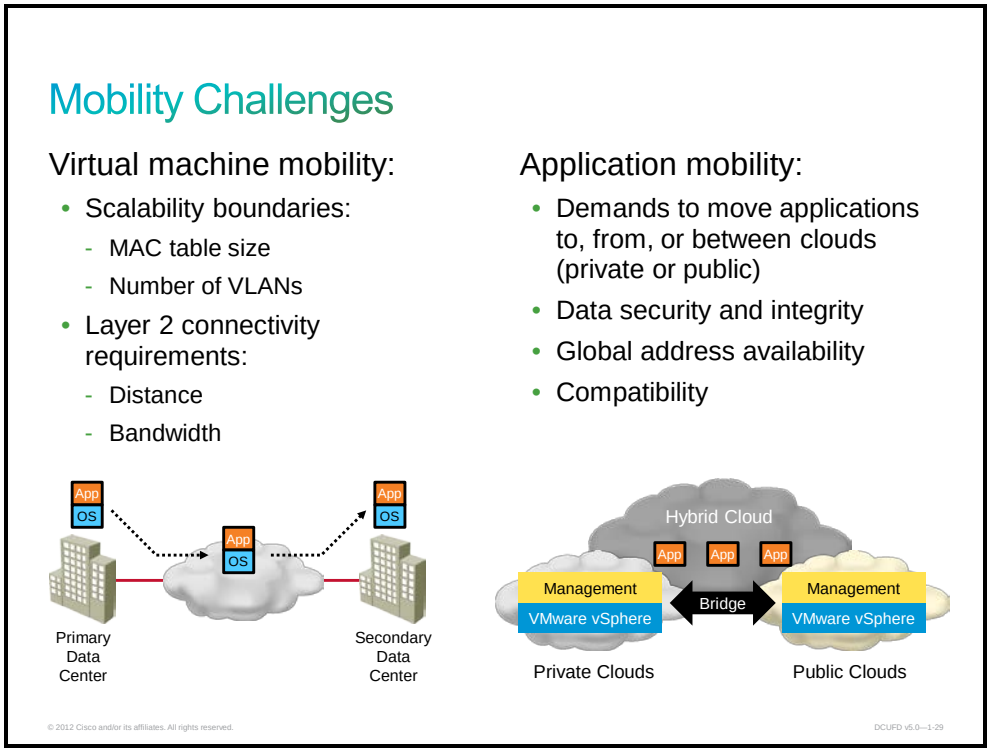
Virtualization thus results in a higher interface count per physical server, and with SAN and LAN infrastructures running in parallel, there are the following implications:

- The network infrastructure costs more and is less efficiently used.
- There are a higher number of adapters, cabling, and network ports, which results in higher costs.
- Multiple interfaces also cause multiple fault domains and more complex diagnostics.
- Having multiple adapters increases management complexity. More management effort is put into proper firmware deployment, driver patching, and version management.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net



These days mobility is of utmost importance. Everyone demands and requires it. IT infrastructure users and businesses demand to be able to access their applications and data from anywhere, which imposes new challenges. At the same time, IT needs to cut costs of the infrastructure, and thus more are considering the cloud.

The virtualization through addressing and solving many of the challenges of “classic” solutions brings its own piece to the mobility puzzle. VM mobility requires that data center architectures are properly conceived, and failure to do so may prevent proper VM mobility. From the VM perspective, the following items limit the solution architecture:

- **Scalability:** MAC address tables and the VLAN address space present a challenge when VMs need to move outside of their own environments. Moving outside the environment means moving the VM from primary to secondary data center, or even from a private IT infrastructure to a public one.
- To ensure proper VM operation, and thus operation of the application hosted by the VM, Layer 2 connectivity is commonly required between the segments where the VM moves. This introduces all sorts of challenges:
  - Distance limitation
  - Selection of an overlay technology that enables seamless Layer 2 connectivity
  - Unwanted traffic carried between sites (broadcast, unknown unicast, and so on) that consumes precious bandwidth
  - Extending IP subnets and split-brain problems upon data center interconnect failure

Today, application mobility means that users can access applications from any device. From the IT perspective, application mobility means the ability to move application load between IT infrastructures (that is, clouds). This issue imposes another set of challenges:

- Data security and integrity when moving application load from a controlled IT infrastructure to an outsourced infrastructure (that is, a public cloud)

https://t.me/learningnets

certcollecion.net

- As with VM mobility, access to applications, that is, enabling application access by the same name regardless of its location
- Because the IT infrastructures typically do not use single standardized underlying architecture, equipment, and so on, incompatibility within the infrastructure arises, which can limit or affect application mobility by requiring the use of conversion tools. This limitation diminishes seamless application mobility and limits the types of applications that can be moved (for example, critical applications do not allow downtime for the conversion process).

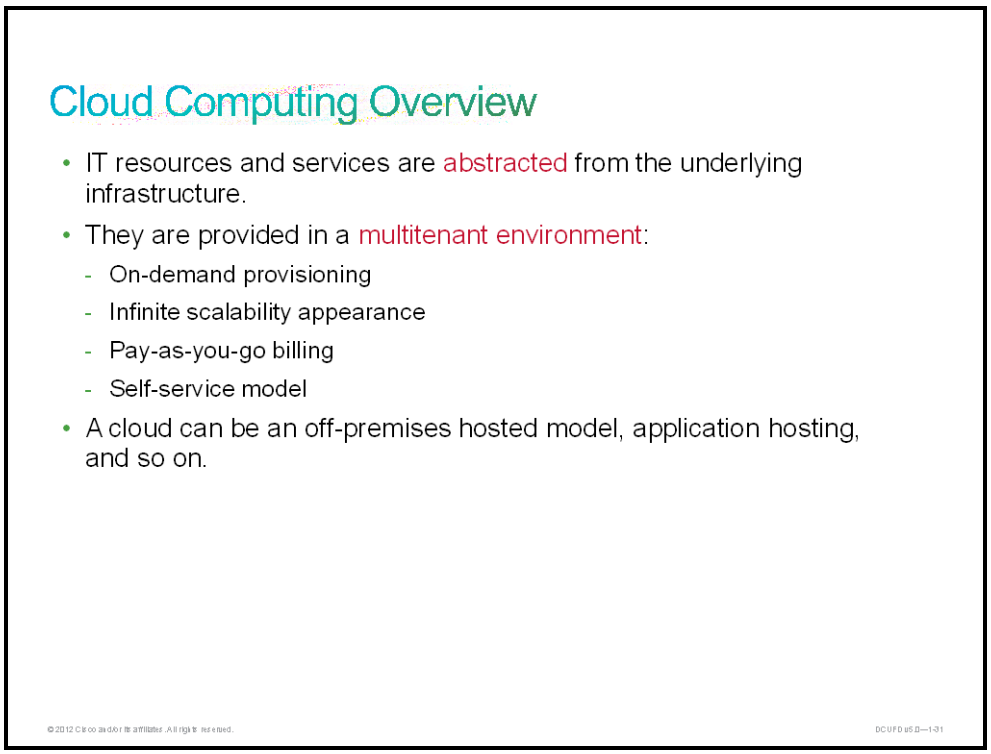
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

Introduction to Cloud Computing

This topic describes how to recognize the cloud computing paradigm, terms, and concepts.



Cloud computing translates into IT resources and services that are abstracted from the underlying infrastructure and provided on demand and “at scale” in a multitenant environment. Today, clouds are associated with an off-premises, hosted model.

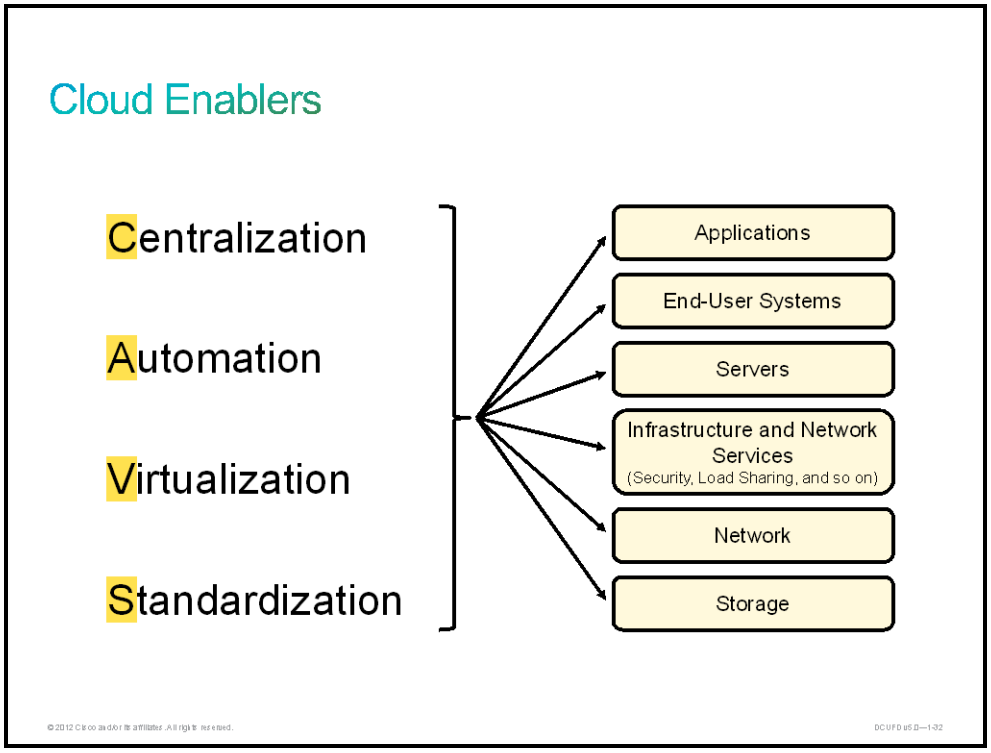
Cloud computing is still new. There are no definite and complete standards yet because the standards are in the process of being defined.

There are some cloud solutions that are available and offered, but not at the scale and ubiquity that is anticipated in the future.

One could argue that a component cloud is very much like a virtualized data center. This description is usually true. However, there are some details that differentiate a component cloud from a virtualized data center. One of those details is on-demand billing. Cloud resource usage is typically tracked at a granular level and billed to the customer on a short interval. Another difference is that, in a cloud, associations of resources to an application are more dynamic and ad hoc than in a virtual data center. However, these differences and definitions are very fragile. Rather than trying to define these differences, it is more useful to focus on the value than on the concepts.

https://t.me/learningnets

certcollecion.net



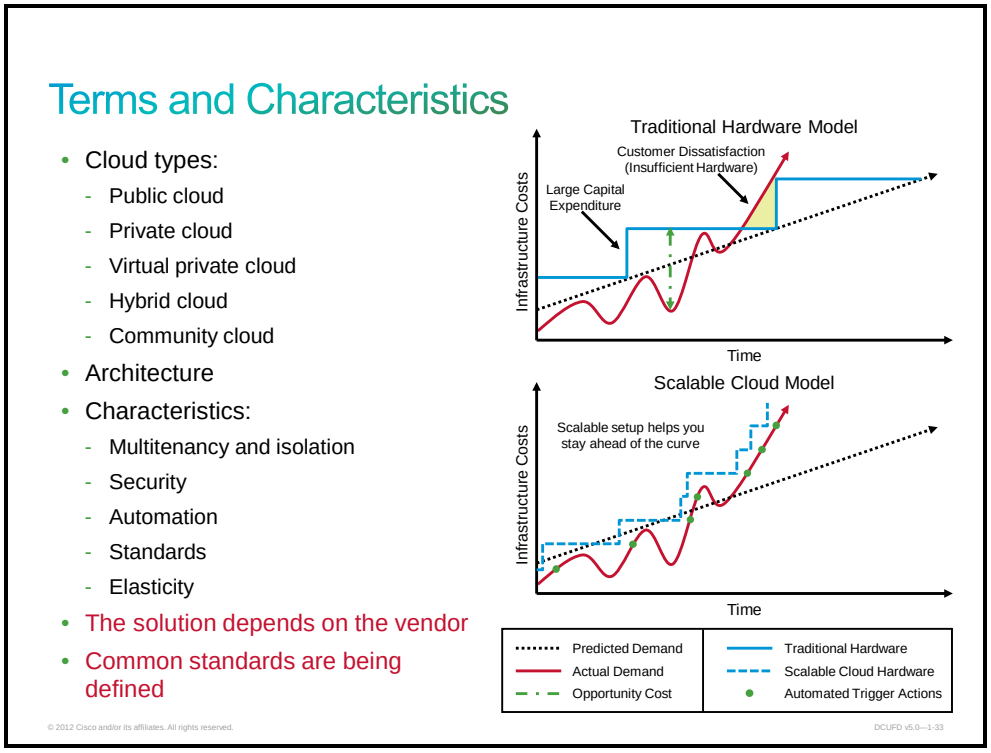
Cloud computing is feasible due to the fundamental principles that are used and applied in modern IT infrastructures and data centers:

- Centralization (that is, consolidation) that aggregates the computing, storage, network, and application resources in central locations—data centers
- Virtualization by which seamless scalability and quick provisioning can be achieved
- Standardization, which makes integration of components from multiple vendors possible
- Automation, which creates time savings and enables user-based self-provisioning of IT services

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net



The terms in the figure are used when talking about cloud computing to define types of cloud solutions. These terms are defined in more detail in the following sections.

The concept of cloud services will evolve toward something that ideally hides complexity and allows control of resources, while providing the automation that removes the complexity.

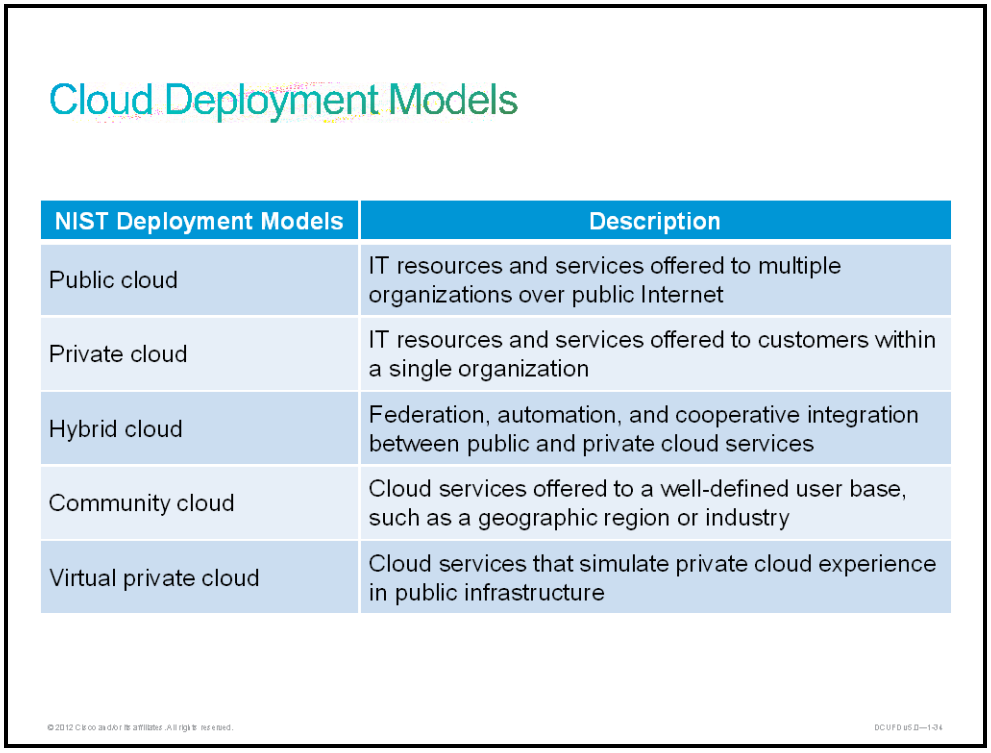
Cloud computing architecture defines the fundamental organization of a system, which is embodied in its components, their relationship to each other, and the principles governing the design and evolution of the system.

This list describes the fundamental characteristics of a cloud solution:

- **Multitenancy and isolation:** This characteristic defines how multiple organizations use and share a common pool of resources (network, computing, storage, and so on) and how the applications and services that are running in the cloud are isolated from each other.
- **Security:** This ability is implicitly included in the previous characteristic. It defines the security policy, mechanisms, and technologies that are used to secure the data and applications of companies that use cloud services. It also defines anything that secures the infrastructure of the cloud itself.
- **Automation:** This feature is an important characteristic that defines how a company that would like to use a cloud-based solution can get the resources and set up its applications and services in the cloud without too much intervention from the cloud service support staff.
- **Standards:** There should be standard interfaces for protocols, packaging, and access to cloud resources so that the companies that are using an external cloud solution (that is, a public cloud or open cloud) can easily move their applications and services between the cloud providers.
- **Elasticity:** Flexibility and elasticity allow users to scale up and down at will—utilizing resources of all kinds (CPU, storage, server capacity, load balancing, and databases).

https://t.me/learningnets

certcollecion.net



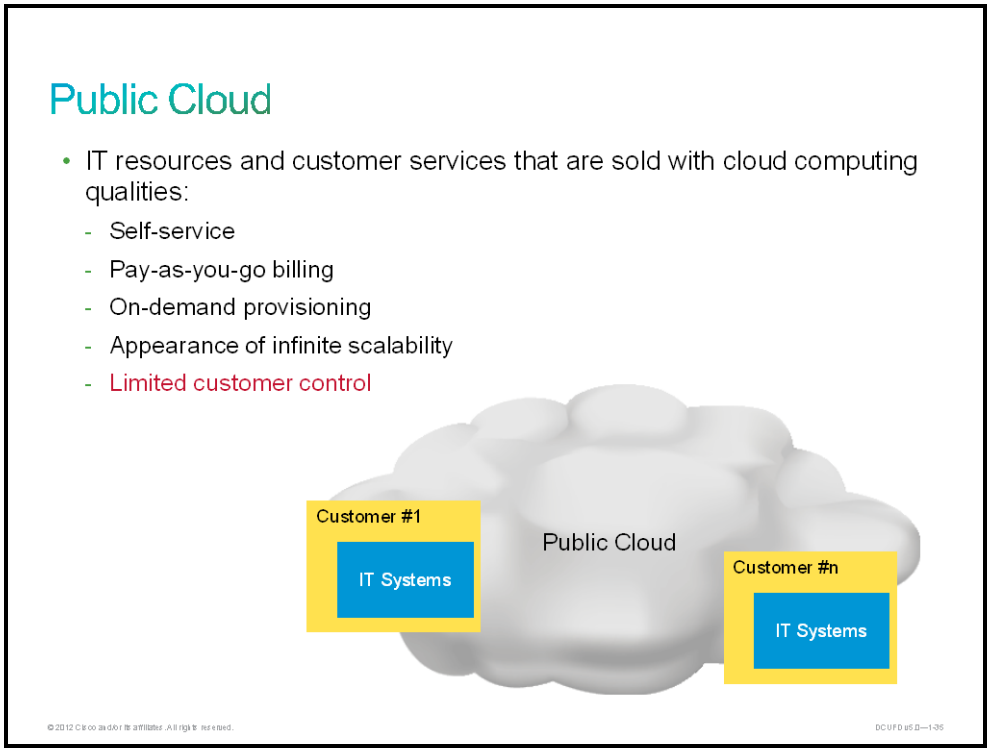
As mentioned, cloud computing is not very well-defined in terms of standards, but there are trends and activities toward defining common cloud computing solutions.

The National Institute of Standards and Technology (NIST), a government agency that is part of the US Department of Commerce, is responsible for establishing standards of all types as needed by industry or government programs.

The NIST cloud definition also defines cloud characteristics that are both essential and common.

https://t.me/learningnets

certcollecion.net



A public cloud can offer IT resources and services that are sold with cloud computing qualities, such as self-service, pay-as-you-go billing, on-demand provisioning, and the appearance of infinite scalability.

Here are some examples of cloud-based offerings:

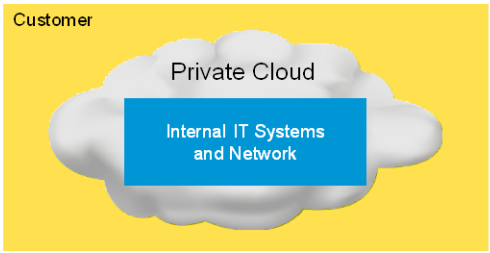
- Google services and applications like Google Docs and Gmail
- Amazon web services and the Amazon Elastic Compute Cloud (EC2)
- Salesforce.com cloud-based customer relationship management (CRM)
- Skype voice services

https://t.me/learningnets

certcollecion.net

Private Cloud

- Enterprise IT infrastructure with cloud computing qualities:
  - Self-service
  - Pay-as-you-go chargeback
  - On-demand provisioning
  - Appearance of infinite scalability
- You can consolidate, virtualize, and automate data center resources with provisioning and cost-metering interfaces to enable self-service IT consumption



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

A private cloud is an enterprise IT infrastructure that is managed with cloud computing qualities, such as self-service, pay-as-you-go charge-back, on-demand provisioning, and the appearance of infinite scalability.

Private clouds will have these characteristics:

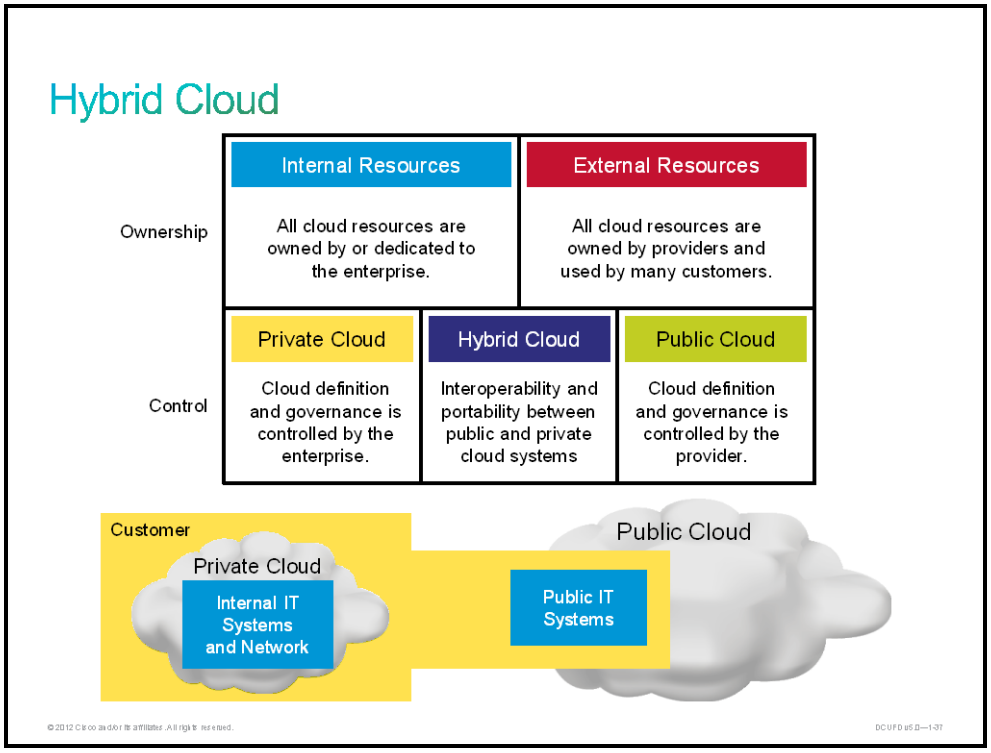
- Consolidated, virtualized, and automated existing data center resources
- Provisioning and cost-metering interfaces to enable self-service IT consumption
- Targeted at one or two noncritical application systems

Once a company has decided that a private cloud service is appropriate, the private cloud will scale by pooling IT resources under a single cloud operating system or management platform. It can support from tens to thousands of applications and services.

This arrangement will enable new architectures to target very large-scale activities.

https://t.me/learningnets

certcollecion.net



Once private and public clouds are well accepted, the tendency to connect them forms the hybrid cloud.

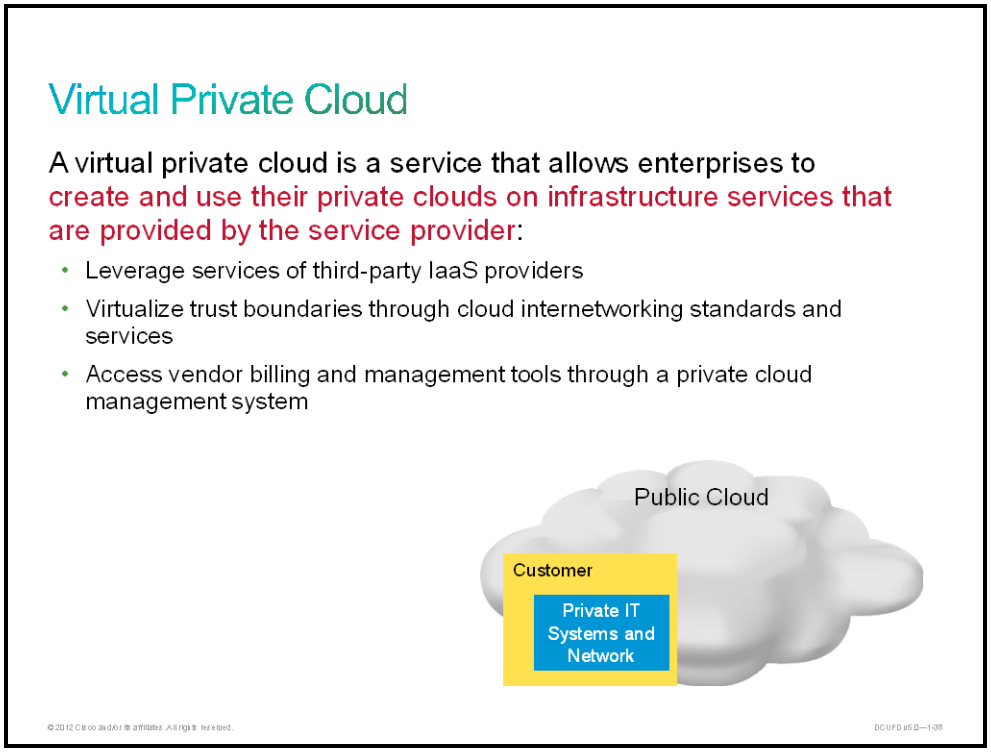
A hybrid cloud links disparate cloud computing infrastructures (that is, an enterprise private cloud with a service provider public cloud) with one another by connecting their individual management infrastructures and allowing the exchange of resources.

The hybrid cloud can enable these options:

- Disparate cloud environments can leverage other cloud-system resources.
- Federation can occur across data center and organization boundaries with cloud internetworking.

https://t.me/learningnets

certcollecion.net



A virtual private cloud is a service offering that allows enterprises to create their private clouds on infrastructure services, such as a public cloud, that are provided by a service provider.

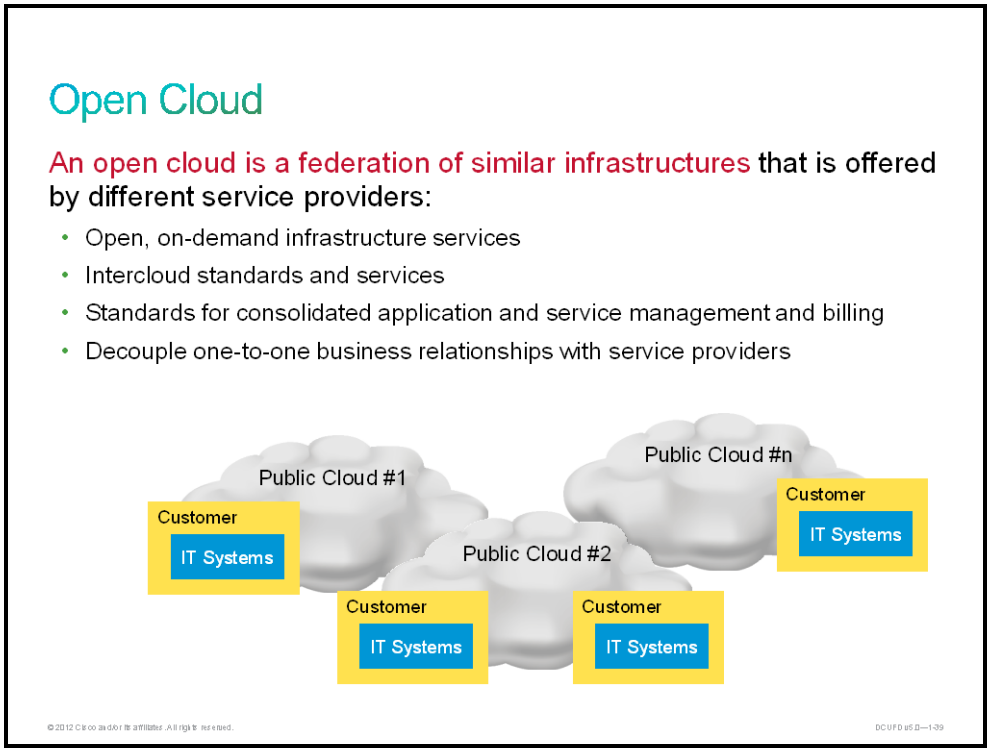
The closed-cloud service provider enables an enterprise to accomplish these activities:

- Leverage services that are offered by third-party Infrastructure as a Service (IaaS) providers
- Virtualize trust boundaries through cloud internetworking standards and services
- Access vendor billing and management tools through a private cloud management system

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net



The largest and most scalable cloud computing system—the open cloud—is a service provider infrastructure that allows a federation with similar infrastructures offered by other providers. Enterprises can choose freely among participants, and service providers can leverage other provider infrastructures to manage exceptional loads on their own offerings.

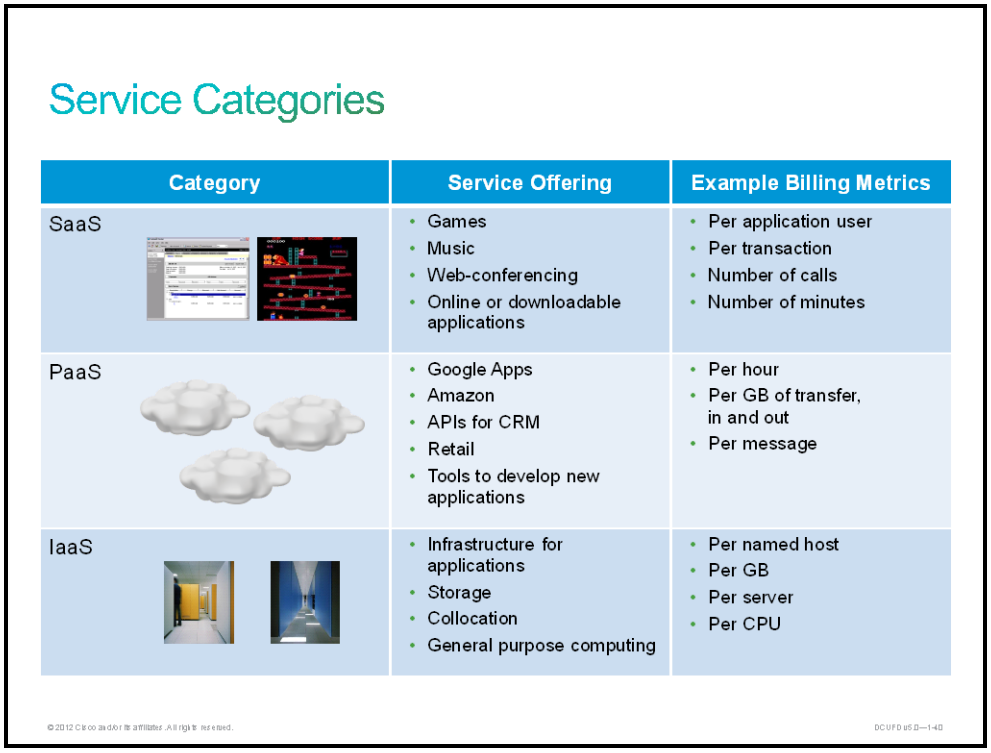
A federation will link disparate cloud computing infrastructures with one another by connecting their individual management infrastructures and allowing the exchange of resources and the aggregation of management and billing streams.

The federation can enable the following options:

- Disparate cloud environments can leverage other cloud-system resources.
- The federation can occur across data center and organization boundaries with cloud internetworking.
- The federation can provide unified metering and billing and “one-stop” self-service provisioning.

https://t.me/learningnets

certcollecion.net



There are three different cloud computing service models:

- Software as a Service (SaaS)
- Platform as a Service (PaaS)
- Infrastructure as a Service (IaaS)

Software as a Service

SaaS is software that is deployed over the Internet or is deployed to run behind a firewall in your LAN or PC. A provider licenses an application to customers as a service on demand, through a subscription or a pay-as-you-go model. SaaS is also called “software on demand.” SaaS vendors develop, host, and operate software for customer use. Rather than installing software onsite, customers can access the application over the Internet. The SaaS vendor may run all or part of the application on its hardware or may download executable code to client machines as needed—disabling the code when the customer contract expires. The software can be licensed for a single user or for a group of users.

Platform as a Service

PaaS is the delivery of a computing platform and solution stack as a service. It facilitates the deployment of applications without the cost and complexity of buying and managing the underlying hardware and software and provisioning hosting capabilities. It provides all of the facilities that are required to support the complete life cycle of building and delivering web applications and services entirely from the Internet. The offerings may include facilities for application design, application development, testing, deployment, and hosting. Offerings may also include application services such as team collaboration, web service integration and marshaling, database integration, security, scalability, storage, persistence, state management, application versioning, application instrumentation, and developer community facilitation. These services may be provisioned as an integrated solution online.

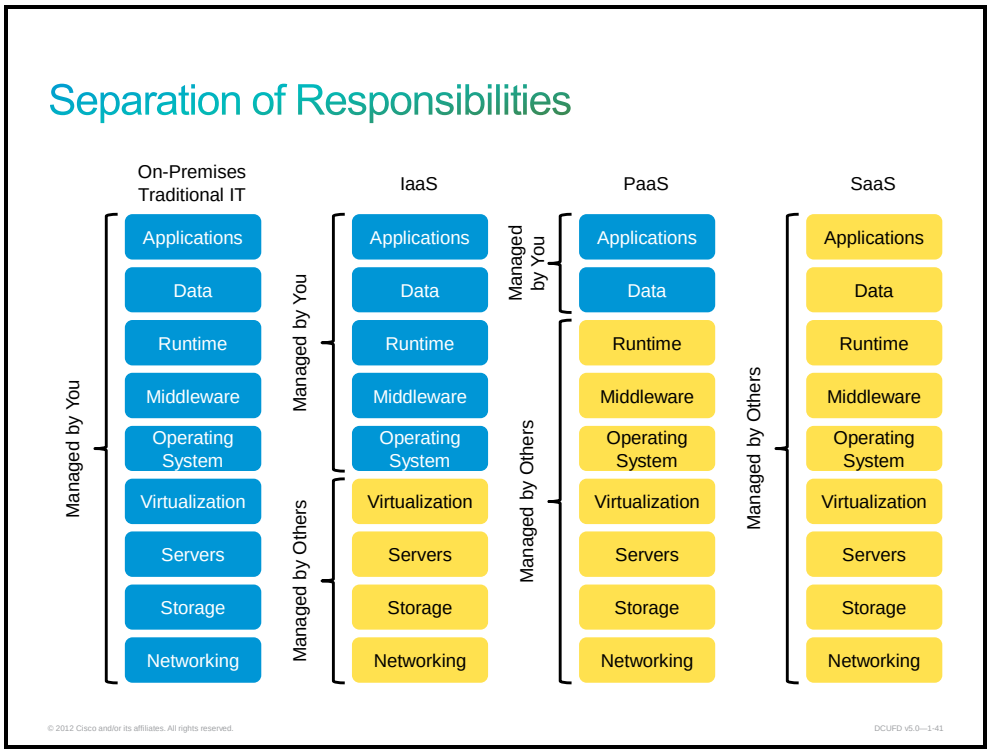
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

Infrastructure as a Service

IaaS, or cloud infrastructure services, can deliver computer infrastructure, typically a platform virtualization environment, as a service. Rather than purchasing servers, software, data center space, or network equipment, clients can buy those resources as a fully outsourced service. The service is typically billed on a utility computing basis and the amount of resources consumed (and therefore the cost) will typically reflect the level of activity. It is an evolution of virtual private server offerings.

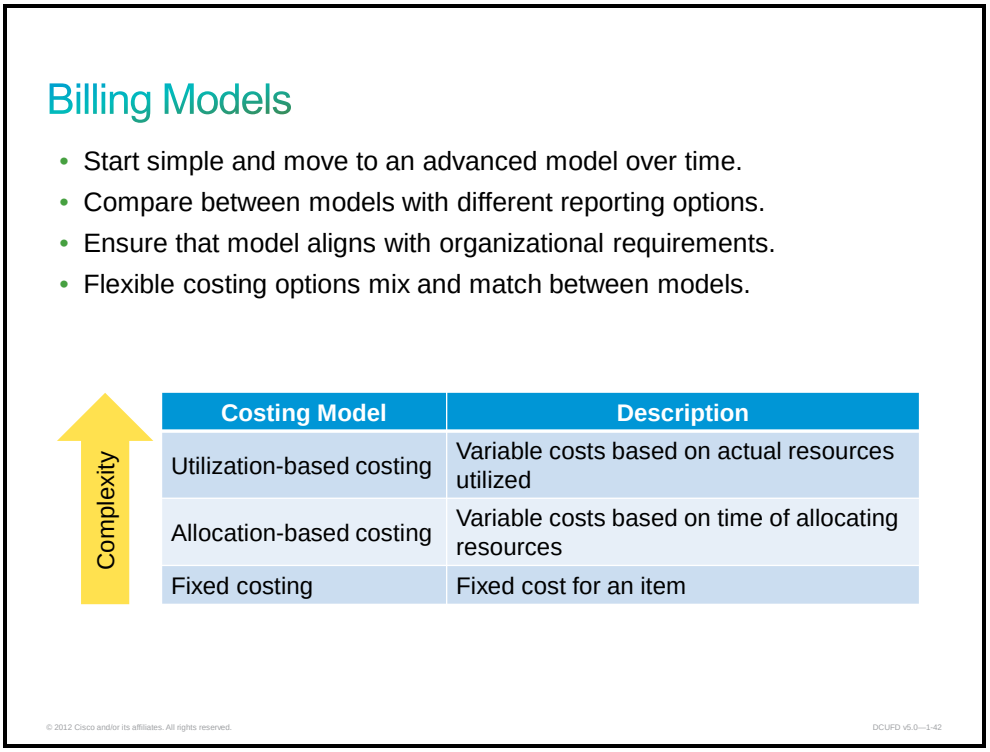


The type of service category also defines the demarcation point for management responsibilities, and ranges from the IaaS model, with shared responsibility between customer and service provider, to the SaaS model, where almost all management responsibilities belong to the service provider.

These demarcation points also mean that the service provider is invested with more trust depending on the model, and must have a better understanding in cases of higher management control. The dependability requirements of the service provider are also higher depending on the model, and are highest in the SaaS model.

https://t.me/learningnets

certcollecion.net



Service providers have an important building block for delivering IT as a service. They need to meter the resources that are offered and used, including broadband network traffic, public IP addresses, and other services such as DHCP, Network Address Translation (NAT), and firewalling.

They need to create a charge-back hierarchy that provides a basis for determining cost structures and delivery of reports.

Multiple cost models provide flexibility in measuring costs.

The table describes three basic cost models that are typically used.

Basic Cost Models

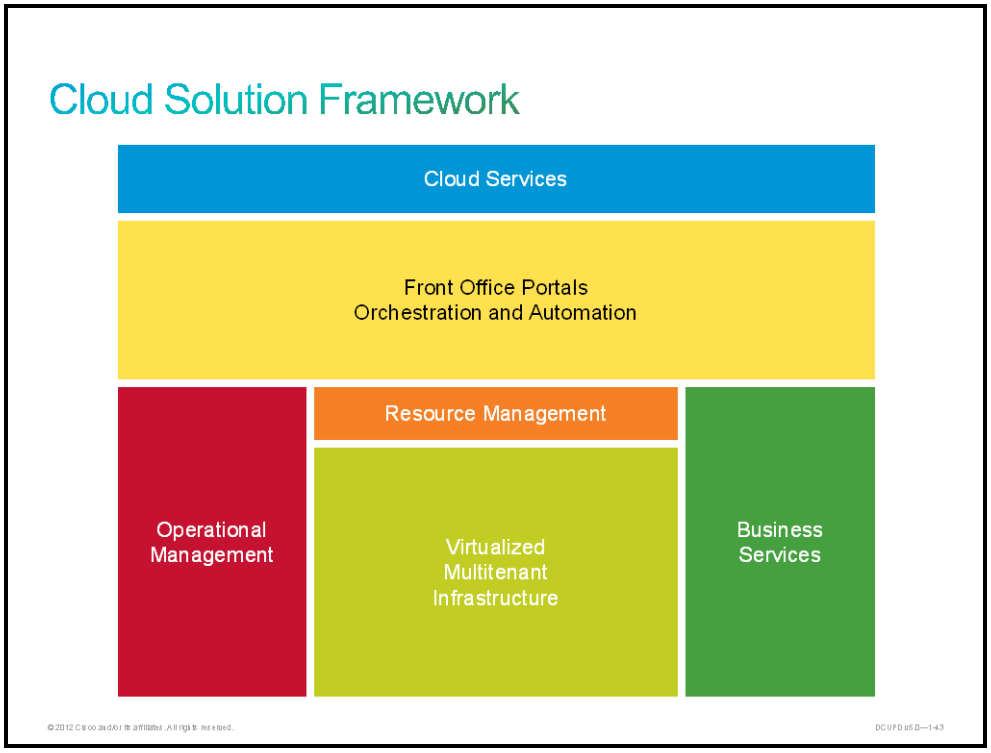
Cost Model	Description
Fixed cost	Specific per-virtual machine instance costs, such as floor space, power, cooling, software, or administrative overhead
Allocation-based costing	Variable costs per virtual machine based on allocated resources, such as the amount of memory, CPU, or storage allocated or reserved for the virtual machine
Utilization-based costing	Variable costs per virtual machine based on actual resources used, including average memory, disk and CPU usage, network I/O, and disk I/O

Cost models can be combined in a cost template, making it easy to start with a simple charge-back model that is aligned with organizational requirements.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net



The transition to cloud-based computing is an evolution that is characterized by different architectures and steps. The focus should be on helping customers to move to the right side of the figure and build their own dynamic data centers.

The evolution from the current infrastructure to cloud-based computing will proceed in phases. Most companies currently operate in application silos or they are already using virtualization that is deployed in zones like infrastructure applications, per department, and so on.

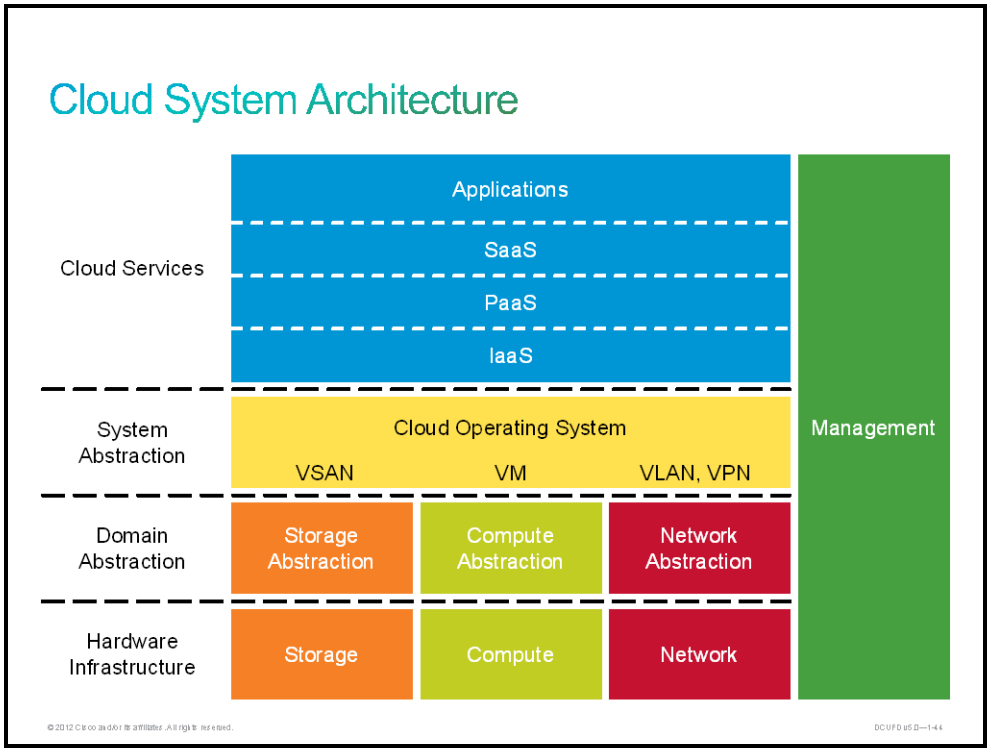
Next, a shared infrastructure that forms an internal cloud will emerge from the virtualization zones.

Finally, the use of external cloud services will become more popular. Service providers in the cloud market are already on the far right of the scheme.

Customers will have more than one architecture (for example, four) and will likely be moving to a more cloud-oriented IT structure.

https://t.me/learningnets

certcollecion.net



The system architecture of a cloud is like the architecture of a data center. The cloud has all the components and aspects that a typical data center has.

Cisco Virtualized Multi-Tenant Data Center

The Cisco Virtualized Multi-Tenant Data Center (VMDC) provides design and implementation guidance for enterprises that are planning to deploy private cloud services and for service providers that are building virtual private and public cloud services.

Cisco VMDC is a validated architecture that delivers a highly available, secure, flexible, and efficient data center infrastructure. It provides these benefits:

- **Reduced time to deployment:** It provides a fully tested and validated architecture that accelerates technology adoption and rapid deployment.
- **Reduced risk:** It enables enterprises and service providers to deploy new architectures and technologies with confidence.
- **Increased flexibility:** It provides rapid, on-demand, workload deployments in a multitenant environment due to a comprehensive automation framework with portal-based resource provisioning and management capabilities.
- **Improved operational efficiency:** It integrates automation with a multitenant resource pool (computing, network, and storage), improves asset use, reduces operational overhead, and mitigates operational configuration errors.

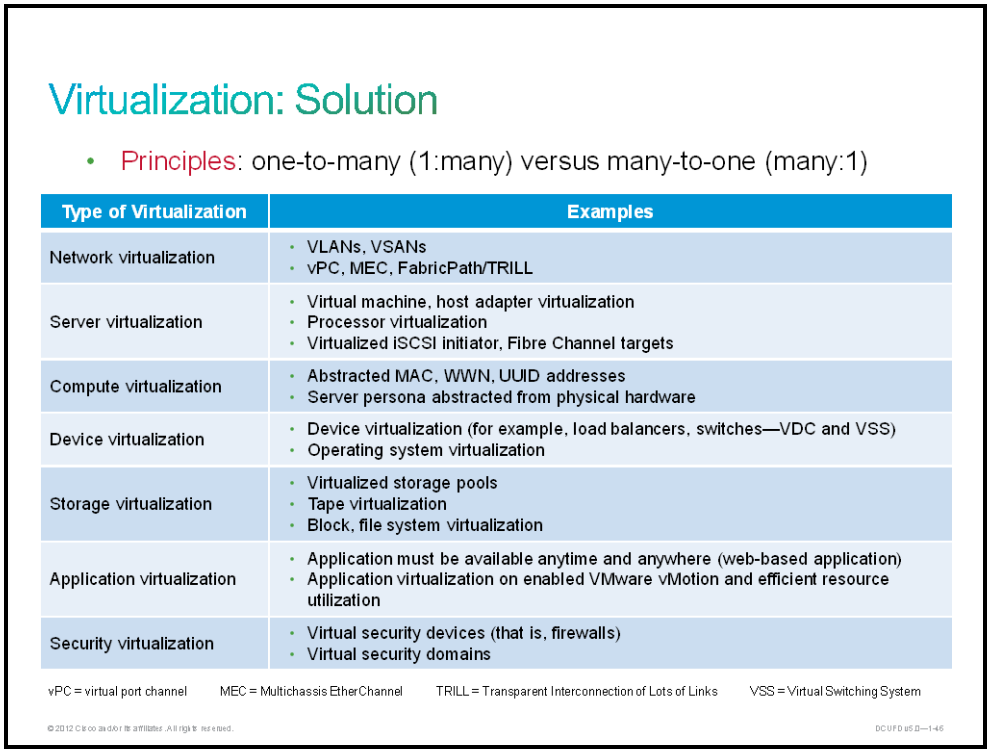
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

Data Center Virtualization

This topic describes how to recognize the importance of virtualization technologies and solutions for data center evolution.



Virtualization delivers tremendous flexibility to build and design data center solutions. Considering the diverse networking needs of different enterprises might require separation of a single user group or a separation of data center resources from the rest of the network for certain reasons. Separation tasks become complex when it is not possible to confine specific users or resources to specific areas in the network. When separation occurs, the physical positioning will no longer address the problem.

Network Virtualization

Network virtualization can address the problem of separation. Network virtualization also provides other types of benefits such as increasing network availability, better security, consolidation of multiple networks, segmentation of networks, and increased network availability. Examples of network virtualization are VLANs and virtual SANs (VSANs) in Fibre Channel SANs. A VLAN virtualizes Layer 2 segments, making them independent of the physical topology. This virtualization presents the ability to connect two servers to the same physical switch, though they participate in different logical broadcast domains (VLANs). A similar concept represents a VSAN in Fibre Channel SANs.

Server Virtualization

Server virtualization enables physical consolidation of servers on the common physical infrastructure. Deployment of another virtual server is easy because there is no need to buy a new adapter and a new server. For a virtual server to be enabled, software just needs to be activated and configured properly. Therefore, server virtualization simplifies server deployment, reduces the cost of management, and increases server utilization. VMware and Microsoft are examples of companies that support server virtualization technologies.

https://t.me/learningnets

certcollecion.net

Device Virtualization

Cisco Nexus 7000 and Catalyst 6500 switches support device virtualization or Cisco Nexus Operating System (Cisco NX-OS) virtualization. A virtual device context (VDC) represents the ability of the switch to enable multiple virtual and independent switches on the common physical switch to participate in data center networks. This feature provides various benefits to the application services, such as higher service availability, fault isolation, separation of logical networking infrastructure that is based on traffic service types, and flexible and scalable data center design.

Storage Virtualization

Storage virtualization is the ability to pool storage on diverse and independent devices into a single view. Features such as copy services, data migration, and multiprotocol and multivendor integration can benefit from storage virtualization.

Application Virtualization

The web-based application must be available anytime and anywhere. It should be able to use idle remote server CPU resources, which implies an extended Layer 2 domain. Application virtualization enables VMware VMotion and efficient resource utilization.

Computing Virtualization

The computing virtualization is a paradigm used in server deployments. The servers have become stateless, with service profiles defining operational properties of the servers, such as: MAC addresses, universally unique identifiers (UUIDs), world wide names (WWNs), and so on. By applying the service profile to another stateless hardware instance, you can move the workload to another server for added capacity or recovery.

Common Goals

There are some common goals across virtualization techniques:

- **Affecting utilization and reducing overprovisioning:** The main goal is to reduce operating costs for maintaining equipment that is not really needed, or is not fully utilized, by reducing the amount of equipment and directing the utilization figures higher. Overprovisioning has been used to provide a safety margin, but with virtualization, a lower overprovisioning percentage can be used because systems are more flexible.
- **Isolation:** Security must be effective enough to prevent any undesired access across the virtual entities that share a common physical infrastructure. Performance (quality of service [QoS] and SLA) must be provided at the desired level independently for each virtual entity. Faults must be contained.
- **Management:** Flexibly managing a virtual resource requires no hardware change in many cases. Individual administration for each virtual entity can be deployed using role-based access control (RBAC).

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

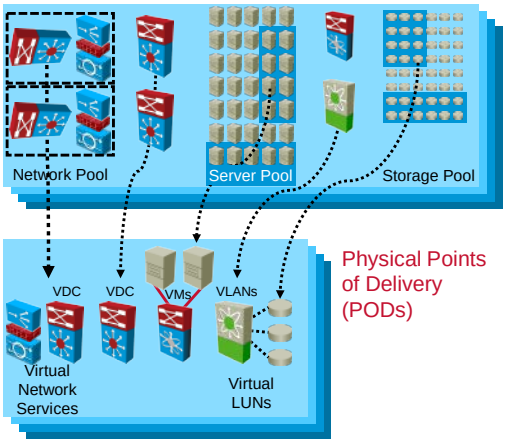
certcollecion.net

Virtualization: Service and Device

Logical and Physical Data Center View

Virtualized data center POD:

- Logical instantiation of entire data center network infrastructure using VDC, VLANs, VSAN, and virtual services
- Fault isolation, high reliability
- Efficient resource pool utilization
- Centralized management
- Scalability



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

Virtualizing data center network services has changed the logical and physical data center network topology view.

Service virtualization enables higher service density by eliminating the need to deploy separate appliances for each application. There are a number of benefits of higher service density:

- Less power consumption
- Less rack space
- Reduced ports and cabling
- Simplified operational management
- Lower maintenance costs

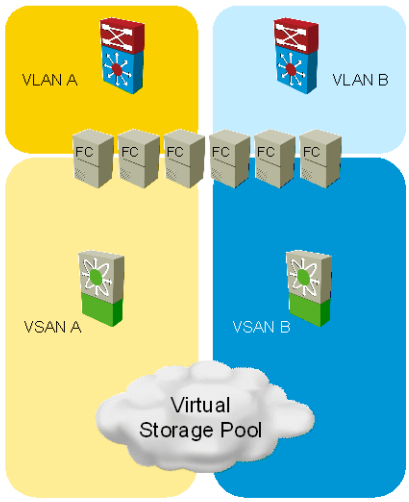
The figure shows how virtual services can be created from the physical infrastructure, using features such as VDC, VLANs, and VSANs. Virtual network services include virtual firewalls with Cisco adaptive security appliances and service modules or Firewall Services Module (FWSM), and virtual server load-balancing contexts with the Cisco Application Control Engine (ACE) and Cisco Intrusion Detection System (IDS).

https://t.me/learningnets

certcollecion.net

Virtualization: Storage and SAN

- Storage device abstracts hosts from physical disks
- SANs allow storage sharing and consolidation across multiple servers:
  - Leads to SAN proliferation (SAN islands) and poor utilization
  - Virtual SANs (VSANs) allow further consolidation of SAN islands:
    - Increased utilization
    - Easier to manage
- Storage device virtualization coupled with VSANs enables dynamic storage allocation.



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

Data center storage virtualization starts with Cisco VSAN technology. Traditionally, SAN islands have been used within the data center to separate traffic on different physical infrastructures, providing security and separation from both a management and traffic perspective. To provide virtualization facilities, VSANs are used within the data center SAN environment to consolidate SAN islands onto one physical infrastructure while maintaining the separation from management and traffic perspectives.

Storage virtualization also involves virtualizing the storage devices themselves. Coupled with VSANs, storage device virtualization enables dynamic allocation of storage. Taking a similar approach to the integration of network services directly into data center switching platforms, the Cisco MDS 9000 NX-OS platform supports third-party storage virtualization applications on an MDS 9000 services module, reducing operational costs by consolidating management processes.

https://t.me/learningnets

certcollecion.net

Summary

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

Summary

- The main data center solution components are the network, storage, and computing components. These allow several services to run on top of the physical infrastructure.
- Various service availability metrics are used in data center environments. High availability, serviceability, and fault tolerance are the most important metrics.
- Data centers have various challenges, including business-oriented, organizational, operational, facility-related, power, and cooling.
- The cloud computing paradigm is an approach where IT resources and services are abstracted from the underlying infrastructure. Clouds are particularly interesting because of various available pay-per-use models.
- Consolidation and virtualization technologies are essential to modern data centers and provide more energy-efficient operation and simplified management.

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

https://t.me/learningnets

certcollecion.net

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

Lesson 2

Identifying the Cisco Data Center Solution

Overview

In this lesson, you will identify Cisco products that are part of the Cisco Data Center solution. The networking equipment is one of the most important components, and comprises Cisco Nexus switches and selected Cisco Catalyst switches. Among the most innovative components is the Cisco Unified Computing System (UCS). Other components are storage, security, and application delivery products. Using all of these components, you can create a flexible, reliable, and highly available data center solution that can fulfill the needs of any data center.

Objectives

Upon completing this lesson, you will be able to provide a high-level overview of the Cisco Data Center solution architectural framework and components within the solution. This ability includes being able to meet these objectives:

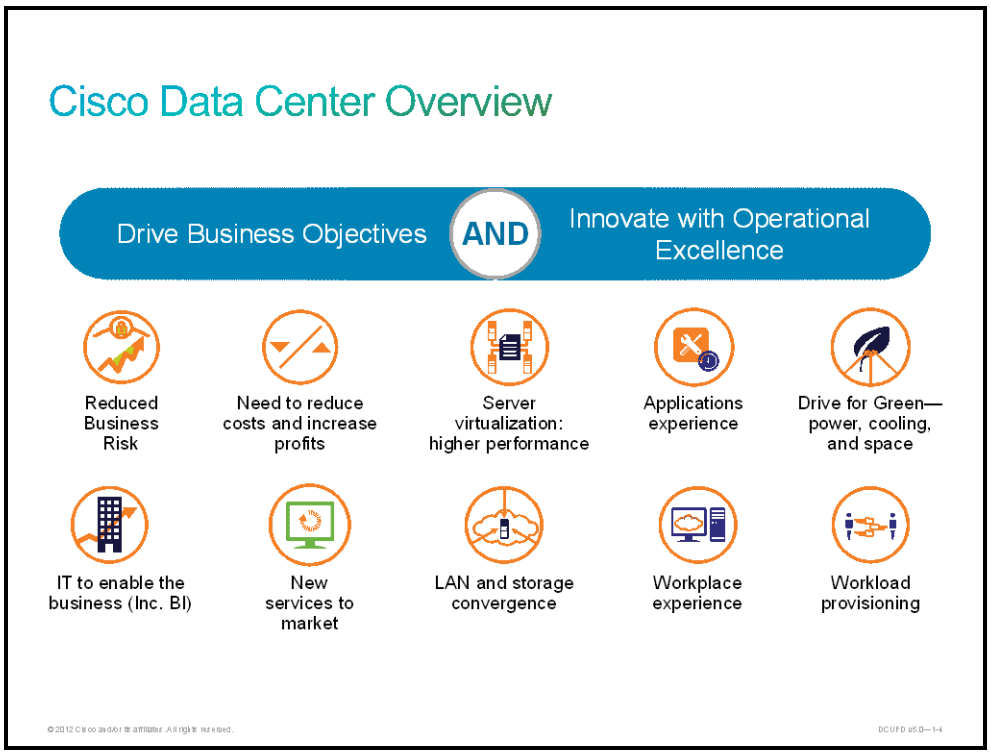
- Evaluate the Cisco Data Center architectural framework
- Evaluate the Cisco Data Center architectural framework network component
- Evaluate the Cisco Data Center architectural framework storage component

https://t.me/learningnets

certcollecion.net

Cisco Data Center Architecture Overview

This topic describes the Cisco Data Center architectural framework.

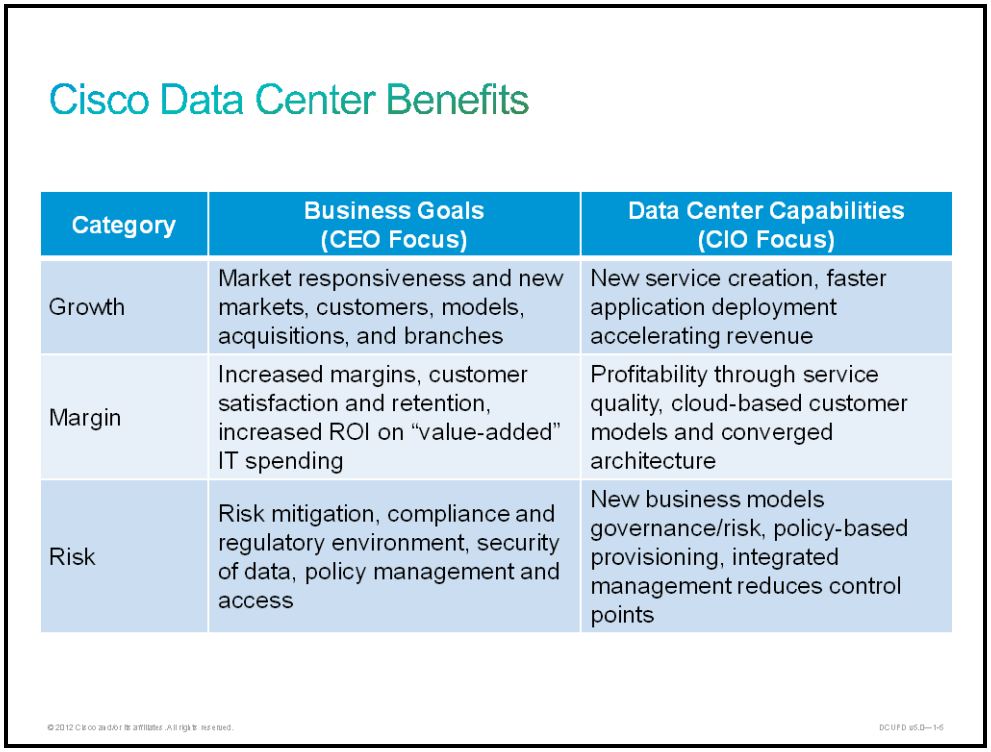


Cisco Data Center represents a fundamental shift in the role of IT into becoming a driver of business innovation. Businesses can create services faster, become more agile, and take advantage of new revenue streams and business opportunities. Cisco Data Center increases efficiency and profitability by reducing capital expenditures, operating expenses, and complexity. It also transforms how a business approaches its market and how IT supports and aligns with the business, to help enable new and innovative business models.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net



Businesses today are under three pressure points. These pressure points are business growth, margin, and risk.

For growth, the business needs to be able to respond to the market quickly and lead market reach into new geographies and branch openings. Businesses also need to gain better insight with market responsiveness (new services), maintain customer relationship management (CRM) for customer satisfaction and retention, and encourage customer expansion.

The data center capabilities can help influence growth and business. By enabling the ability to affect new service creation and faster application deployment through service profiling and rapid provision of resource pools, the business can enable service creation without spending on infrastructure, and provide increased service level agreements.

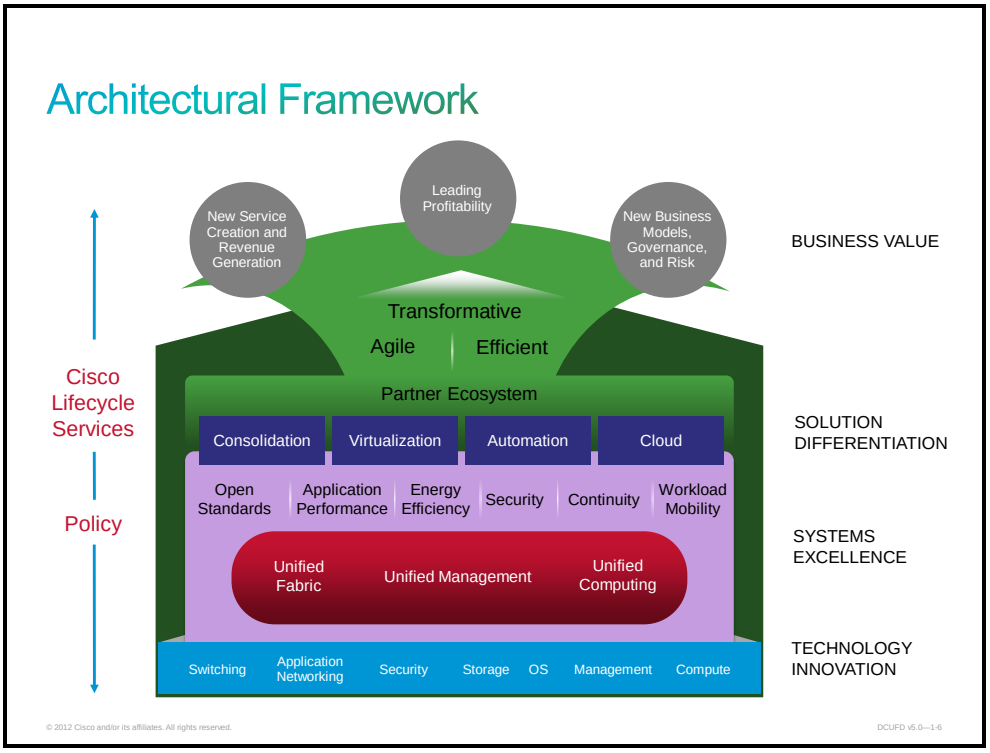
Cost cutting, margin, and efficiencies are all critical elements for businesses today in the current economic climate. When a business maintains focus on cutting costs, increasing margins through customer retention and satisfaction, and product brand awareness and loyalty, the result is a higher return on investment (ROI). The data center works toward a service-robust, converged architecture to reduce costs. At the same time, the data center enhances application experience and increases productivity through a scalable platform for collaboration tools.

The element of risk in a business must be minimized. While the business focuses on governing and monitoring changing compliance rules and a regulatory environment, it is also highly concerned with security of data, policy management, and access. The data center must ensure a consistent policy across services so there is no compromise on quality of service versus quantity of service. Furthermore, the business needs the flexibility to implement and try new services quickly, while being sure they can retract them quickly if they prove unsuccessful, all with limited impact.

These areas show how the IT environment, and the data center in particular, can have a major impact on business.

https://t.me/learningnets

certcollecion.net



Cisco Data Center is an architectural framework that connects technology innovation with business innovation. It is the foundation for a model of the dynamic networked organization and can enable the following important aspects:

- Quick and efficient innovation
- Control of data center architecture
- Freedom to choose technologies

The Cisco Data Center architectural framework is delivered as a portfolio of technologies and systems that can be adapted to meet organizational needs. You can adopt the framework in an incremental and granular fashion to control when and how you implement data center innovations. This framework allows you to easily evolve and adapt the data center to keep pace with changing organizational needs.

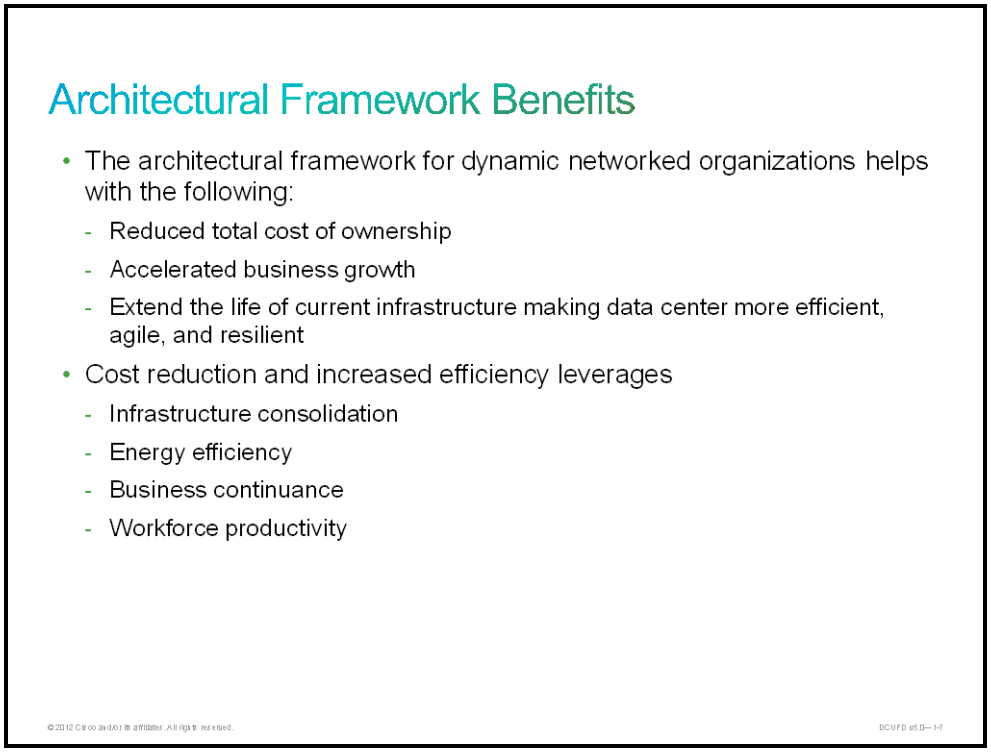
The Cisco approach to the data center is to provide an open and standards-based architecture. System-level benefits, such as performance, energy efficiency, and resiliency, are addressed, along with workload mobility and security. Cisco offers tested, preintegrated, and validated designs that provide businesses with a faster deployment model and a quicker time to market.

The components of the Cisco Data Center architecture are categorized into four areas: technology innovations, systems excellence, solution differentiation, and business advantages.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net



The Cisco Data Center framework is an architecture for dynamic networked organizations. The framework allows organizations to create services faster, improve profitability, and reduce the risk of implementing new business models. It can provide the following benefits:

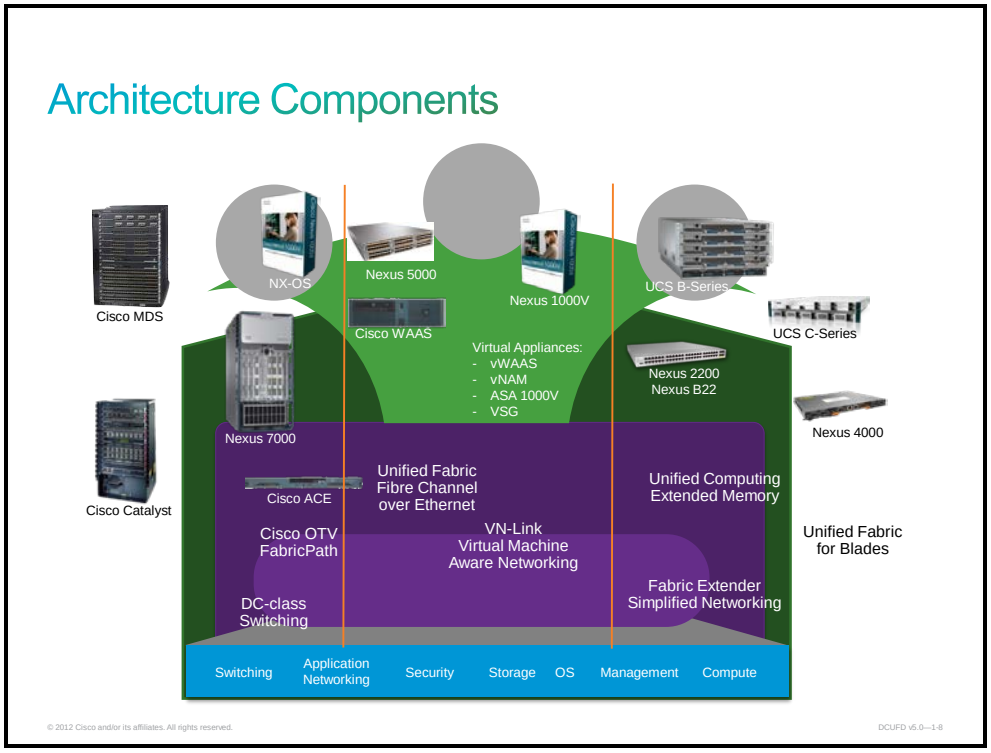
- Business value
- Flexibility and choice with an open ecosystem
- Innovative data center services

Cisco Data Center is a portfolio of practical solutions that are designed to meet IT and business needs and can help accomplish these goals:

- Reduce total cost of ownership (TCO)
- Accelerate business growth
- Extend the life of the current infrastructure by making your data center more efficient, agile, and resilient

https://t.me/learningnets

certcollecion.net



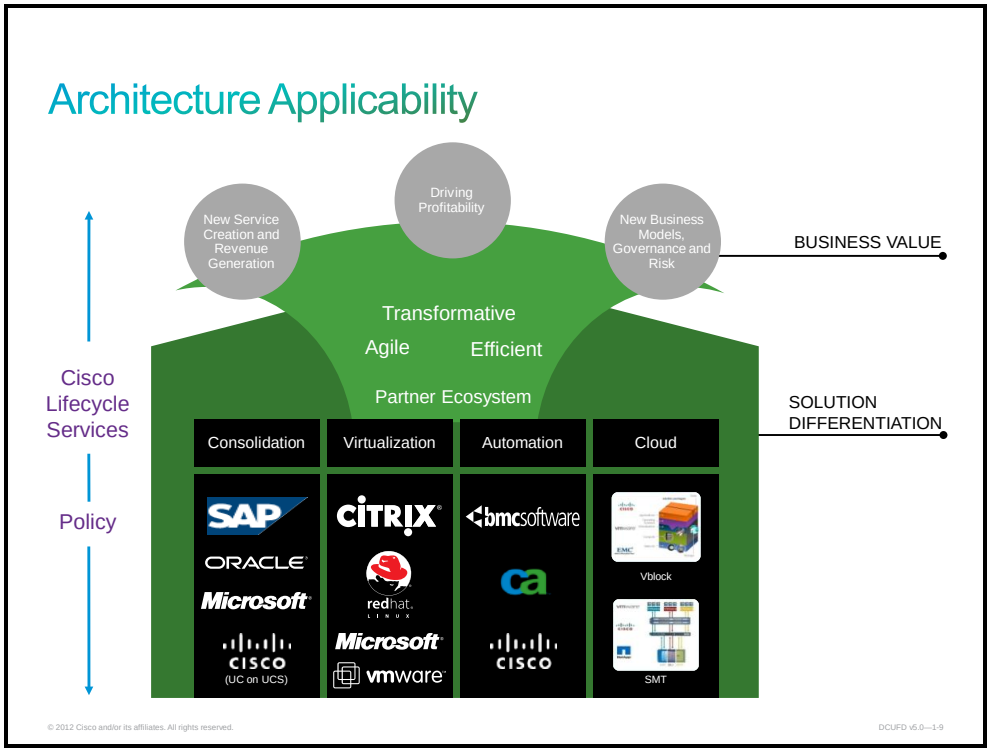
The framework brings together multiple technologies:

- **Data center switching:** These next-generation virtualized data centers need a network infrastructure that delivers the complete potential of technologies such as server virtualization and unified fabric.
- **Storage networking solutions:** SANs are central to the Cisco Data Center architecture. They provide a networking platform that helps IT departments to achieve a lower TCO, enhanced resiliency, and greater agility through Cisco Data Center storage solutions.
- **Cisco Application Networking Solutions:** You can improve application performance, availability, and security, while simplifying your data center and branch infrastructure. Cisco Application Networking Services (ANS) solutions can help you lower your TCO and improve IT flexibility.
- **Data center security:** Cisco Data Center security solutions enable you to create a trusted data center infrastructure that is based on a systems approach and uses industry-leading security solutions.
- **Cisco UCS:** You can improve IT responsiveness to rapidly changing business demands with the Cisco UCS. This next-generation data center platform accelerates the delivery of new services simply, reliably, and securely through end-to-end provisioning and migration support.
- **Cisco Virtualization Experience Infrastructure (VXI):** Cisco VXI can deliver a superior collaboration and rich-media user experience with a best-in-class ROI in a fully integrated, open, and validated desktop virtualization solution.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net



The architectural framework encompasses the network, storage, application services, security, and computing equipment.

The architectural framework is open and can integrate with other vendor solutions and products, such as VMware vSphere, VMware View, Microsoft Hyper-V, Citrix XenServer, and Citrix XenDesktop.

Starting from the top down, virtual machines (VMs) are one of the most important components of the framework. VMs are entities that run an application within the client operating system, which is further virtualized and running on common hardware.

The logical server personality is defined by using management software, and it defines the properties of the server: the amount of memory, percentage of total computing power, number of network interface cards, boot image, and so on.

The network hardware for consolidated connectivity serves as one of the most important technologies for fabric unification.

VLANs and virtual storage area networks (VSANs) provide for virtualized LAN and SAN connectivity, separating physical networks and equipment into virtual entities.

On the lowest layer of the framework is the virtualized hardware. Storage devices can be virtualized into storage pools, and network devices are virtualized by using device contexts.

The Cisco Data Center switching portfolio is built on the following common principles:

- **Design flexibility:** Modular, rack, and integrated blade switches are optimized for both Gigabit Ethernet and 10 Gigabit Ethernet environments.
- **Industry-leading switching capabilities:** Layer 2 and Layer 3 functions can build stable, secure, and scalable data centers.
- **Investment protection:** The adaptability of the Cisco Nexus and Catalyst families simplifies capacity and capability upgrades.
- **Operational consistency:** A consistent interface and consistent tools simplify management, operations, and problem resolution.

https://t.me/learningnets

certcollecion.net

- **Virtualization:** Cisco Data Center switches provide VM mobility support, management, and operations tools for a virtualized environment. In addition, the Cisco Catalyst 6500 Switch offers Virtual Switching System (VSS) capabilities, and the Cisco Nexus 7000 Switch offers hypervisor-like virtual switch capabilities in the form of virtual device contexts (VDCs).

The Cisco storage solution provides the following:

- **Multiprotocol storage networking:** By providing flexible options for Fibre Channel, fiber connectivity (FICON), Fibre Channel over Ethernet (FCoE), Internet Small Computer Systems Interface (iSCSI), and Fibre Channel over IP (FCIP), any business risk is reduced.
- **A unified operating system and management tools:** Operational simplicity, simple interoperability, and feature consistency can reduce operating expenses.
- **Enterprise-class storage connectivity:** Significantly larger virtualized workloads can be supported, providing availability, scalability, and performance.
- **Services-oriented SANs:** The “any network service to any device” model can be extended regardless of protocol, speed, vendor, or location.

Cisco ANS provides the following attributes:

- **Application intelligence:** You can take control of applications and the user experience.
- **Cisco Unified Network Services:** You can connect any person to any resource with any device.
- **Integrated security:** There is built-in protection for access, identity, and data.
- **Nonstop communications:** Users can stay connected with a resilient infrastructure that enables business continuity.
- **Virtualization:** This feature allows simplification of the network and the ability to maximize resource utilization.
- **Operational manageability:** You can deploy services faster and automate routine tasks.

The Cisco Data Center security solutions enable businesses to create a trusted data center infrastructure that is based on a systems approach and industry-leading security solutions. These solutions enable the rapid deployment of data center technologies without compromising the ability to identify and respond to evolving threats, protect critical assets, and enforce business policies.

The Cisco UCS provides the following benefits:

- Streamlining of data center resources to reduce TCO
- Scaling of service delivery to increase business agility
- Reducing the number of devices that require setup, management, power, cooling, and cabling

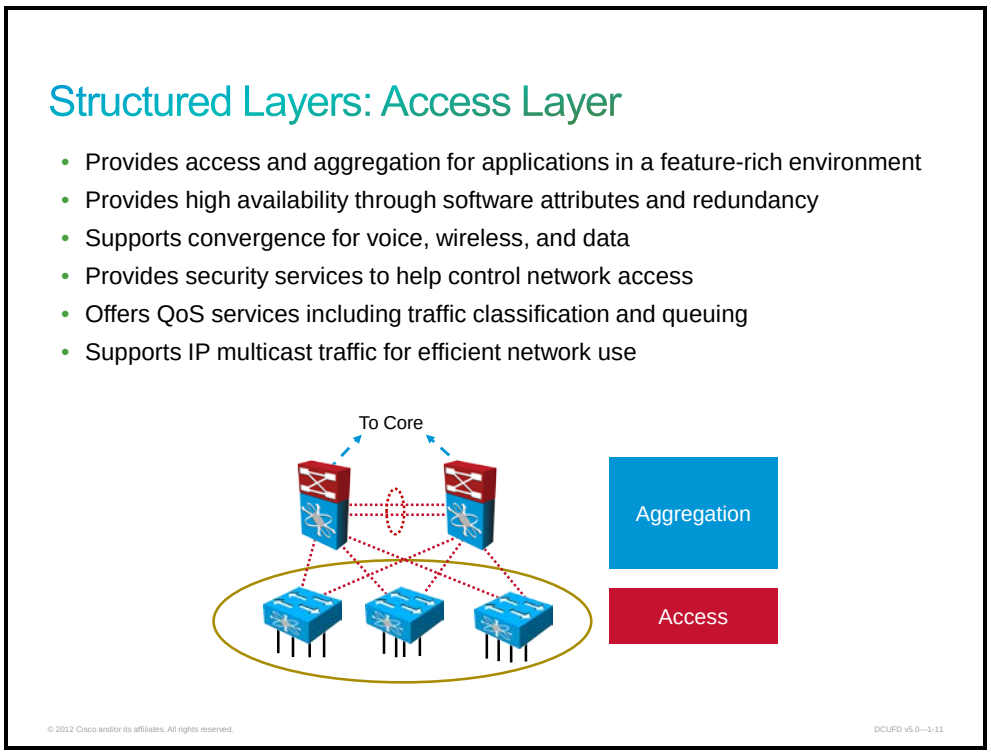
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

Cisco Data Center Architecture Network

This topic describes the Cisco Data Center architectural framework network component.



The architectural components of the infrastructure are the access layer, the aggregation layer, and the core layer. The principal advantages of this model are its hierarchical structure and its modularity. A hierarchical design avoids the need for a fully meshed network in which all network nodes are interconnected. Modules in a layer can be put into service and taken out of service without affecting the rest of the network. This ability facilitates troubleshooting, problem isolation, and network management.

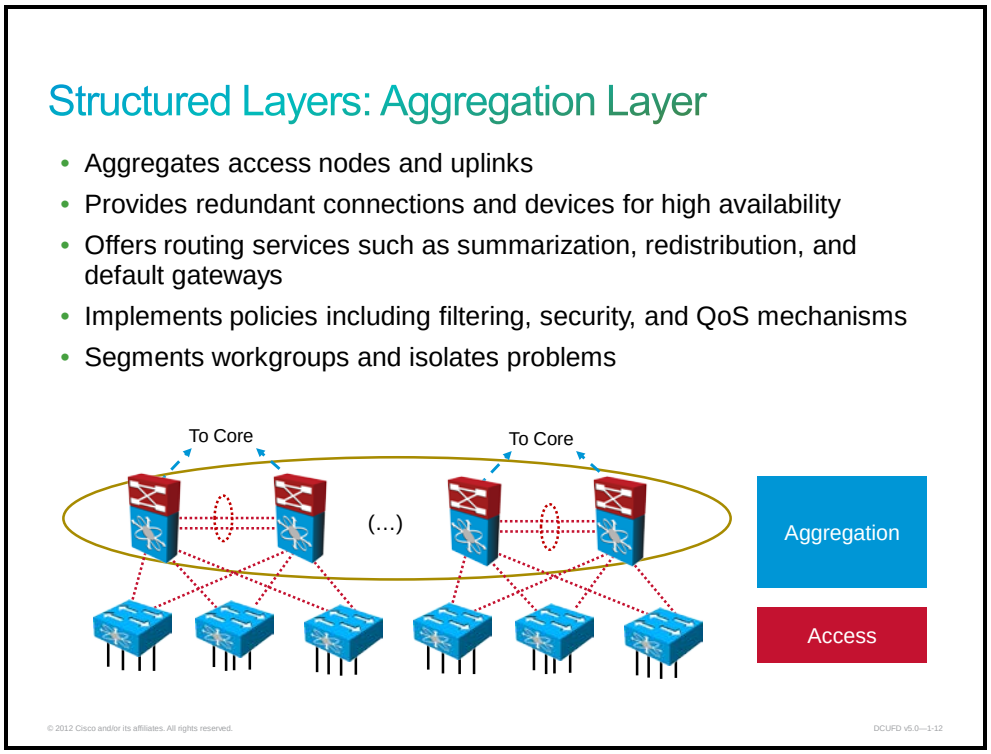
The access layer aggregates end users and provides uplinks to the aggregation layer. The access layer can be an environment with many features:

- **High availability:** The access layer is supported by many hardware and software attributes. This layer offers system-level redundancy by using redundant supervisor engines and redundant power supplies for crucial application groups. The layer also offers default gateway redundancy by using dual connections from access switches to redundant aggregation layer switches that use a First Hop Redundancy Protocol (FHRP), such as the Hot Standby Router Protocol (HSRP).
- **Convergence:** The access layer supports inline Power over Ethernet (PoE) for IP telephony and wireless access points. This support allows customers to converge voice onto their data networks and provides roaming wireless LAN (WLAN) access for users.
- **Security:** The access layer provides services for additional security against unauthorized access to the network. This security is provided by using tools such as IEEE 802.1X, port security, DHCP snooping, Dynamic Address Resolution Protocol (ARP) Inspection (DAI), and IP Source Guard.
- **Quality of service (QoS):** The access layer allows prioritization of mission-critical network traffic by using traffic classification and queuing as close to the ingress of the network as possible. The layer supports the QoS trust boundary.

https://t.me/learningnets

certcollecion.net

- **IP multicast:** The access layer supports efficient network and bandwidth management by using software features such as Internet Group Management Protocol (IGMP) snooping.



Availability, load balancing, QoS, and provisioning are the important considerations at this layer. High availability is typically provided through dual paths from the aggregation layer to the core and from the access layer to the aggregation layer. Layer 3 equal-cost load sharing allows both uplinks from the aggregation to the core layer to be used.

The aggregation layer is the layer in which routing and packet manipulation is performed and can be a routing boundary between the access and core layers. The aggregation layer represents a redistribution point between routing domains or the demarcation between static and dynamic routing protocols. This layer performs tasks such as controlled-routing decision making and filtering to implement policy-based connectivity and QoS. To further improve routing protocol performance, the aggregation layer summarizes routes from the access layer. For some networks, the aggregation layer offers a default route to access layer routers and runs dynamic routing protocols when communicating with core routers.

The aggregation layer uses a combination of Layer 2 and multilayer switching to segment workgroups and to isolate network problems, so that they do not affect the core layer. This layer is commonly used to terminate VLANs from access layer switches. The aggregation layer also connects network services to the access layer and implements policies regarding QoS, security, traffic loading, and routing. In addition, this layer provides default gateway redundancy by using an FHRP such as HSRP, Gateway Load Balancing Protocol (GLBP), or Virtual Router Redundancy Protocol (VRRP). Default gateway redundancy allows for the failure or removal of one of the aggregation nodes without affecting endpoint connectivity to the default gateway.

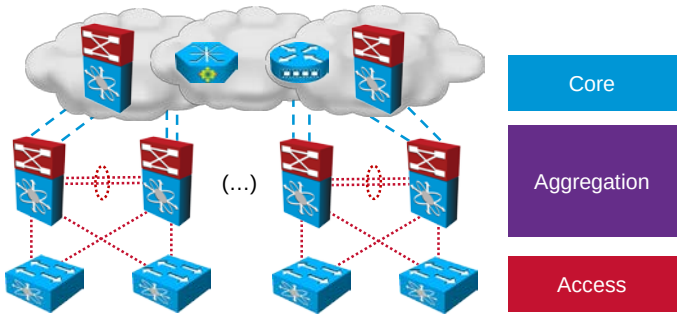
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

Structured Layers: Core Layer

- The core layer is a high-speed backbone and aggregation point for the enterprise.
- It provides reliability through redundancy and fast convergence.
- Separate core layer helps in scalability during future growth.



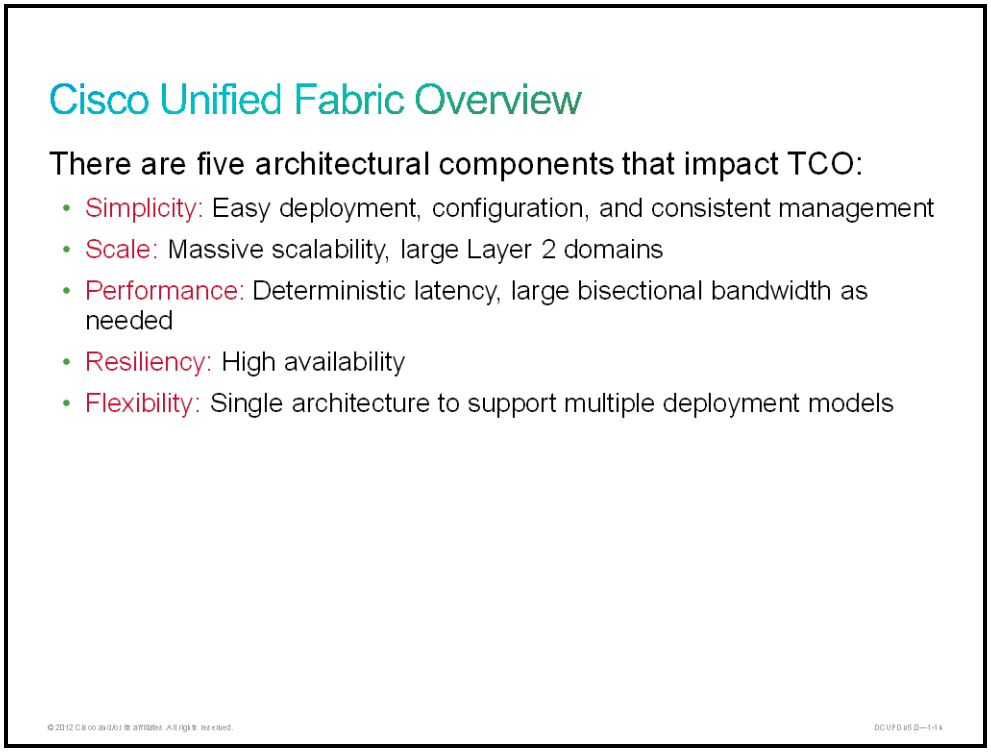
© 2012 Cisco and/or its affiliates. All rights reserved.

The core layer is the backbone for connectivity and is the aggregation point for the other layers and modules in the Cisco Data Center Business Advantage architecture. The core must provide a high level of redundancy and must adapt to changes very quickly. Core devices are most reliable when they can accommodate failures by rerouting traffic and can respond quickly to changes in the network topology. The core devices must be able to implement scalable protocols and technologies, alternate paths, and load balancing. The core layer helps in scalability during future growth.

The core should be a high-speed Layer 3 switching environment that uses hardware-accelerated services. For fast convergence around a link or node failure, the core uses redundant point-to-point Layer 3 interconnections in the core. That type of design yields the fastest and most deterministic convergence results. The core layer should not perform any packet manipulation, such as checking access lists and filtering, which would slow down the switching of packets.

https://t.me/learningnets

certcollecion.net



Cisco Unified Fabric delivers transparent convergence, massive three-dimensional scalability, and sophisticated intelligent services to provide the following benefits:

- Support for traditional and virtualized data centers
- Reduction in TCO
- An increase in ROI

The five architectural components that impact TCO include the following:

- **Simplicity:** Businesses require the data center to be able to provide easy deployment, configuration, and consistent management of existing and new services.
- **Scale:** Data centers need to be able to support large Layer 2 domains that can provide massive scalability without the loss of bandwidth and throughput.
- **Performance:** Data centers should be able to provide deterministic latency and large bisectional bandwidth to applications and services as needed.
- **Resiliency:** The data center infrastructure and implemented features need to provide high availability to the applications and services they support.
- **Flexibility:** Businesses require a single architecture that can support multiple deployment models to provide the flexible component of the architecture.

Universal I/O brings efficiency to the data center through “wire-once” deployment and protocol simplification. This efficiency, in the Cisco WebEx data center, has shown the ability to increase workload density by 30 percent in a flat power budget. In a 30-megawatt (MW) data center, this increase accounts for an annual US$60 million cost deferral. Unified fabric technology enables a wire-once infrastructure in which there are no physical barriers in the network to redeploying applications or capacity, thus delivering hardware freedom.

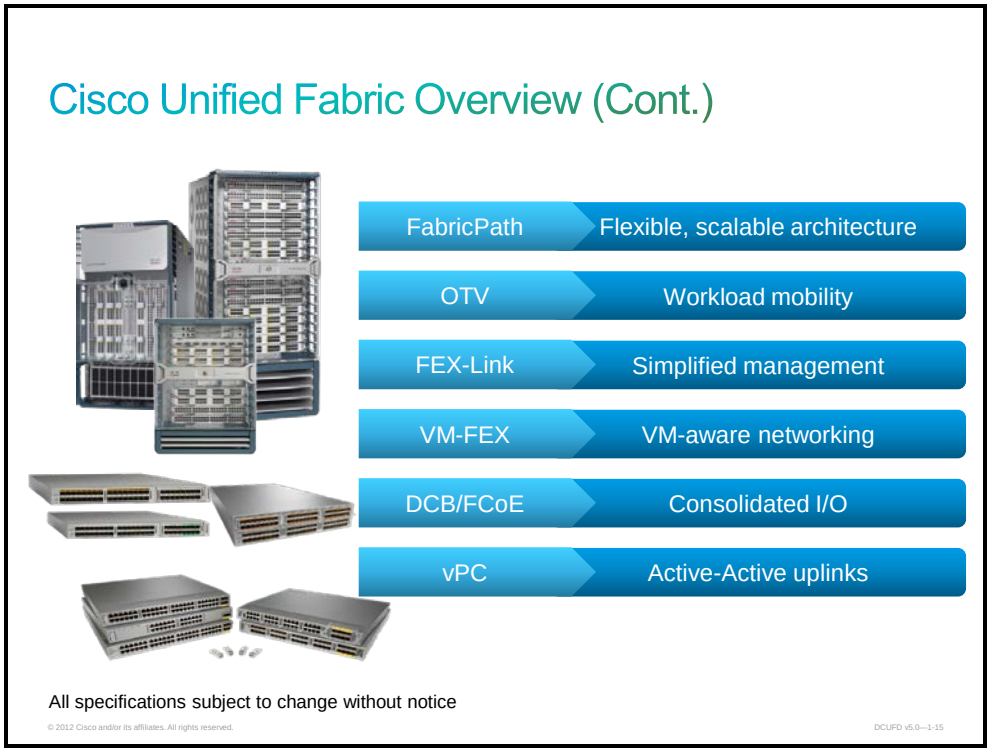
The main advantage of Cisco Unified Fabric is that it offers LAN and SAN infrastructure consolidation. It is no longer necessary to plan for and maintain two completely separate infrastructures. The network comes in as a central component to the evolution of the virtualized data center and to the enablement of cloud computing.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

Cisco Unified Fabric offers a low-latency and lossless connectivity solution that is fully virtualization-enabled. Cisco Unified Fabric offers you a massive reduction of cables, various adapters, switches, and pass-through modules.



Cisco Unified Fabric is a foundational pillar for the Cisco Data Center Business Advantage architectural framework. Cisco Unified Fabric complements Unified Network Services and Unified Computing to enable IT and business innovation.

- Cisco Unified Fabric convergence offers the best of both SANs and LANs by enabling users to take advantage of the Ethernet economy of scale, extensive vendor community, and future innovations.
- Cisco Unified Fabric scalability delivers performance, magnitude of ports and bandwidth, and geographic span.
- Cisco Unified Fabric intelligence embeds critical policy-based intelligent functionality into the unified fabric for both traditional and virtualized data centers.

To support the five architectural attributes, the Cisco Unified Fabric evolution is continuing to provide architectural innovations.

- **Cisco FabricPath:** Cisco FabricPath is a set of capabilities within the Cisco Nexus Operating System (NX-OS) Software that combines the plug-and-play simplicity of Ethernet with the reliability and scalability of Layer 3 routing. Cisco FabricPath enables companies to build highly scalable Layer 2 multipath networks without the Spanning Tree Protocol (STP). These networks are particularly suitable for large virtualization deployments, private clouds, and high-performance computing environments.
- **Cisco Overlay Transport Virtualization (OTV):** Cisco OTV is an industry-first solution that significantly simplifies extending Layer 2 applications across distributed data centers. Cisco OTV allows companies to deploy virtual computing resources and clusters across geographically distributed data centers to deliver transparent workload mobility, business resiliency, and superior computing resource efficiencies.

https://t.me/learningnets

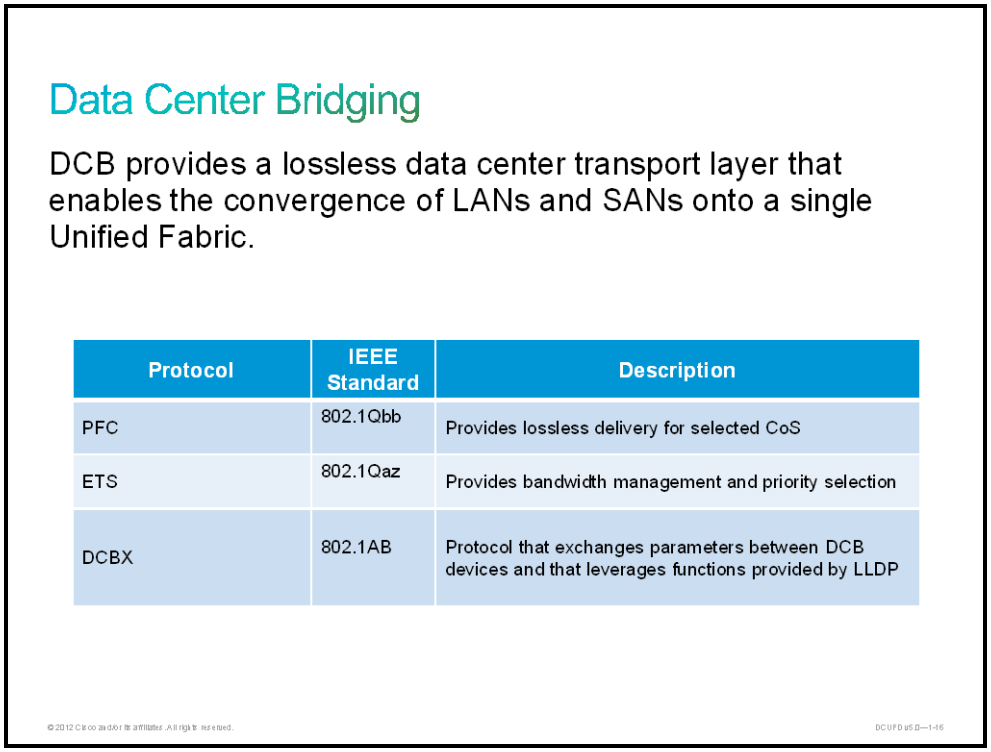
certcollecion.net

- **Cisco FEX-Link:** Cisco Fabric Extender (FEX)-Link technology enables data center designers to gain new design flexibility while simplifying cabling infrastructure and management complexity. Cisco FEX-Link uses the Cisco Nexus 2000 Series FEXs to extend the capacities and benefits that are offered by upstream Cisco Nexus switches.
- **Cisco VM-FEX**: Cisco Virtual Machine Fabric Extender (VM-FEX) provides advanced hypervisor switching as well as high-performance hardware switching. It is flexible, extensible, and service enabled. Cisco VM-FEX architecture provides virtualization-aware networking and policy control.
- **Data Center Bridging (DCB) and FCoE**: Cisco Unified Fabric provides the flexibility to run Fibre Channel, IP-based storage such as network-attached storage (NAS) and Small Computer System Interface over IP or FCoE, or a combination of these technologies, on a converged network.
- **vPC**: Cisco virtual port channel (vPC) technology enables the deployment of a link aggregation from a generic downstream network device to two individual and independent Cisco NX-OS devices (vPC peers). This multichassis link aggregation path provides both link redundancy and active-active link throughput that scales high-performance failover characteristics.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net



Data Center Bridging

Cisco Unified Fabric is a network that can transport many different protocols, such as LAN, SAN, and high-performance computing (HPC) protocols, over the same physical network.

10 Gigabit Ethernet is the basis for a new DCB protocol that, through enhanced features, provides a common platform for lossy and lossless protocols that carry LAN, SAN, and HPC data.

The IEEE 802.1 DCB is a collection of standards-based extensions to Ethernet and it can enable a Converged Enhanced Ethernet (CEE). It provides a lossless data center transport layer that enables the convergence of LANs and SANs onto a single unified fabric. In addition to supporting FCoE, DCB enhances the operation of iSCSI, NAS, and other business-critical traffic.

- **Priority-based flow control (PFC, IEEE 802.1Qbb):** Provides lossless delivery for selected classes of service
- **Enhanced Transmission Selection (ETS, IEEE 802.1Qaz):** Provides bandwidth management and priority selection
- **Quantized congestion notification (QCN, IEEE 802.1au):** Provides congestion awareness and avoidance (optional)

Note	Cisco equipment does not use QCN as a means to control congestion. Instead, PFC and ETS are used. Currently, Cisco does not have plans to implement QCN in its equipment.

- **Data Center Bridging Exchange (DCBX, IEEE 802.1AB):** Exchanges parameters between DCB devices and leverages functions that are provided by Link Layer Discovery Protocol (LLDP)

https://t.me/learningnets

certcollecion.net

Different organizations have created different names to identify the specifications. IEEE has used the term Data Center Bridging, or DCB. IEEE typically calls a standard specification by a number: for example, IEEE 802.1az. IEEE did not have a way to identify the group of specifications with a standard number, so the organization grouped the specifications into DCB.

The term Converged Enhanced Ethernet was created by IBM, again, to reflect the core group of specifications and to gain consensus among industry vendors (including Cisco) as to what a Version 0 list of the specifications would be before they all become standards.

FCoE

FCoE is a new protocol that is based on the Fibre Channel layers that are defined by the ANSI T11 committee, and it replaces the lower layers of Fibre Channel traffic. FCoE addresses the following:

- **Jumbo frames:** An entire Fibre Channel frame (2180 B in length) can be carried in the payload of a single Ethernet frame.
- **Fibre Channel port:** World wide name (WWN) addresses are encapsulated in the Ethernet frames and MAC addresses are used for traffic forwarding in the converged network.
- **FCoE Initialization Protocol (FIP):** This protocol provides a login for Fibre Channel devices into the fabric.
- **Quality of service (QoS) assurance:** This ability monitors the Fibre Channel traffic with respect to lossless delivery of Fibre Channel frames and bandwidth reservations for Fibre Channel traffic.
- **A minimum 10-Gb/s Ethernet platform**.

FCoE traffic consists of a Fibre Channel frame that is encapsulated within an Ethernet frame. The Fibre Channel frame payload may in turn carry SCSI messages and data or, in the future, may use FICON for mainframe traffic.

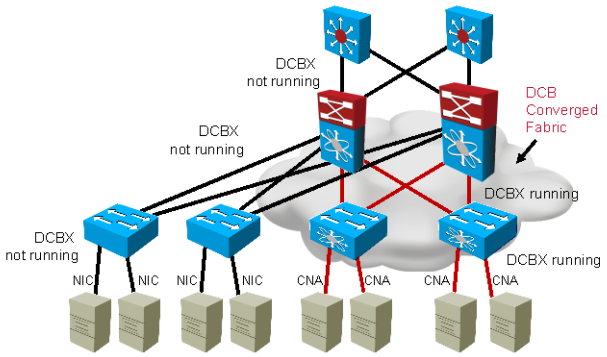
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

DCBX

- DCBX is a discovery and capability exchange protocol that is used by devices enabled for DCB.
- Autonegotiation of capabilities between DCB devices:
  - PFC
  - ETS
  - Congestion notification (as BCN and QCN)
  - Logical link down
  - NIV



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

The DCBX protocol allows each DCB device to communicate with other devices and to exchange capabilities within a unified fabric. Without DCBX, each device would not know if it could send lossless protocols like FCoE to another device that was not capable of dealing with lossless delivery.

DCBX is a discovery and capability exchange protocol that is used by devices that are enabled for Data Center Ethernet to exchange configuration information. The following parameters of the Data Center Ethernet features can be exchanged:

- Priority groups in ETS
- PFC
- Congestion notification (as backward congestion notification [BCN] or as Quantized Congestion Notification [QCN])
- Application types and capabilities
- Logical link down to signify the loss of a logical connection between devices even though the physical link is still up
- Network interface virtualization (NIV)

(See http://www.ieee802.org/1/files/public/docs2008/az-wadekar-dcbcxp-overview-rev0.2.pdf.)

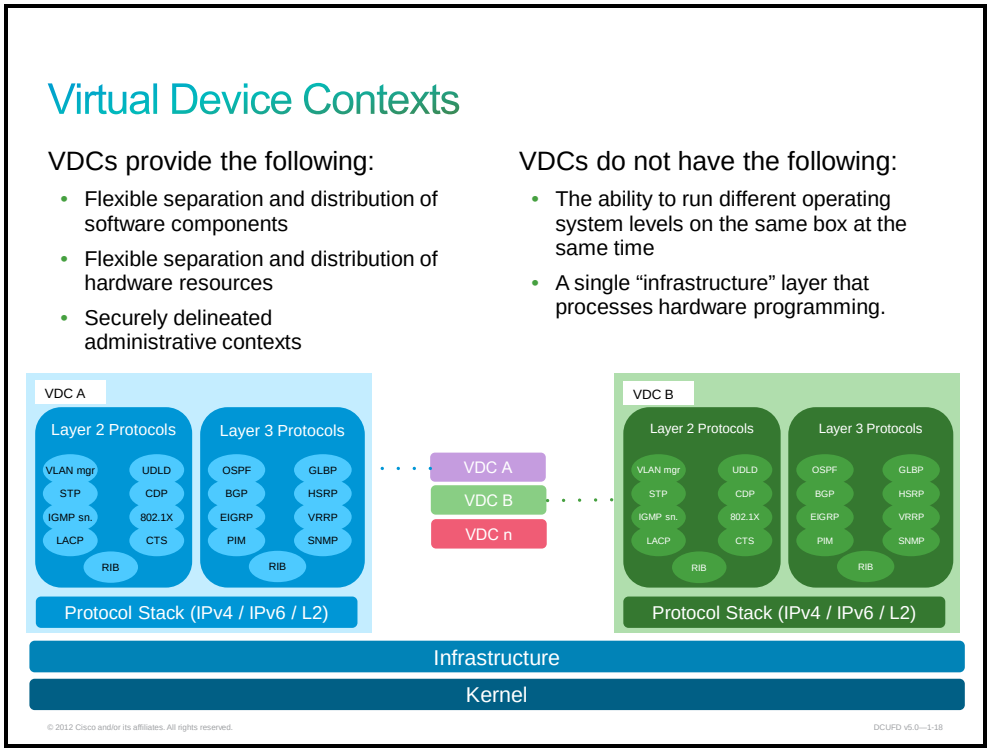
Devices need to discover the edge of the enhanced Ethernet cloud:

- Each edge switch needs to learn that it is connected to an existing switch.
- Servers need to learn whether they are connected to enhanced Ethernet devices.
- Within the enhanced Ethernet cloud, devices need to discover the capabilities of peers.
- The Data Center Bridging Capability Exchange Protocol (DCBCXP) utilizes the LLDP and processes the local operational configuration for each feature.
- Link partners can choose supported features and willingness to accept configurations from peers.

Details on DCBCXP can be found at http://www.intel.com/technology/eedc/index.htm.

https://t.me/learningnets

certcollecion.net



A VDC is used to virtualize the Cisco Nexus 7000 Switch and presents the physical switch as multiple logical devices. Each VDC contains its own unique and independent VLANs and virtual routing and forwarding (VRF), and each VDC is assigned its own physical ports.

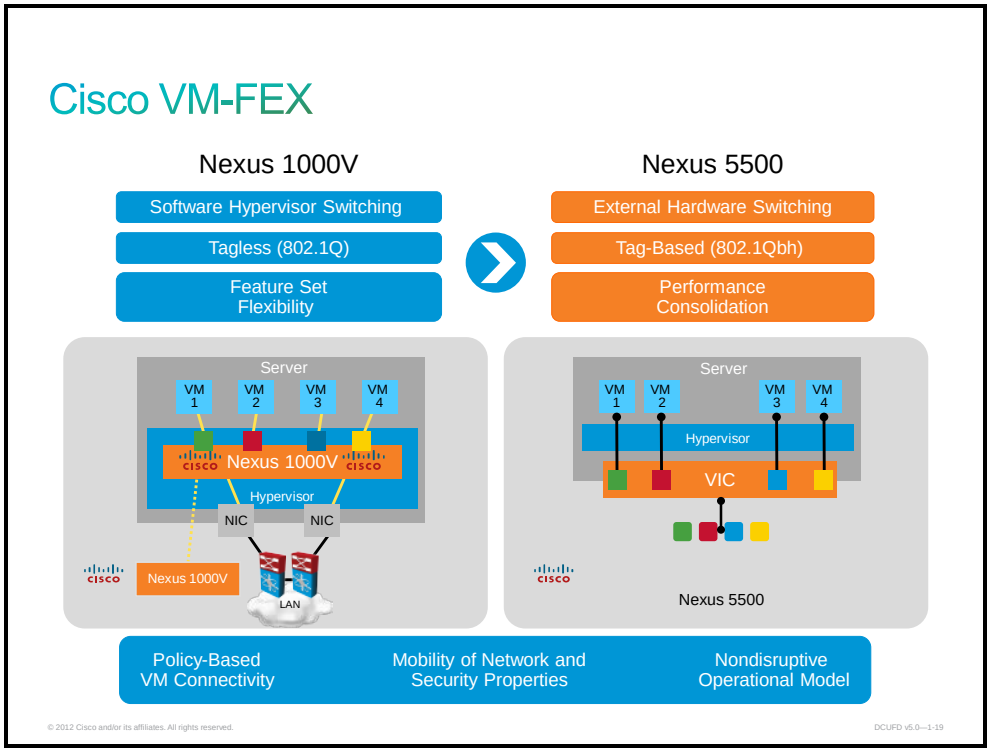
VDCs provide the following benefits:

- They can secure the network partition between different users on the same physical switch.
- They can provide departments with the ability to administer and maintain their own configurations.
- They can be dedicated for testing purposes without impacting production systems.
- They can consolidate the switch platforms of multiple departments onto a single physical platform.
- They can be used by network administrators and operators for training purposes.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net



Cisco VM-FEX encompasses a number of products and technologies that work together to improve server virtualization strategies:

- **Cisco Nexus 1000V Virtual Distributed Switch:** This switch is a software-based switch that was developed in collaboration with VMware. The switch integrates directly with the VMware ESXi hypervisor. Because the switch can combine the network and server resources, the network and security policies automatically follow a VM that is being migrated with VMware VMotion.
- **NIV:** This VM networking protocol was jointly developed by Cisco and VMware and allows the Cisco VM-FEX functions to be performed in hardware.
- **Cisco N-Port Virtualizer (NPV):** This function is currently available on the Cisco MDS 9000 Series Multilayer Switches and the Cisco Nexus 5000 and 5500 Series Switches. The Cisco NPV allows storage services to follow a VM as the VM moves.

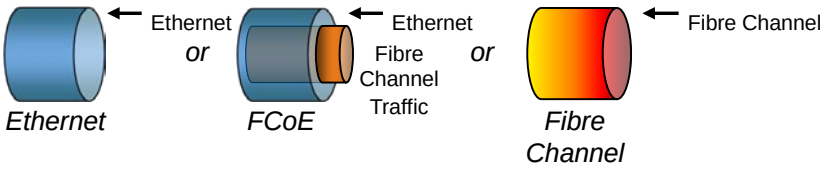
Cisco VM-FEX provides visibility down to the VM level, simplifying management, troubleshooting, and regulatory compliance.

https://t.me/learningnets

certcollecion.net

Unified Ports

- One port for all types of server I/O
  - Any port can be configured as
    - 1/10 Gigabit Ethernet,
    - DCB (lossless Ethernet),
    - FCoE on 10 Gigabit Ethernet (dedicated or converged link) or
    - 8/4/2/1G native Fibre Channel
- Flexibility of use
  - One standard chassis for all data center I/O needs



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

Unified ports are ports that can be configured as Ethernet or Fibre Channel. Unified ports are supported on Cisco Nexus 5500UP Switches and Cisco UCS 6200 Series Fabric Interconnects.

Unified ports support all existing port types including 1 Gigabit and 10 Gigabit Ethernet, FCoE, pure 1/2/4/8-Gb/s Fibre Channel interfaces such as any Nexus 5500UP port can be configured as 1 Gigabit/10 Gigabit Ethernet, DCB (lossless Ethernet), FCoE on 10 Gigabit Ethernet (dedicated or converged link) or 8/4/2/1-G native Fibre Channel port.

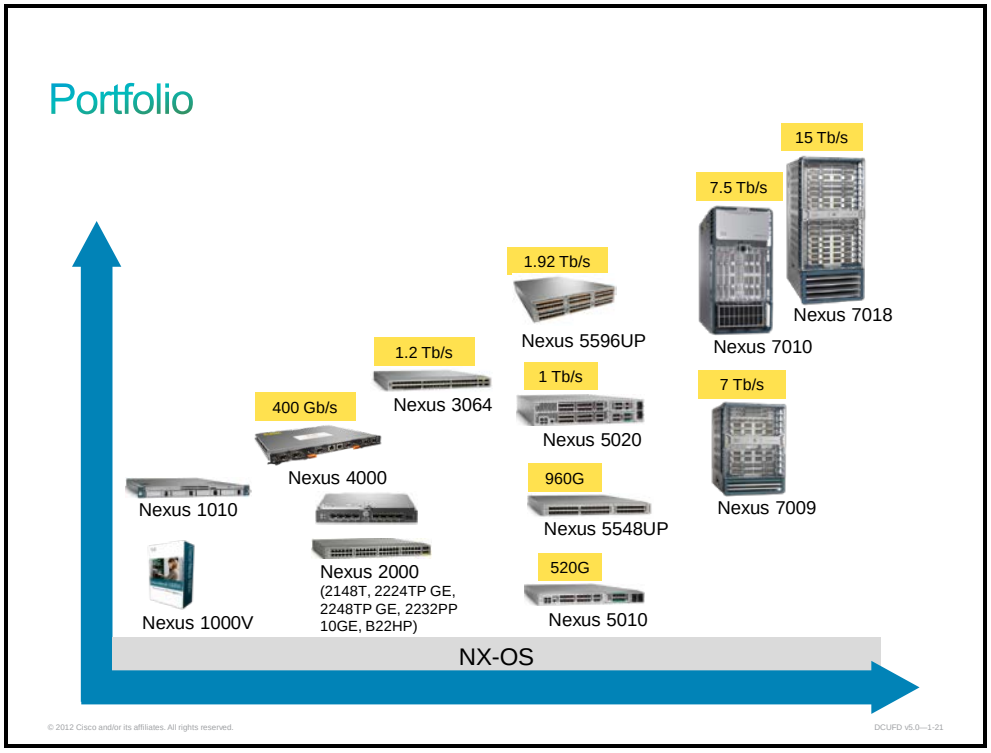
The benefits are as follows:

- Deploy a switch, such as the Cisco Nexus 5500UP, as a data center switch standard capable of all important I/O
- Mix Fibre Channel SAN to host, as well as switch and target with FCoE SAN
- Implement with native Fibre Channel today and enable smooth migration to FCoE in the future.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net



The Cisco Nexus product family comprises the following switches:

- **Cisco Nexus 1000V:** A virtual machine access switch that is an intelligent software switch implementation for VMware vSphere environments that run the Cisco Nexus Operating System (NX-OS) software. The Cisco Nexus 1000V operates inside the VMware ESX or ESXi hypervisor, and supports the Cisco Virtual Network Link (VN-Link) server virtualization technology to provide the following:
  - Policy-based virtual machine connectivity
  - Mobile virtual machine security and network policy
  - Nondisruptive operational model for server virtualization and networking teams
- **Cisco Nexus 1010 Virtual Services Appliance:** The appliance is a member of the Cisco Nexus 1000V Series Switches and hosts the Cisco Nexus 1000V Virtual Supervisor Module (VSM). It also supports the Cisco Nexus 1000V Network Analysis Module (NAM) Virtual Service Blade and provides a comprehensive solution for virtual access switching. The Cisco Nexus 1010 provides dedicated hardware for the VSM, making access switch deployment much easier for the network administrator.
- **Cisco Nexus 2000 Series FEX:** A category of data center products that are designed to simplify data center access architecture and operations. The Cisco Nexus 2000 Series uses the Cisco FEX-Link architecture to provide a highly scalable unified server-access platform across a range of 100-Mb/s Ethernet, Gigabit Ethernet, 10 Gigabit Ethernet, unified fabric, copper and fiber connectivity, and rack and blade server environments. The Cisco Nexus 2000 Series FEXs act as remote line cards for the Cisco Nexus 5000 and 5500 Series and the Cisco Nexus 7000 Series Switches.
- **Cisco Nexus 3000 Series Switches:** The Cisco Nexus 3000 Series Switches extend the comprehensive, proven innovations of the Cisco Data Center Business Advantage architecture into the High-Frequency Trading (HFT) market. The Cisco Nexus 3064 Switch supports 48 fixed 1/10-Gb/s Enhanced small form-factor pluggable plus (SFP+) ports and 4 fixed quad SFP+ (QSFP+) ports, which allow smooth transition from 10 Gigabit Ethernet to 40 Gigabit Ethernet. The Cisco Nexus 3064 Switch is well suited for financial

https://t.me/learningnets

certcollecion.net

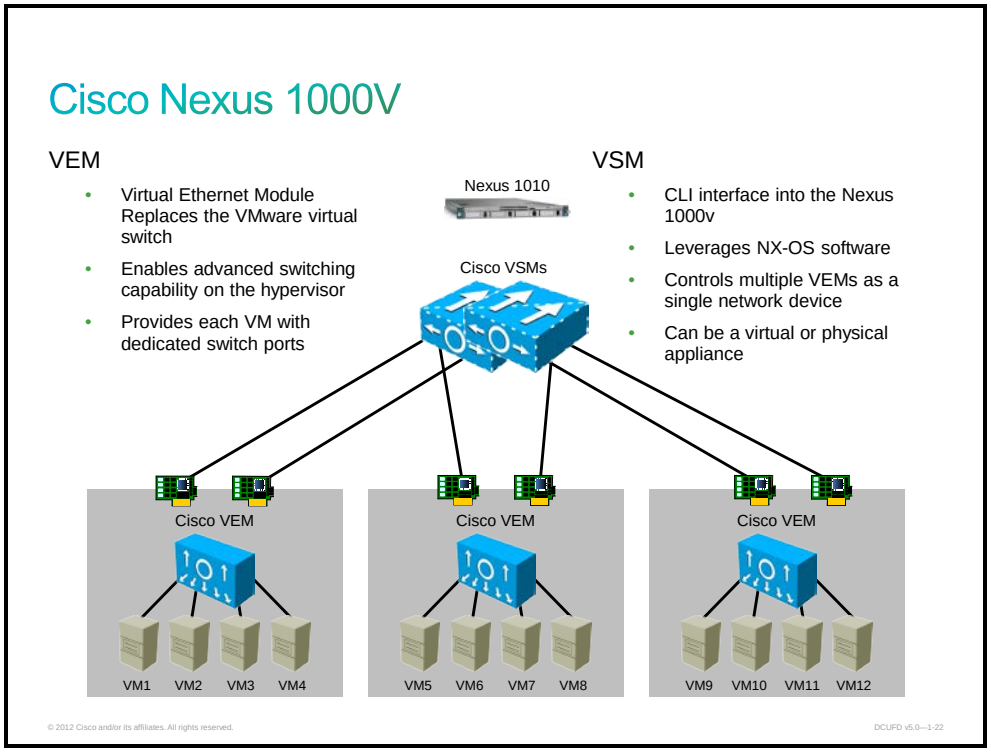
collocation deployments and delivers features such as latency of less than a microsecond, line-rate Layer 2 and 3 unicast and multicast switching, and support for 40 Gigabit Ethernet standards technologies.

- **Cisco Nexus 4000 Switch Module for IBM BladeCenter:** A blade switch solution for IBM BladeCenter H and HT chassis. This switch provides the server I/O solution that is required for high-performance, scale-out, virtualized and nonvirtualized x86 computing architectures. It is a line-rate, extremely low-latency, nonblocking, Layer 2, 10-Gb/s blade switch that is fully compliant with the INCITS FCoE and IEEE 802.1 DCB standards.
- **Cisco Nexus 5000 and 5500 Series Switches:** A family of line-rate, low-latency, lossless 10 Gigabit Ethernet and FCoE switches for data center applications. The Cisco Nexus 5000 Series Switches are designed for data centers transitioning to 10 Gigabit Ethernet as well as data centers ready to deploy a unified fabric that can manage LAN, SAN, and server clusters. This capability provides networking over a single link, with dual links used for redundancy.
- **Cisco Nexus 7000 Series Switches:** A modular data center-class switch that is designed for highly scalable 10 Gigabit Ethernet networks with a fabric architecture that scales beyond 15 Tb/s. The switch is designed to deliver continuous system operation and virtualized services. The Cisco Nexus 7000 Series Switches incorporate significant enhancements in design, power, airflow, cooling, and cabling. The 10-slot chassis has front-to-back airflow, making it a good solution for hot-aisle and cold-aisle deployments. The 18-slot chassis uses side-to-side airflow to deliver high density in a compact form factor.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net



The Cisco Nexus 1000V Virtual Ethernet Module provides Layer 2 switching functions in a virtualized server environment, and replaces virtual switches within the VMware ESX servers. This allows users to configure and monitor the virtual switch using the Cisco NX-OS CLI. The Cisco Nexus 1000V provides visibility into the networking components of the ESX servers and access to the virtual switches within the network.

The VMware vCenter server defines the data center that the Cisco Nexus 1000V will manage, with each server being represented as a line card, and managed as if it were a line card in a physical Cisco switch.

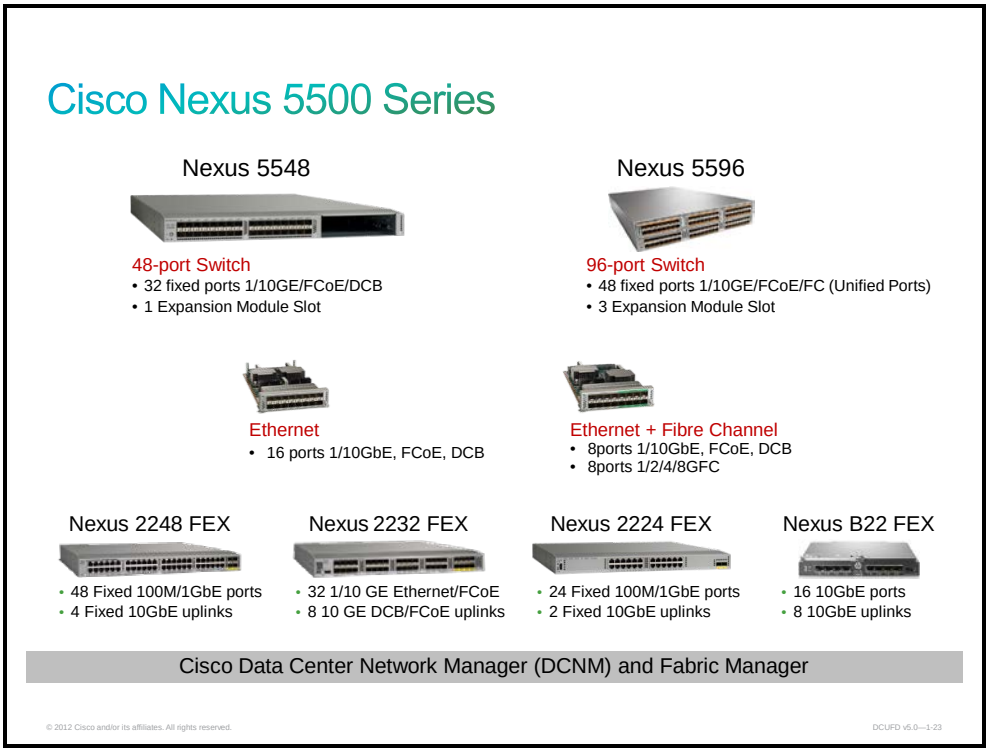
There are two components that are part of the Cisco Nexus 1000V implementation:

- **Virtual Supervisor Module (VSM):** This is the control software of the Cisco Nexus 1000V distributed virtual switch, and runs on either a virtual machine (VM) or as an appliance. It is based on the Cisco NX-OS Software.
- **Virtual Ethernet Module (VEM):** This is the part that actually switches the data traffic and runs on a VMware ESX 4.0 host. The VSM can control several VEMs, with the VEMs forming a switch domain that should be in the same virtual data center that is defined by VMware vCenter.

The Cisco Nexus 1000V is effectively a virtual chassis. It is modular, and ports can either be physical or virtual. The servers are modules on the switch, with each physical NIV port on a module being a physical Ethernet port. Modules 1 and 2 are reserved for the VSM, with the first server or host automatically being assigned to the next available module number. The ports to which the virtual network interface card (vNIC) interfaces connect are virtual ports on the Cisco Nexus 1000V, where they are assigned a global number.

https://t.me/learningnets

certcollecion.net



Cisco Nexus 5548UP Switch

The Cisco Nexus 5548UP Switch is the first of the Cisco Nexus 5500 platform. It is a 1-RU, 10 Gigabit Ethernet and FCoE switch offering up to 960-Gb/s throughput and up to 48 ports. The switch has 32 1- or 10-Gb/s FCOE Fibre Channel ports and one expansion slot.

Cisco Nexus 5596UP Switch

The Cisco Nexus 5596UP Switch has 48 fixed ports capable of supporting 1- or 10-Gb/s FCOE Fibre Channel and three additional slots. The three additional slots will accommodate any of the expansion modules for the Cisco Nexus 5500 Series Switches, taking the maximum capacity for the switch to 96 ports. These ports are unified ports that provide flexibility regarding the connectivity requirements, and the switch offers 1.92-Tb/s throughput.

Both of the Cisco Nexus 5500 Series Switches support Cisco FabricPath, and with the Layer 3 routing module, both Layer 2 and Layer 3 support is provided. The Cisco Nexus 5500 Series supports the same Cisco Nexus 2200 Series FEXs.

Expansion Modules for the Cisco 5500 Series Switches

The Cisco Nexus 5500 Series Switches support the following expansion modules:

- Ethernet module that provides 16 x 1/10 Gigabit Ethernet and FCoE ports using SFP+ interfaces.
- Fibre Channel plus Ethernet module that provides 8 x 1/10 Gigabit Ethernet and FCoE ports using the SFP+ interface, and eight ports of 8/4/2/1-Gb/s native Fibre Channel connectivity using the SFP interface.
- A Layer 3 daughter card for routing functionality

The modules for the Cisco Nexus 5500 Series Switches are not backward-compatible with the Cisco Nexus 5000 Series Switches.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

Cisco Nexus 2000 Series FEXs

The Cisco Nexus 2000 Series FEXs offer front-to-back cooling, compatibility with data center hot-aisle and cold-aisle designs, placement of all switch ports at the rear of the unit in close proximity to server ports, and accessibility of all user-serviceable components from the front panel. The Cisco Nexus 2000 Series has redundant hot-swappable power supplies, a hot-swappable fan tray with redundant fans, and is a 1-RU form factor.

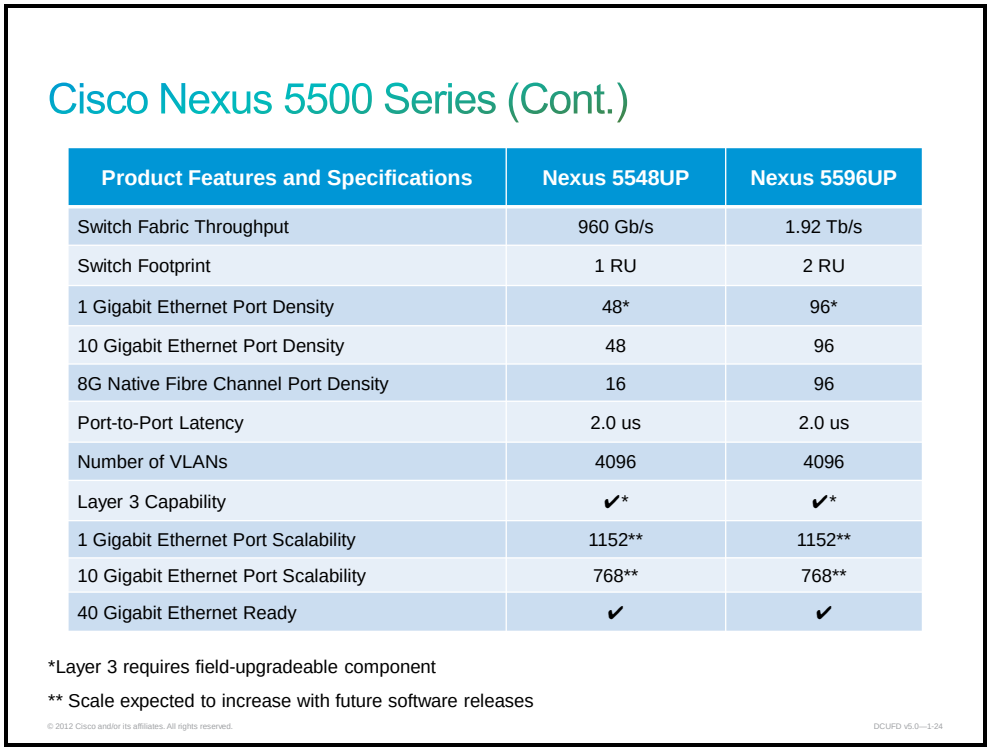
The Cisco Nexus 2000 Series has two types of ports: ports for end-host attachment and uplink ports. The Cisco Nexus 2000 Series is an external line module for the Cisco Nexus 5000 and 5500 Series Switches, and for the Cisco Nexus 7000 Series.

- **Cisco Nexus 2148T FEX:** 48 x 1000BASE-T ports and 4 x 10 Gigabit Ethernet uplinks (SFP+)
- **Cisco Nexus 2224TP GE FEX:** 24 x 100/1000BASE-T ports and 2 x 10 Gigabit Ethernet uplinks (SFP+)
- **Cisco Nexus 2248TP GE FEX:** 48 x 100/1000BASE-T ports and 4 x 10 Gigabit Ethernet uplinks (SFP+). This model is supported as an external line module for the Cisco Nexus 7000 Series using Cisco NX-OS 5.1(2) software.
- **Cisco Nexus 2232PP 10GE FEX:** 32 x 1/10 Gigabit Ethernet and FCoE ports (SFP+) and 8 x 10 Gigabit Ethernet FCoE uplinks (SFP+).
- **Cisco B22 Blade FEX for HP:** 16 x 10 Gigabit Ethernet host interfaces and 8 x 10 Gigabit Ethernet FCoE uplinks (SFP+).

Note Cisco Nexus 5500 Series Switches and Cisco Nexus 2200 Series FEXs can be ordered with front-to-back or back-to-front airflow direction, depending on the fan tray that is ordered. This way you can achieve the desired switch orientation and still fit into your hot aisle-cold aisle thermal model.

https://t.me/learningnets

certcollecion.net



The table in the figure describes the differences between the Cisco Nexus 5000 and 5500 Series Switches. The port counts are based on 24 Cisco Nexus 2000 FEXs per Cisco Nexus 5500 Switch.

Cisco Nexus 5500 Platform Features

The Cisco Nexus 5500 Series is the second generation of access switches for 10 Gigabit Ethernet connectivity. The Cisco Nexus 5500 platform provides a rich feature set that makes it well suited for top-of-rack (ToR), middle-of-row (MoR), or end-of-row (EoR) access-layer applications. It protects investments in data center racks with standards-based 1 and 10 Gigabit Ethernet and FCoE features, and virtual machine awareness features that allow IT departments to consolidate networks. The combination of high port density, lossless Ethernet, wire-speed performance, and extremely low latency makes this switch family well suited to meet the growing demand for 10 Gigabit Ethernet. The family can support unified fabric in enterprise and service provider data centers, which protects the investments of enterprises. The switch family has sufficient port density to support single and multiple racks that are fully populated with blade and rack-mount servers.

- **High density and high availability:** The Cisco Nexus 5548P provides 48 1/10-Gb/s ports in 1 RU, and the upcoming Cisco Nexus 5596UP Switch provides a density of 96 1/10-Gb/s ports in 2RUs. The Cisco Nexus 5500 Series is designed with redundant and hot-swappable power and fan modules that can be accessed from the front panel, where status lights offer an at-a-glance view of switch operation. To support efficient data center hot- and cold-aisle designs, front-to-back cooling is used for consistency with server designs.
- **Nonblocking line-rate performance:** All the 10 Gigabit Ethernet ports on the Cisco Nexus 5500 platform can manage packet flows at wire speed. The absence of resource sharing helps ensure the best performance of each port regardless of the traffic patterns on other ports. The Cisco Nexus 5548P can have 48 Ethernet ports at 10 Gb/s sending packets simultaneously without any effect on performance, offering true 960-Gb/s bidirectional bandwidth. The upcoming Cisco Nexus 5596UP can have 96 Ethernet ports at 10 Gb/s, offering true 1.92-Tb/s bidirectional bandwidth.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

- **Low latency:** The cut-through switching technology that is used in the ASICs of the Cisco Nexus 5500 Series enables the product to offer a low latency of 2 microseconds, which remains constant regardless of the size of the packet being switched. This latency was measured on fully configured interfaces, with access control lists (ACLs), QoS, and all other data path features turned on. The low latency on the Cisco Nexus 5500 Series, together with a dedicated buffer per port and the congestion management features, make the Cisco Nexus 5500 platform an excellent choice for latency-sensitive environments.
- **Single-stage fabric:** The crossbar fabric on the Cisco Nexus 5500 Series is implemented as a single-stage fabric, thus eliminating any bottleneck within the switches. Single-stage fabric means that a single crossbar fabric scheduler has complete visibility into the entire system and can therefore make optimal scheduling decisions without building congestion within the switch. With a single-stage fabric, the congestion becomes exclusively a function of your network design; the switch does not contribute to it.
- **Congestion management:** Keeping latency low is not the only critical element for a high-performance network solution. Servers tend to generate traffic in bursts, and when too many bursts occur at the same time, a short period of congestion occurs. Depending on how the burst of congestion is smoothed out, the overall network performance can be affected. The Cisco Nexus 5500 platform offers a complete range of congestion management features to reduce congestion. These features address congestion at different stages and offer granular control over the performance of the network.
  - **Virtual output queues:** The Cisco Nexus 5500 platform implements virtual output queues (VOQs) on all ingress interfaces, so that a congested egress port does not affect traffic that is directed to other egress ports. Every IEEE 802.1p class of service (CoS) uses a separate VOQ in the Cisco Nexus 5500 platform architecture, resulting in a total of 8 VOQs per egress on each ingress interface, or a total of 384 VOQs per ingress interface on the Cisco Nexus 5548P, and a total of 768 VOQs per ingress interface on the Cisco Nexus 5596UP. The extensive use of VOQs in the system helps ensure high throughput on a per-egress, per-CoS basis. Congestion on one egress port in one CoS does not affect traffic that is destined for other classes of service or other egress interfaces. This ability avoids head-of-line (HOL) blocking, which would otherwise cause congestion to spread.
  - **Separate egress queues for unicast and multicast:** Traditionally, switches support eight egress queues per output port, each servicing one IEEE 802.1p CoS. The Cisco Nexus 5500 platform increases the number of egress queues by supporting 8 egress queues for unicast and 8 egress queues for multicast. This support allows separation of unicast and multicast that are contending for system resources within the same CoS and provides more fairness between unicast and multicast. Through configuration, the user can control the amount of egress port bandwidth for each of the 16 egress queues.
  - **Lossless Ethernet with PFC**: By default, Ethernet is designed to drop packets when a switching node cannot sustain the pace of the incoming traffic. Packet drops make Ethernet very flexible in managing random traffic patterns that are injected into the network. However, they effectively make Ethernet unreliable and push the burden of flow control and congestion management up to a higher level in the network stack.

    PFC offers point-to-point flow control of Ethernet traffic that is based on IEEE 802.1p CoS. With a flow-control mechanism in place, congestion does not result in drops, which transforms Ethernet into a reliable medium. The CoS granularity allows some classes of service to have a reliable no-drop behavior, while allowing other classes to retain traditional best-effort Ethernet behavior. The no-drop benefits are significant for any protocol that assumes reliability at the media level, such as FCoE.

https://t.me/learningnets

certcollecion.net

- **Explicit congestion notification (ECN) marking**: ECN is an extension to TCP/IP that is defined in RFC 3168. ECN allows end-to-end notification of network congestion without dropping packets. Traditionally, TCP detects network congestion by observing dropped packets. When congestion is detected, the TCP sender takes action by controlling the flow of traffic. However, dropped packets can sometimes lead to long TCP timeouts and consequent loss of throughput. The Cisco Nexus 5500 platform can set a mark in the IP header so that instead of dropping a packet, it sends a signal impending congestion. The receiver of the packet echoes the congestion indicator to the sender, which must respond as though congestion had been indicated by packet drops.

- **FCoE:** FCoE is a standards-based encapsulation of Fibre Channel frames into Ethernet frames. By implementing FCoE, the Cisco Nexus 5500 platform enables storage I/O consolidation in addition to Ethernet.

- **NIV architecture:** The introduction of blade servers and server virtualization has increased the number of access-layer switches that need to be managed. In both cases, an embedded switch or softswitch requires separate management. NIV enables a central switch to create an association with the intermediate switch, whereby the intermediate switch becomes the data path to the central forwarding and policy enforcement under control of the central switch. This scheme enables both a single point of management and a uniform set of features and capabilities across all access-layer switches.

  One critical implementation of NIV in the Cisco Nexus 5000 and 5500 Series is the Cisco Nexus 2000 Series FEXs and their deployment in data centers. A Cisco Nexus 2000 Series FEX behaves as a virtualized remote I/O module, enabling the Cisco Nexus 5500 platform to operate as a virtual modular chassis.

- **IEEE 1588 Precision Time Protocol (PTP):** In financial environments, particularly high-frequency trading environments, transactions occur in less than a millisecond. For accurate application performance monitoring and measurement, the systems supporting electronic trading applications must be synchronized with extremely high accuracy (to less than a microsecond). IEEE 1588 is designed for local systems that require very high accuracy beyond that attainable using Network Time Protocol (NTP). The Cisco Nexus 5500 platform supports IEEE 1588 boundary clock synchronization. In other words, the Cisco Nexus 5500 platform will run PTP and synchronize to an attached master clock, and the boundary clock will then act as a master clock for all attached slaves. The Cisco Nexus 5500 platform also supports packet time stamping by including the IEEE 1588 time stamp in the Encapsulated Remote Switched Port Analyzer (ERSPAN) header.

- **Cisco FabricPath and Transparent Interconnection of Lots of Links (TRILL):** Existing Layer 2 networks that are based on STP have a number of challenges to overcome. These challenges include suboptimal path selection, underutilized network bandwidth, control-plane scalability, and slow convergence. Although enhancements to STP and features such as Cisco vPC technology help mitigate some of these limitations, these Layer 2 networks lack fundamentals that limit their scalability.

  Cisco FabricPath and TRILL are two emerging solutions for creating scalable and highly available Layer 2 networks. Cisco Nexus 5500 Series hardware is capable of switching packets that are based on Cisco FabricPath headers or TRILL headers. This capability enables customers to deploy scalable Layer 2 networks with native Layer 2 multipathing.

- **Layer 3:** The design of the access layer varies depending on whether Layer 2 or Layer 3 is used at the access layer. The access layer in the data center is typically built at Layer 2. Building at Layer 2 allows better sharing of service devices across multiple servers and allows the use of Layer 2 clustering, which requires the servers to be near Layer 2. In some designs, such as two-tier designs, the access layer may be Layer 3, although this may not

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

imply that every port on these switches is a Layer 3 port. The Cisco Nexus 5500 platform can operate in Layer 3 mode with the addition of a routing module.

- **Hardware-level I/O consolidation:** The Cisco Nexus 5500 platform ASICs can transparently forward Ethernet, Fibre Channel, FCoE, Cisco FabricPath, and TRILL, providing true I/O consolidation at the hardware level. The solution that is adopted by the Cisco Nexus 5500 platform reduces the costs of consolidation through a high level of integration in the ASICs. The result is a full-featured Ethernet switch and a full-featured Fibre Channel switch that is combined into one product.

Cisco Nexus 7000 Series

- 15+ Tb/s system
- DCB and FCoE support
- Continuous operations
- Device virtualization
- Modular OS
- Cisco TrustSec





	Nexus 7009	Nexus 7010	Nexus 7018
Slots	7 I/O + 2 sup	8 I/O + 2 sup	16 I/O + 2 sup
Height	14 RU	21 RU	25 RU
BW / Slot Fab 1	N/A	230 Gig / slot	230 Gig / slot
BW / Slot Fab 2	550 Gig / Slot	550 Gig / slot	550 Gig / slot

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

The Cisco Nexus 7000 Series Switches offer a modular data center-class product that is designed for highly scalable 10 Gigabit Ethernet networks with a fabric architecture that scales beyond 15 Tb/s. The Cisco Nexus 7000 Series provides integrated resilience that is combined with features optimized specifically for the data center for availability, reliability, scalability, and ease of management.

The Cisco Nexus 7000 Series Switches run the Cisco NX-OS software to deliver a rich set of features with nonstop operation.

- Front-to-back airflow with 10 front-accessed vertical module slots and an integrated cable management system facilitates installation, operation, and cooling in both new and existing facilities.
- 18 front-accessed module slots with side-to-side airflow in a compact horizontal form factor with purpose-built integrated cable management ease operation and reduce complexity.
- Designed for reliability and maximum availability, all interface and supervisor modules are accessible from the front. Redundant power supplies, fan trays, and fabric modules are accessible completely from the rear to ensure that cabling is not disrupted during maintenance.

https://t.me/learningnets

certcollecion.net

- The system uses dual dedicated supervisor modules and fully distributed fabric architecture. There are five rear-mounted fabric modules, which combined with the chassis midplane, deliver up to 230 Gb/s per slot for 4.1 Tb/s of forwarding capacity in the 10-slot form factor, and 7.8 Tb/s in the 18-slot form factor using the Cisco Fabric Module 1. Migrating to the Cisco Fabric Module 2 increases the bandwidth per slot to 550 Gb/s. This increases the forwarding capacity on the 10-slot form factor to 9.9 Tb/s and on the 18-slot form factor to 18.7 Tb/s.
- The midplane design supports flexible technology upgrades as your needs change and provides ongoing investment protection.

Cisco Nexus 7000 Series 9-Slot Chassis

- The Cisco Nexus 7000 Series 9-slot chassis with up to 7 I/O module slots supports up to 224 10 Gigabit Ethernet or 336 Gigabit Ethernet ports.
- Airflow is side-to-side.
- The integrated cable management system is designed to support the cabling requirements of a fully configured system to either or both sides of the switch, allowing maximum flexibility. All system components can easily be removed with the cabling in place, providing ease of maintenance tasks with minimal disruption.
- A series of LEDs at the top of the chassis provides a clear summary of the status of the major system components. The LEDs alert operators to the need to conduct further investigation. These LEDs report the power supply, fan, fabric, supervisor, and I/O module status.
- The purpose-built optional front module door provides protection from accidental interference with both the cabling and modules that are installed in the system. The transparent front door allows easy observation of cabling and module indicators and status lights without any need to open the doors. The door supports a dual-opening capability for flexible operation and cable installation while attached. The door can be completely removed for both initial cabling and day-to-day management of the system.
- Independent variable-speed system and fabric fans provide efficient cooling capacity to the entire system. Fan tray redundancy features help ensure reliability of the system and support for hot swapping of fan trays.
- The crossbar fabric modules are located in the front of the chassis, with support for two supervisors.

Cisco Nexus 7000 Series 10-Slot Chassis

- The Cisco Nexus 7000 Series 10-slot chassis, with up to 8 I/O module slots, supports up to 256 10 Gigabit Ethernet or 384 Gigabit Ethernet ports, meeting the demands of large deployments.
- Front-to-back airflow helps ensure that use of the Cisco Nexus 7000 Series 10-slot chassis addresses the requirement for hot-aisle and cold-aisle deployments without additional complexity.
- The system uses dual system and fabric fan trays for cooling. Each fan tray is redundant and composed of independent variable-speed fans that automatically adjust to the ambient temperature. This adjustment helps reduce power consumption in well-managed facilities while providing optimum operation of the switch. The system design increases cooling efficiency and provides redundancy capabilities, allowing hot swapping without affecting the system. If either a single fan or a complete fan tray fails, the system continues to operate without a significant degradation in cooling capacity.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

- The integrated cable management system is designed for fully configured systems. The system allows cabling either to a single side or to both sides for maximum flexibility without obstructing any important components. This flexibility eases maintenance even when the system is fully cabled.
- The system supports an optional air filter to help ensure clean airflow through the system. The addition of the air filter satisfies Network Equipment Building System (NEBS) requirements.
- A series of LEDs at the top of the chassis provides a clear summary of the status of the major system components. The LEDs alert operators to the need to conduct further investigation. These LEDs report the power supply, fan, fabric, supervisor, and I/O module status.
- The cable management cover and optional front module doors provide protection from accidental interference with both the cabling and modules that are installed in the system. The transparent front door allows observation of cabling and module indicator and status lights.

Cisco Nexus 7000 Series 18-Slot Chassis

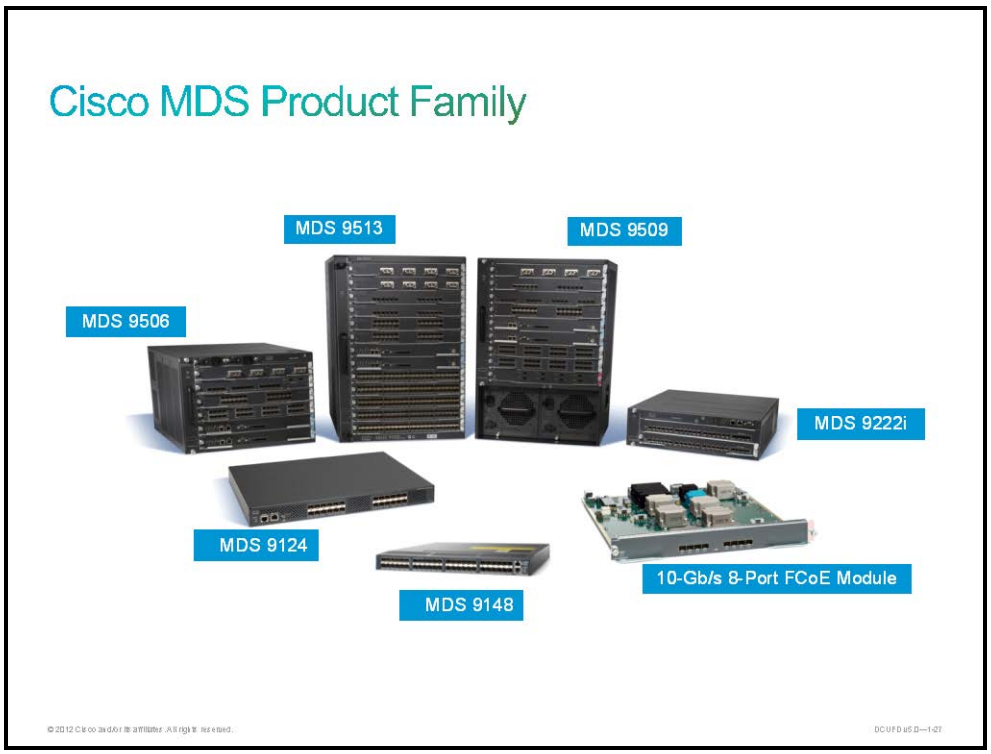
- The Cisco Nexus 7000 Series 18-slot chassis with up to 16 I/O module slots supports up to 512 10 Gigabit Ethernet or 768 Gigabit Ethernet ports, meeting the demands of the largest deployments.
- Side-to-side airflow increases the system density within a 25-RU footprint, optimizing the use of rack space. The optimized density provides more than 16 RU of free space in a standard 42-RU rack for cable management and patching systems.
- The integrated cable management system is designed to support the cabling requirements of a fully configured system to either or both sides of the switch, allowing maximum flexibility. All system components can easily be removed with the cabling in place, providing ease of maintenance tasks with minimal disruption.
- A series of LEDs at the top of the chassis provides a clear summary of the status of the major system components. The LEDs alert operators to the need to conduct further investigation. These LEDs report the power supply, fan, fabric, supervisor, and I/O module status.
- The purpose-built optional front module door provides protection from accidental interference with both the cabling and modules that are installed in the system. The transparent front door allows easy observation of cabling and module indicators and status lights without any need to open the doors. The door supports a dual-opening capability for flexible operation and cable installation while fitted. The door can be completely removed for both initial cabling and day-to-day management of the system.
- Independent variable-speed system and fabric fans provide efficient cooling capacity to the entire system. Fan tray redundancy features help ensure reliability of the system and support for hot swapping of fan trays.

https://t.me/learningnets

certcollecion.net

Cisco Data Center Architecture Storage

This topic describes the Cisco Data Center architectural framework storage component.



Cisco MDS SAN Switches

Cisco Multilayer Director Switches (MDS) and directors and the line card modules provide connectivity in a SAN. Since their introduction in 2002, the Cisco MDS switches and directors have embodied many innovative features that help improve performance and help overcome some of the limitations present in many SANs today. One of the benefits of the Cisco MDS products is that the chassis supports several generations of line card modules without modification. As Fibre Channel speeds have increased from 2 to 4 Gb/s and are now 8 Gb/s, new line card modules have been introduced to support those faster data rates. These line card modules can be installed in existing chassis without having to replace them with new ones.

Multilayer switches are switching platforms with multiple layers of intelligent features, such as the following:

- Ultrahigh availability
- Scalable architecture
- Comprehensive security features
- Ease of management
- Advanced diagnostics and troubleshooting capabilities
- Transparent integration of multiple technologies
- Multiprotocol support

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

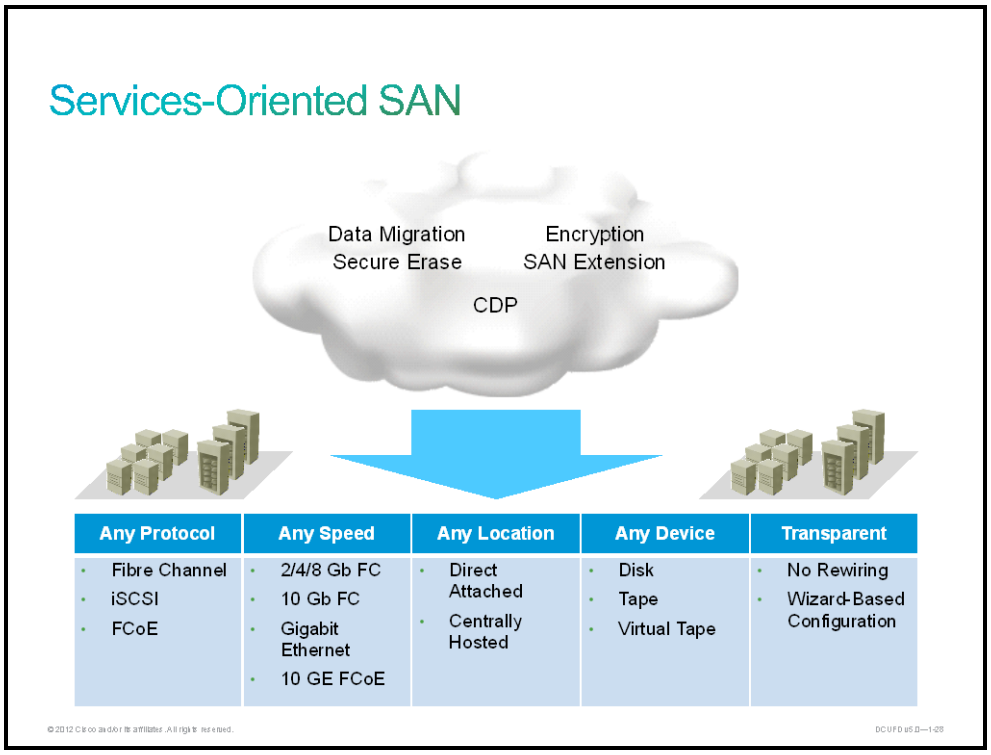
certcollecion.net

The Cisco MDS 9500 Series products offer industry-leading investment protection and offer a scalable architecture with highly available hardware and software. Based on the Cisco MDS 9000 Series operating system and a comprehensive management platform in Cisco Fabric Manager, the Cisco MDS 9500 Series offers various application line card modules and a scalable architecture from an entry-level fabric switch to director-class systems.

The Cisco MDS DS-X9708-K9 module has eight 10 Gigabit Ethernet multihop-capable FCoE ports. It enables extension of FCoE beyond the access layer into the core of the data center with a full line-rate FCoE module for the Cisco MDS 9500 Series Multilayer Directors.

The Cisco MDS 9222i Multiservice Modular Switch uses the 18/4 architecture of the DS-X9304-18K9 line card and includes native support for Cisco MDS Storage Media Encryption (SME) along with all the features of the Cisco MDS 9216i Multilayer Fabric Switch.

The Cisco MDS 9148 switch is an 8-Gb/s Fibre Channel switch providing forty-eight 2-, 4-, or 8-Gb/s Fibre Channel ports. The base license supports 16-, 32-, or 48-port models but can be expanded to use the 8-port license.

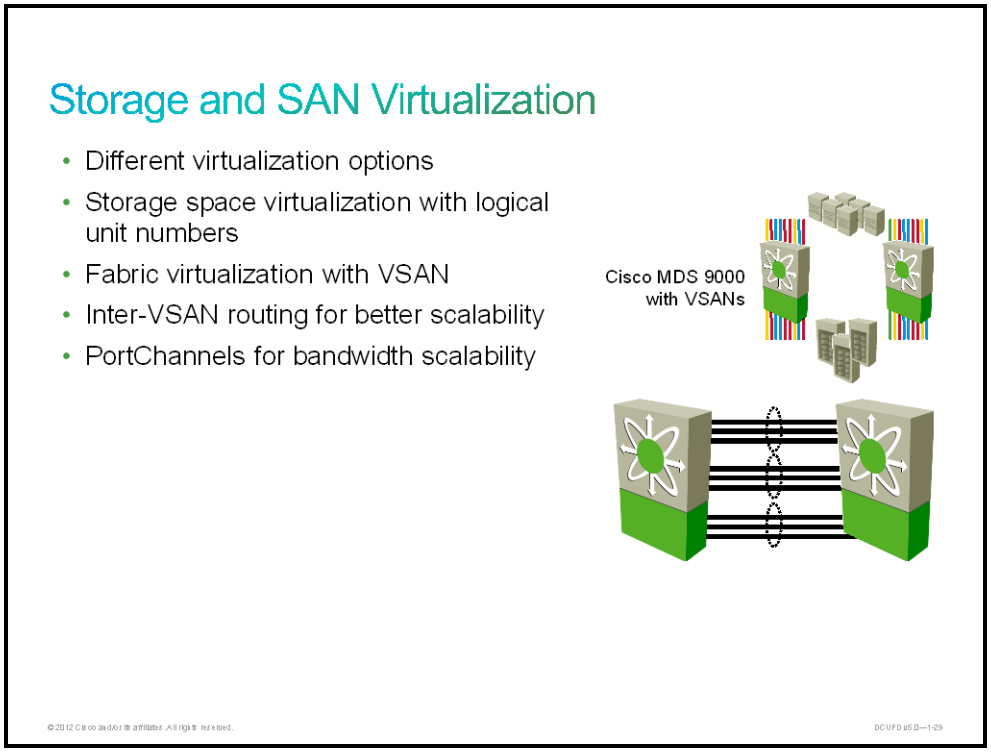


Services-oriented SAN fabrics can transparently extend any of the following SAN services to any device within the fabric:

- Data Mobility Manager (DMM), which provides online data migration
- LinkSec encryption, which encrypts Fibre Channel traffic
- Secure Erase, which permanently erases data
- SAN extension features like Write Acceleration and Tape Read Acceleration, which reduce overall latency
- Continuous data protection (CDP) is enabled by the Cisco SANTap feature

https://t.me/learningnets

certcollecion.net



SAN Islands

A SAN island refers to a physically isolated switch or group of switches that is used to connect hosts to storage devices. Today, SAN designers build separate fabrics, otherwise known as SAN islands, for various reasons.

Reasons for building SAN islands may include the desire to isolate different applications in their own fabrics or to raise availability by minimizing the impact of fabric-wide disruptive events. In addition, separate SAN islands also offer a higher degree of security because each physical infrastructure contains its own separate set of fabric services and management access.

VSAN Scalability

VSAN functionality is a feature that was developed by Cisco that can leverage the advantages of isolated SAN fabrics with capabilities that address the limitations of isolated SAN islands.

VSANs provide a method for allocating ports within a physical fabric to create virtual fabrics. Independent physical SAN islands are virtualized onto a common SAN infrastructure.

An analogy is that VSANs on Fibre Channel switches are like VDCs on Cisco Nexus 7000 Series Ethernet switches. VSANs can virtualize the physical switch into many virtual switches.

Using VSANs, SAN designers can raise the efficiency of a SAN fabric and alleviate the need to build multiple physically isolated fabrics to meet organizational or application needs. Instead, fewer and less-costly redundant fabrics can be built, each housing multiple applications, and can still provide island-like isolation.

VSANs provide not only hardware-based isolation but also a complete replicated set of Fibre Channel services for each VSAN. Therefore, when a VSAN is created, a completely separate set of fabric services, configuration management capability, and policies are created within the new VSAN.

Each separate virtual fabric is isolated from another by using a hardware-based frame-tagging mechanism on VSAN member ports and Enhanced Inter-Switch Links (EISLs). The EISL type has been created and includes added tagging information for each frame within the fabric. The EISL is supported on links that interconnect any Cisco MDS 9000 Series Switch product.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

Membership in a VSAN is based on the physical port, and no physical port may belong to more than one VSAN. Therefore, the node that is connected to a physical port becomes a member of that VSAN port.

Cisco NPV-Enabled Environment

NPV Edge Switches

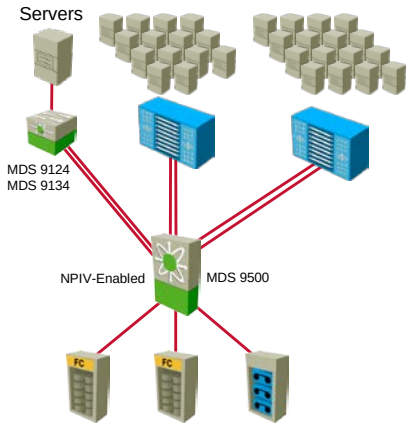
- Need to enable switch in NPV mode
- Changing to and from NPV mode is disruptive:
  - The switch reboots.
  - The configuration is not saved.
- Supports only F, SD, and NP modes
- Supports 16 VSANs
- Local switching is not supported.
  - Switching is done at the core,

NPV Core Switches

- Must enable the N-Port ID Virtualization (NPIV) feature
- Supports up to 100 NPV edge switches

NPV-enabled switches are standards-based and interoperable with other third-party switches in the SAN.

F = fabric mode, SD = SPAN destination mode, NP = proxy N mode



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

The Fibre Channel standards as defined by the ANSI T11 committee allow for up to 239 Fibre Channel domains per fabric or VSAN. However, the original storage manufacturer has only qualified up to 70 domains per fabric or VSAN.

Each Fibre Channel switch is identified by a single domain ID, so effectively there can be no more than 40 switches that are connected together.

Blade switches and top-of-rack access layer switches will also consume a domain ID, which will limit the number of domains that can be deployed in data centers.

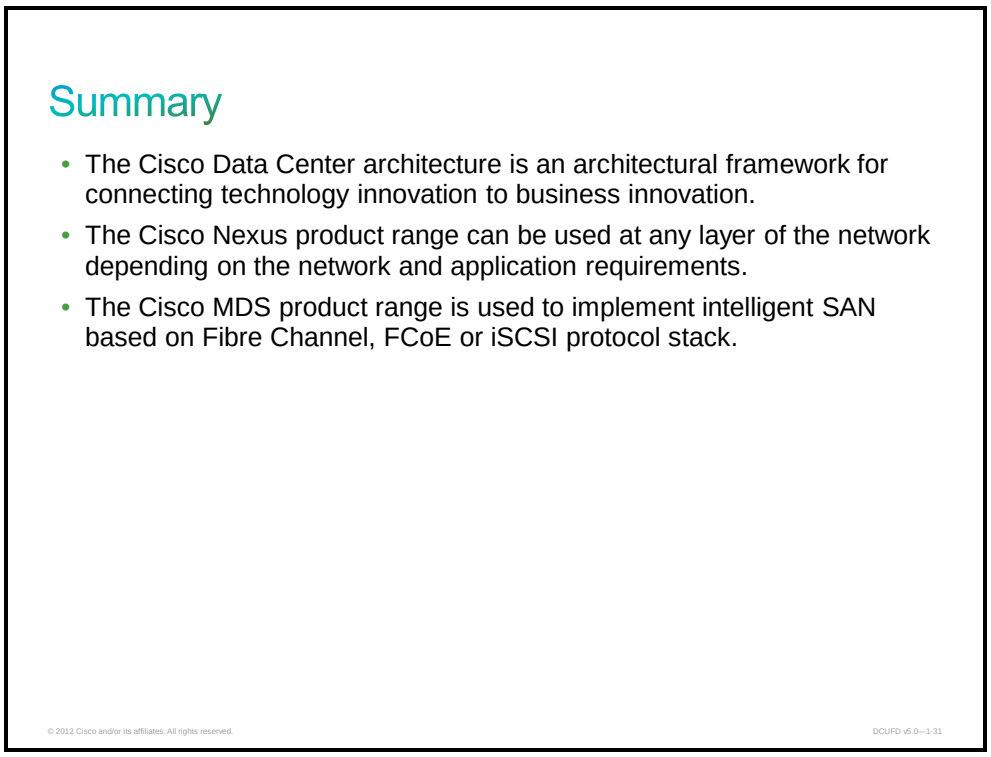
The Cisco NPV addresses the increase in the number of domain IDs that are needed to deploy many ports by making a fabric or module switch appear as a host to the core Fibre Channel switch, and as a Fibre Channel switch to the servers in the fabric or blade switch. Cisco NPV aggregates multiple locally connected N Ports into one or more external N-Port links, which share the domain ID of the NPV core switch among multiple NPV switches. Cisco NPV also allows multiple devices to attach to the same port on the NPV core switch, and it reduces the need for more ports on the core.

https://t.me/learningnets

certcollecion.net

Summary

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



 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

Lesson 3

Designing the Cisco Data Center Solution

Overview

In this lesson, you will gain insight into the data center solution design process. This lesson provides an overview of how a data center solution is designed and the documentation that is necessary. There is a difference in the design phase for new scenarios and designs that involve an existing production environment and usually a migration from the old to the new environment.

Objectives

Upon completing this lesson, you will be able to define the tasks and phases of the design process for the Cisco Data Center solution. This ability includes being able to meet these objectives:

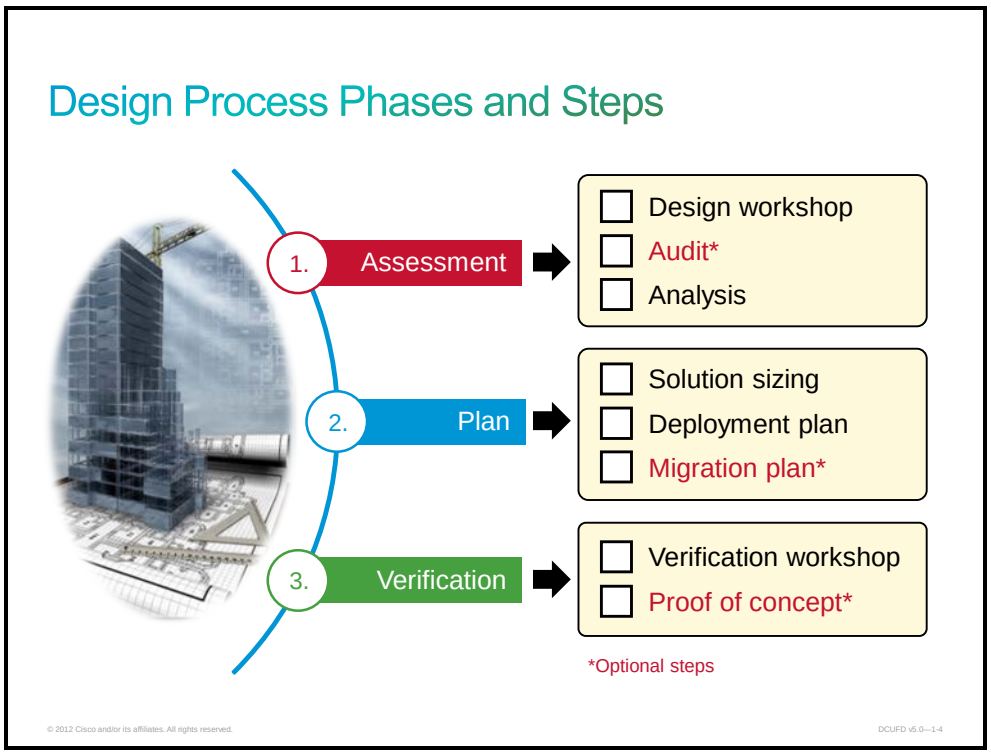
- Describe the design process for the Cisco Data Center solution
- Assess the deliverables of the Cisco Data Center solution
- Describe Cisco Validated Designs

https://t.me/learningnets

certcollecion.net

Design Process

This topic describes the design process for the Cisco Data Center solution.



To design a solution that meets customer needs, it is important to identify the organizational goals, organizational constraints, technical goals, and technical constraints. In general, the design process can be divided into three major phases:

1. **Assessment phase:** This phase is vital for the project to be successful and to meet the customer needs and expectations. In this phase, all information that is relevant for the design has to be collected.
2. **Plan phase:** In this phase, the solution designer creates the solution architecture by using the assessment phase results as input data.
3. **Verification phase:** To ensure that the designed solution architecture does meet the customer expectations, the solution should be verified and confirmed by the customer.

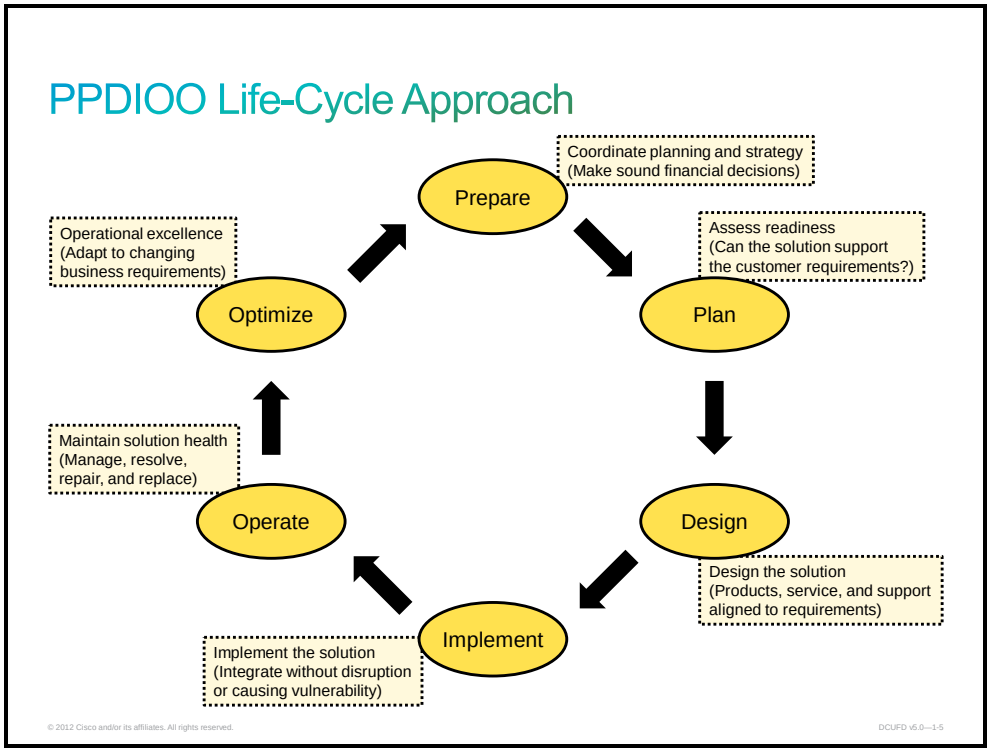
Each phase of the design process has steps that need to be taken in order to complete the phase. Some of the steps are mandatory and some are optional. The decision about which steps are necessary is governed by the customer requirements and the type of the project (for example, new deployment versus migration).

To track the design process progress and completed and open actions, the checklist in the figure can aid the effort.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net



Cisco has formalized the life cycle of a solution into six phases: Prepare, Plan, Design, Implement, Operate, and Optimize (PPDIOO). For the design of the Cisco Data Center solution, the first three phases are used.

The PPDIOO solution life-cycle approach reflects the life-cycle phases of a standard solution. The PPDIOO phases are as follows:

- **Prepare:** The prepare phase involves establishing the organizational requirements, developing a solution strategy, and proposing a high-level conceptual architecture that identifies technologies that can best support the architecture. The prepare phase can establish a financial justification for the solution strategy by assessing the business case for the proposed architecture.
- **Plan:** The plan phase involves identifying initial solution requirements based on goals, facilities, user needs, and so on. The plan phase involves characterizing sites, assessing any existing environment, and performing a gap analysis to determine whether the existing system infrastructure, sites, and operational environment are able to support the proposed system. A project plan is useful to help manage the tasks, responsibilities, critical milestones, and resources that are required to implement changes to the solution. The project plan should align with the scope, cost, and resource parameters established in the original business requirements.
- **Design:** The initial requirements that were derived in the planning phase lead the activities of the solution design specialists. The solution design specification is a comprehensive detailed design that meets current business and technical requirements and incorporates specifications to support availability, reliability, security, scalability, and performance. The design specification is the basis for the implementation activities.
- **Implement:** After the design has been approved, implementation (and verification) begins. The solution is built or additional components are incorporated according to the design specifications, with the goal of integrating devices without disrupting the existing environment or creating points of vulnerability.

https://t.me/learningnets

certcollecion.net

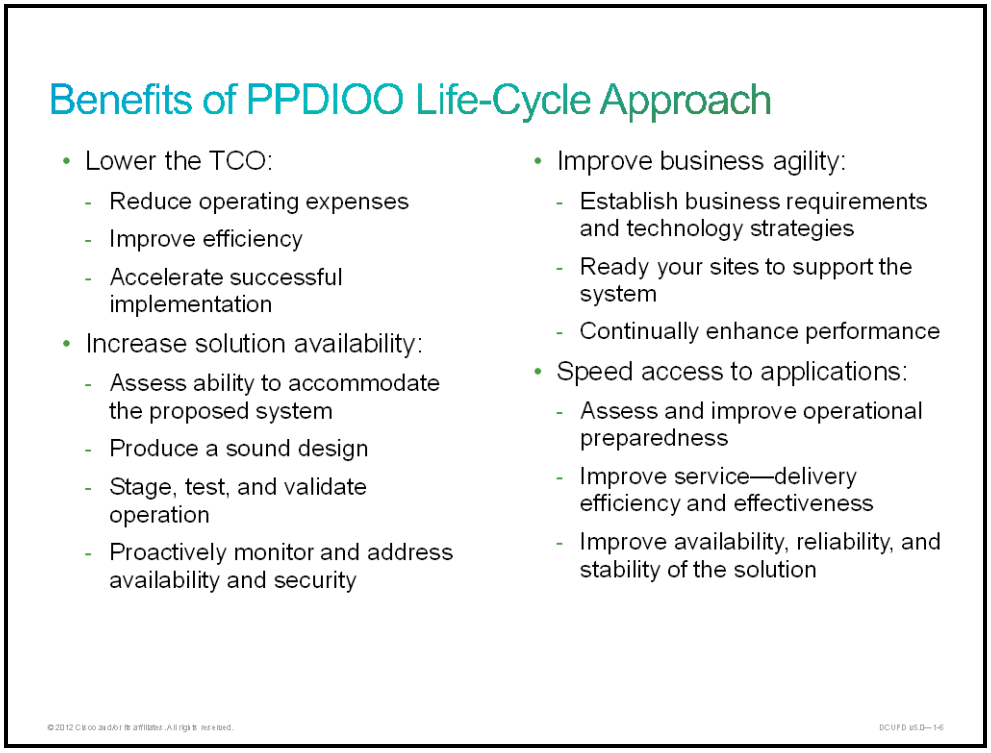
- **Operate:** Operation is the final test of the appropriateness of the design. The operational phase involves maintaining solution health through day-to-day operations, including maintaining high availability and reducing expenses. The fault detection, correction, and performance monitoring that occur in daily operations provide initial data for the optimization phase.
- **Optimize:** The optimization phase involves proactive management of the solution. The goal of proactive management is to identify and resolve issues before they affect the organization. Reactive fault detection and correction (troubleshooting) are needed when proactive management cannot predict and mitigate failures. In the PPDIOO process, the optimization phase may prompt a network redesign if too many solution problems and errors arise, if performance does not meet expectations, or if new applications are identified to support organizational and technical requirements.

Note	Although design is listed as one of the six PPDIOO phases, some design elements may be present in all the other phases.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net



The solution life-cycle approach provides four main benefits:

- Lowering the total cost of ownership (TCO)
- Increasing solution availability
- Improving business agility
- Speeding access to applications and services

The TCO is lowered by these actions:

- Identifying and validating technology requirements
- Planning for infrastructure changes and resource requirements
- Developing a sound solution design aligned with technical requirements and business goals
- Accelerating successful implementation
- Improving the efficiency of your solution and of the staff supporting it
- Reducing operating expenses by improving the efficiency of operation processes and tools

Solution availability is increased by these actions:

- Assessing the security state of the solution and its ability to support the proposed design
- Specifying the correct set of hardware and software releases and keeping them operational and current
- Producing a sound operations design and validating solution operation
- Staging and testing the proposed system before deployment
- Improving staff skills
- Proactively monitoring the system and assessing availability trends and alerts
- Proactively identifying security breaches and defining remediation plans

https://t.me/learningnets

certcollecion.net

Business agility is improved by these actions:

- Establishing business requirements and technology strategies
- Readying sites to support the system that you want to implement
- Integrating technical requirements and business goals into a detailed design and demonstrating that the solution is functioning as specified
- Expertly installing, configuring, and integrating system components
- Continually enhancing performance

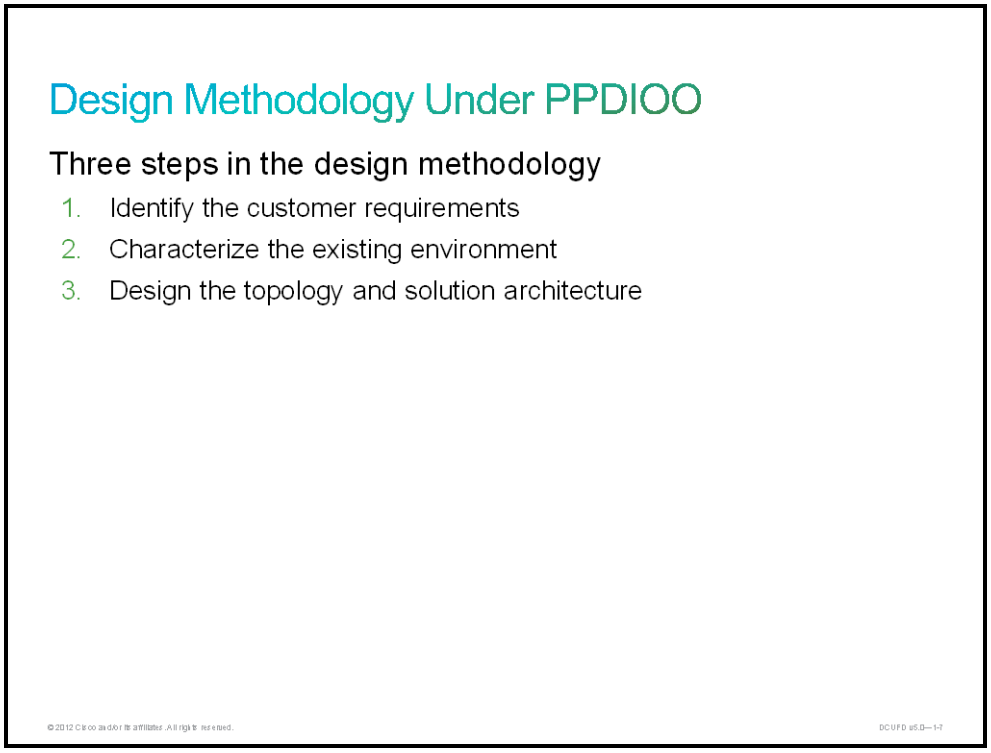
Access to applications and services is accelerated by these actions:

- Assessing and improving operational preparedness to support current and planned solution technologies and services
- Improving service-delivery efficiency and effectiveness by increasing availability, resource capacity, and performance
- Improving the availability, reliability, and stability of the solution and the applications running on it
- Managing and resolving problems affecting your system and keeping software applications current

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net



The design methodology under PPDIOO consists of three basic steps:

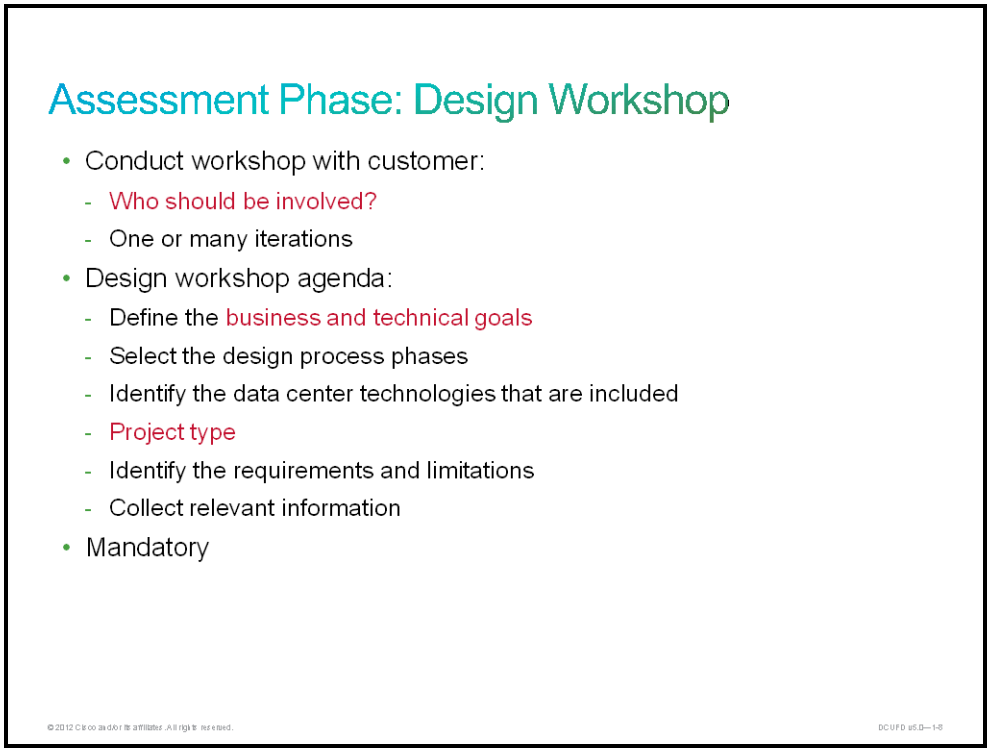
Step 1 **Identify customer requirements:** In this step, important decision makers identify the initial requirements. Based on these requirements, a high-level conceptual architecture is proposed. This step is typically done within the PPDIOO prepare phase.

Step 2 **Characterize the existing network and sites:** The plan phase involves characterizing sites and assessing any existing networks and performing a gap analysis to determine whether the existing system infrastructure, sites, and operational environment can support the proposed system. Characterization of the existing environment includes existing environment audit and analysis. During the audit, the existing environment is thoroughly checked for integrity and quality. During the analysis, environment behavior (traffic, congestion, and so on) is analyzed. This investigation is typically done within the PPDIOO plan phase.

Step 3 **Design the network topology and solutions:** In this step, you develop the detailed design. Decisions on solution infrastructure, intelligent services, and solutions are made. You may also build a pilot or prototype solution to verify the design. You also write a detailed design document.

https://t.me/learningnets

certcollecion.net



The first action of the design process and the first step of the assessment phase is the design workshop. The workshop has to be conducted with proper customer IT personnel and can take several iterations in order to collect relevant and valid information. In the design workshop, a draft high-level architecture may already be defined.

The high-level agenda of the design workshop should include these tasks:

- **Define the business goals:** This step is important for several reasons. First, you should ensure that the project follows customer business goals, which will help you ensure that the project is successful. With the list of goals, the solution designers can then learn and write down what the customer wants to achieve with the project and what the customer expects from the project.
- **Define the technical goals:** This step ensures that the project also follows customer technical goals and expectations and thus likewise ensures that the project is successful. With this information, the solution designer will know the technical requirements of the project.
- **Identify the data center technologies:** This task is used to clarify which data center technologies are covered by the project and is the basis for how the experts determine what is needed for the solution design.
- **Define the project type:** There are two main types of projects. They are new deployments or the migration of existing solutions.
- **Identify the requirements and limitations:** The requirements and limitations are the details that significantly govern the equipment selection, the connectivity that is used, the integration level, and the equipment configuration details. For migration projects, this step is the first part of identifying relevant requirements and limitations. The second part is the audit of the existing environment with proper reconnaissance and analysis tools.

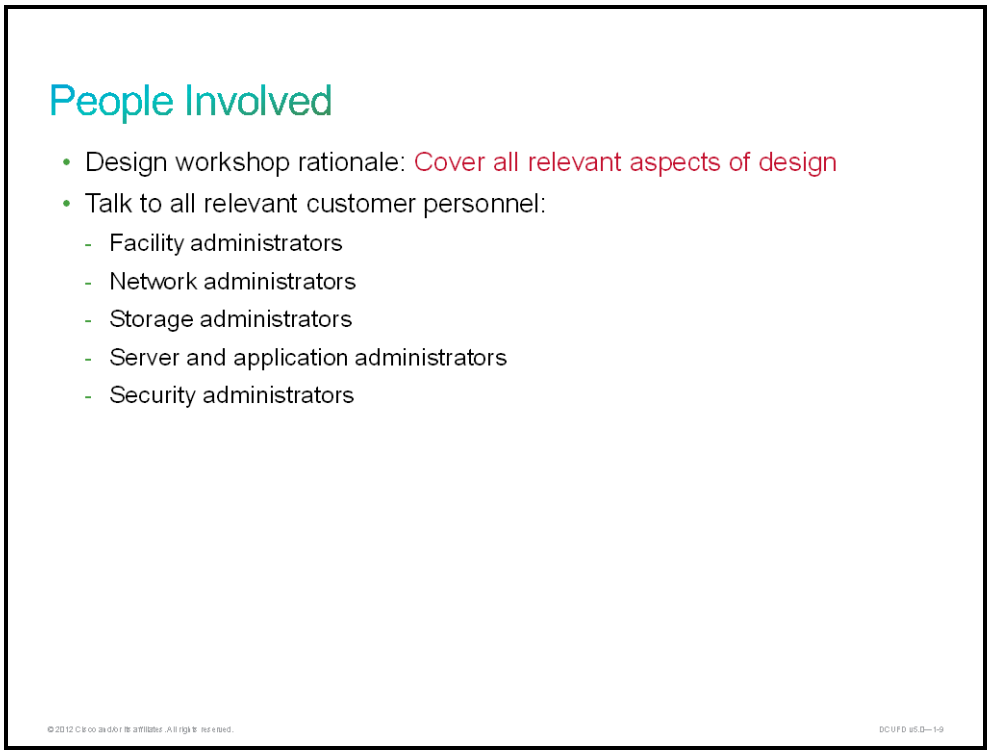
The workshop can be conducted in person or it can be done virtually by using Cisco WebEx or a Cisco TelePresence solution.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

The design workshop is a mandatory step of the assessment phase because without it, there is no relevant information with which the design can be created.



It is very important to gather all the relevant people in the design workshop to cover all the aspects of the solution (the design workshop can be a multiday event).

The Cisco Unified Computing System (UCS) solution is effectively part of the data center, and as such, the system must comply with all data center policies and demands. The following customer personnel must attend the workshop (or should at least provide information that is requested by the solution designer):

- **Facility administrators:** They are in charge of the physical facility and have the relevant information about environmental conditions like available power, cooling capacity, available space and floor loading, cabling, physical security, and so on. This information is important for the physical deployment design and can also influence the equipment selection.
- **Network administrators:** They ensure that the network properly connects all the bits and pieces of the data center and thus also the equipment of the future Cisco UCS solution. It is vital to receive all the information about the network: throughput, port and connector types, Layer 2 and Layer 3 topologies, high-availability mechanisms, addressing, and so on. The network administrators may report certain requirements for the solution.
- **Storage administrators:** Here, the relevant information encompasses storage capacity (available and used), storage design and redundancy mechanisms (logical unit numbers [LUNs], Redundant Array of Independent Disks [RAID] groups, service processor ports, and failover), storage access speed, type (Fibre Channel, Internet Small Computer Systems Interface [iSCSI], Network File System [NFS]), replication policy and access security, and so on.

https://t.me/learningnets

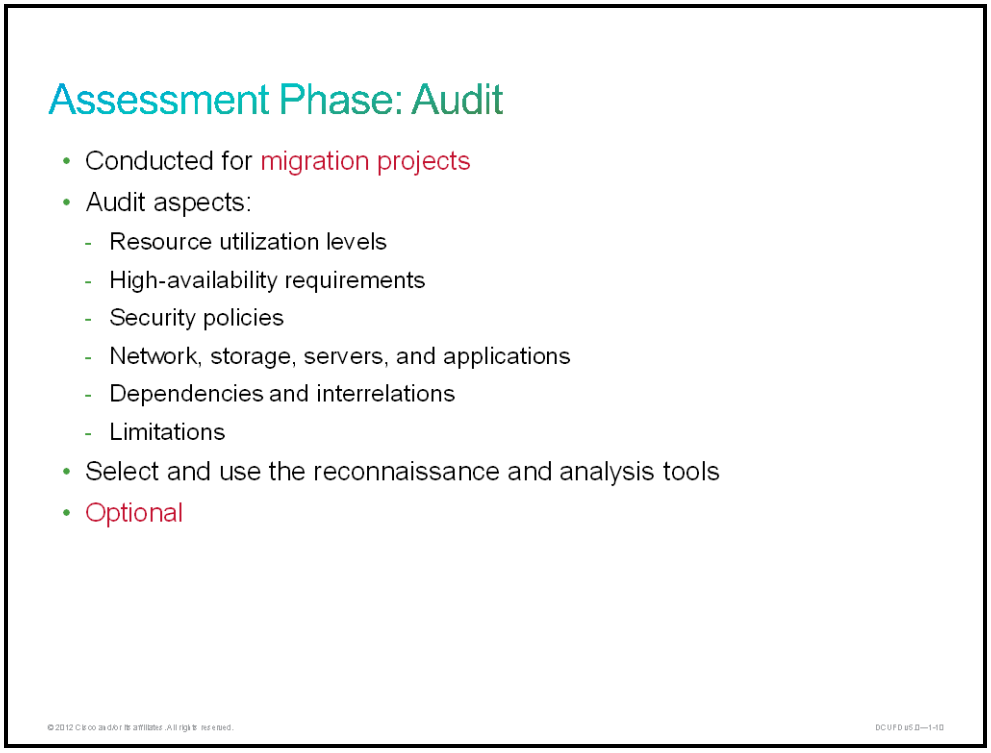
certcollecion.net

- **Server and application administrators:** They know the details for the server requirements, operating systems, and application dependencies and interrelations. The solution designer learns which operating systems and versions are or will be used, what the requirements of the operating systems are from the connectivity perspective (one network interface card [NIC], two NICs, NIC teaming, and so on). The designer also learns which applications will be deployed on which operating systems and what the application requirements will be (connectivity, high availability, traffic throughput, typical memory and CPU utilization, and so on).
- **Security administrators:** The solution limitations can also be known from the customer security requirements (for example, the need to use separate physical VMware vSphere hosts for a demilitarized zone [DMZ] and private segments). The security policy also defines the control of equipment administrative access and allowed and restricted services (for example, Telnet versus Secure Shell [SSH]), and so on.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net



The audit step of the assessment should typically be undertaken for migration projects. It is not necessary, but it is strongly advised in order to audit the existing environment.

For proper design, it is of the utmost importance to have the relevant information upon which the design is based:

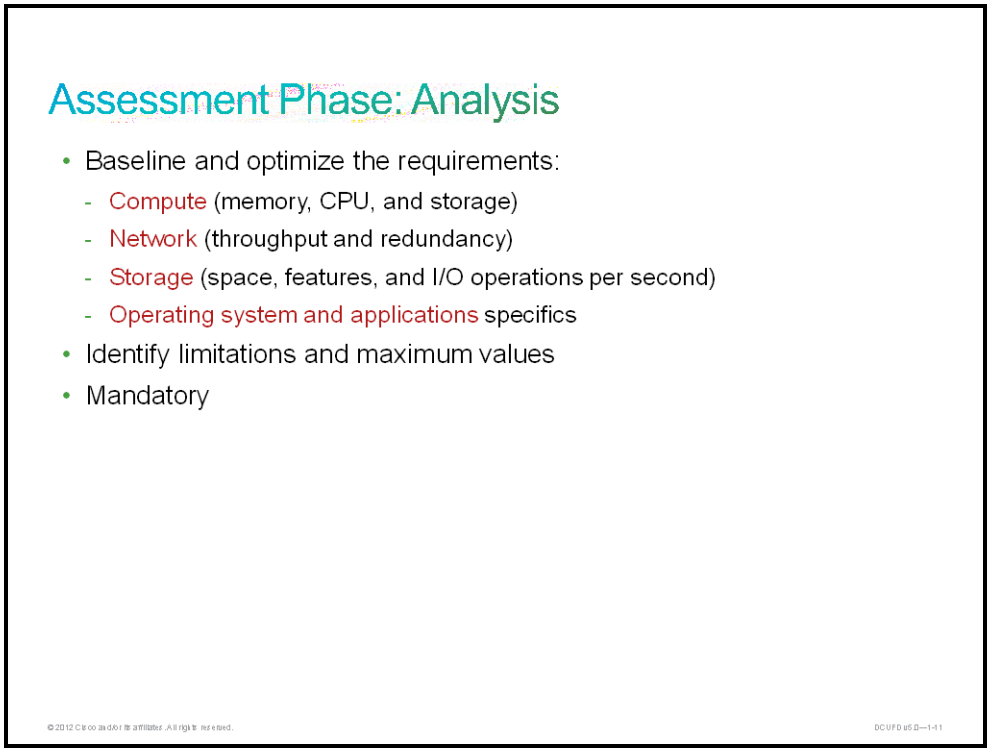
- Memory and CPU resources
- Storage space
- Inventory details
- Historical growth report
- Security policies that are in place
- High-availability mechanisms
- Dependencies between the data center elements (that is, applications, operating system, server, storage, and so on)
- The limitations of the current infrastructure

From the description of the audit, it is clear that some information should be collected over a longer time to be relevant (for example, the levels of server memory and CPU utilization that are measured over a weekend are significantly lower than during weekdays). Other details can be gathered by inspecting the equipment configuration (for example, administrative access, logging, Simple Network Management Protocol (SNMP) management, and so on).

Information can be collected with the various reconnaissance and analysis tools that are available from different vendors. If the project involves a migration to a VMware vSphere environment from physical servers, the VMware Capacity Planner will help the designer collect information about the servers and it can even suggest the type of servers that are appropriate for the new design (regarding processor power, memory size, and so on).

https://t.me/learningnets

certcollecion.net



The analysis is the last part of the assessment phase. The solution designer must review all the collected information and then select only the important details.

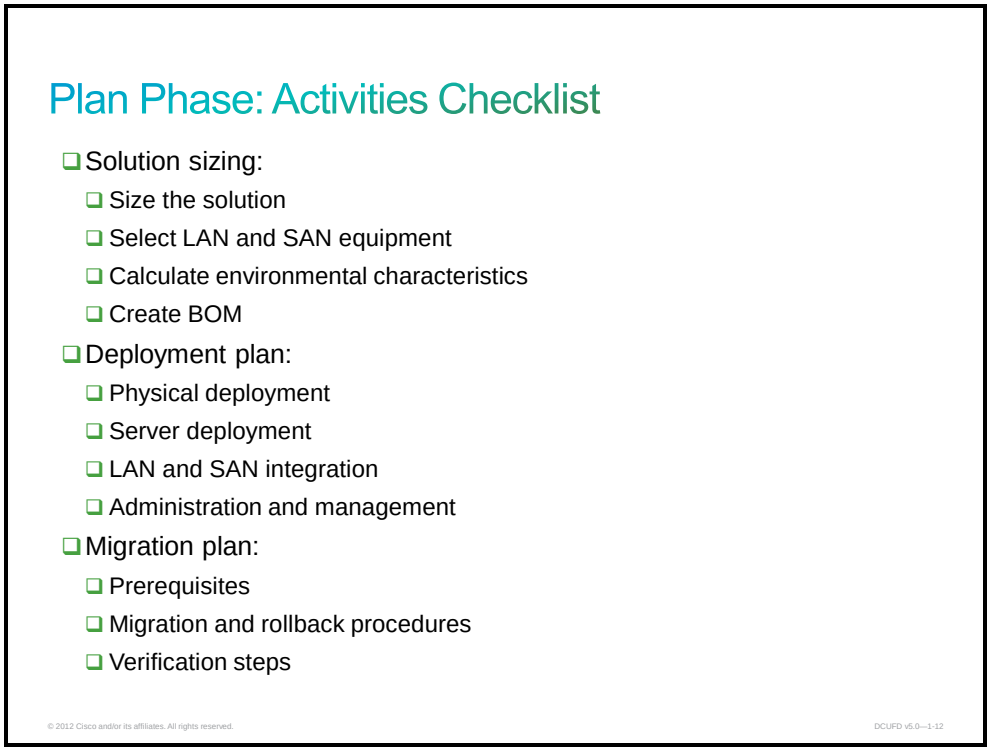
The designer must baseline and optimize the requirements, which can then be directly translated into the proper equipment, software, and configurations.

The analysis is mandatory for creating a designed solution that will meet project goals and customer expectations.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net



Once the assessment phase is completed and the solution designer has the analysis results, the design or plan phase can commence.

This phase (like the assessment phase) contains several steps and substeps. Some steps are mandatory and some are optional. There are three major steps:

- **Solution sizing:** In this step, the hardware and software that are used will be defined.
  - LAN and SAN equipment that is required for connecting the system has to be selected. The equipment can be small form-factor pluggable (SFP) modules, a new module, or even Cisco Nexus and Multilayer Director Switch (MDS) switches and licenses.
  - Once the equipment is selected, the environmental requirements need to be determined by using the Cisco Power Calculator. You will need to calculate the power, cooling, and weight measurements for the Cisco UCS, Nexus, MDS, and other devices.
  - Last but not least, the Bill of Materials (BOM), which is a detailed list of the equipment parts, needs to be created. The BOM includes not only the Cisco UCS and Nexus products, but also all the necessary patch cables, power inlets, and so on.
- **Deployment plan:** This step can be divided into the following substeps:
  - The physical deployment plan details where and how the equipment will be placed into the racks for racking and stacking.
  - The server deployment plan details the server infrastructure configuration, such as the LAN and SAN access layer configuration, VLANs and VSANs, and port connectivity. This plan also details MAC, world wide name (WWN), and universally unique identifier (UUID) addressing, and management access, firmware versions, and high-availability settings. All Cisco UCS details are defined from a single management point in the Cisco UCS Manager.

https://t.me/learningnets

certcollecion.net

- The LAN and SAN integration plan details the physical connectivity and configuration of core data center devices (the Cisco Nexus and MDS switches, VLAN and VSAN configuration on the core side, and the high-availability settings).
- The administration and management plan details how the new solution will be managed and how it integrates into the existing management infrastructure (when present).

- **Migration plan:** Applicable for migration projects, this plan details when, how, and with which technologies the migration from an existing solution to a new deployment will be performed. A vital part of the migration plan is the series of verification steps that confirm or disprove the successfulness of migration. Equally important (although hopefully not used) are the rollback procedures that should be used in case of failures or problems during migration.

Different deployments have different requirements and thus different designs. Typical solutions to common requirements are described in the Cisco Validated Designs (for example, Citrix XenDesktop with VMware and Cisco UCS, an Oracle database and the Cisco UCS, and so on).

Note Details about such designs are discussed later in the course.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

Verification Phase

- Verification workshop with customer is mandatory
  - Confirm and sign off on the solution design
- Proof of concept is optional
  - Implement partial solution in a lab environment
  - Confirm and verify the designed solution



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

DCUFD v5.0—1-13

Once the design phase is completed, the solution must be verified and approved by the customer. This approval is typically received by conducting a verification workshop with the customer personnel who are responsible for the project. The customer also receives complete information about the designed solution.

The second step of the verification phase can be the proof of concept, which is how the customer and designer can confirm that the proposed solution meets the expected goals. The proof of concept is typically a smaller set of the proposed solution that encompasses all the vital and necessary components to confirm the proper operation.

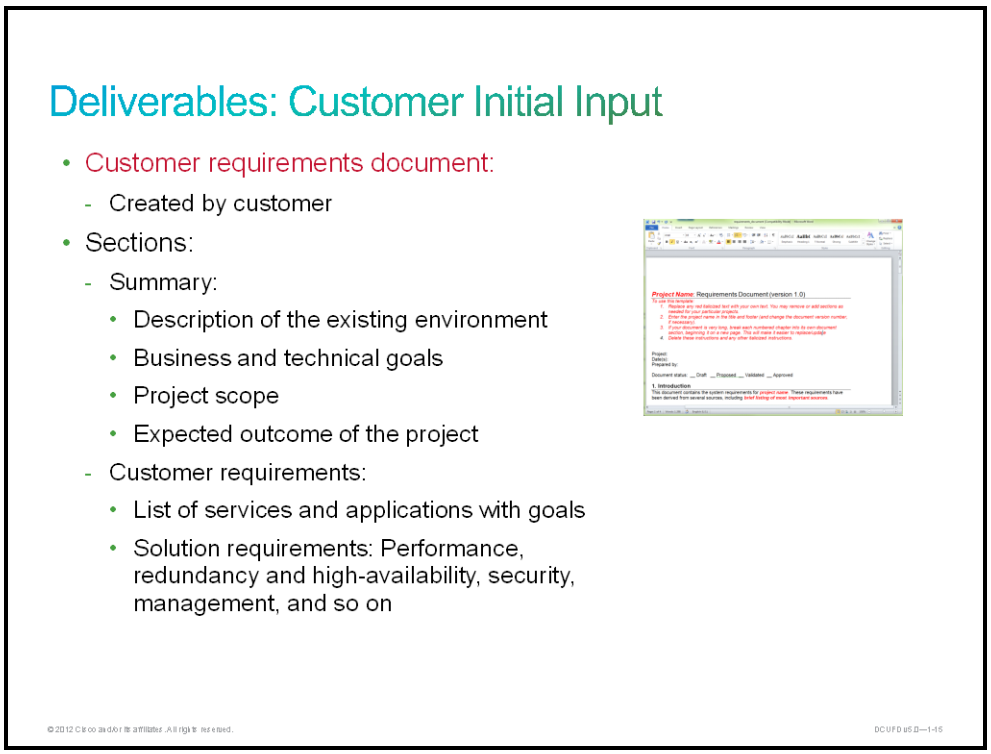
The solution designer must define the subset of the designed solution that needs to be tested and must conduct the necessary tests with expected results.

https://t.me/learningnets

certcollecion.net

Design Deliverables

This topic describes how to assess the deliverables of the Cisco Data Center solution.



Every project should start with a clear understanding of the requirements of the customer. Thus the customer requirements document (CRD) should be used to detail the customer requirements for a project for which a solution will be proposed. It is required to be completed upon the request of the department or project leader from the customer side.

The following sections should be part of the CRD:

- **Existing environment:** This section describes the current customer environment.
- **Expected outcome:** This section provides an overview of the intentions and future direction, and summarizes the services and applications that the customer intends to introduce. This section also defines the strategic impact of this project to the customer (for example, is the solution required to solve an important issue, make the customer more profitable, and give the customer a competitive edge?).
- **Project scope:** This section defines the scope of the project regarding the design (for example, which data center components are involved, which technologies should be covered, and so on).
- **List of services and applications with goals:** This section provides a list of the objectives and requirements for this service, that is, details about the type of services the customer plans to offer and introduce with the proposed solution. Apart from connectivity, security, and so on, the list includes details about the applications and services planned to be deployed.
- **Solution requirements:** This section defines the following characteristics for the solution as a whole as well as for the individual parts:
  - Requirements concerning system availability, behavior under failure scenarios, and service restoration

 © 2012 Cisco Systems, Inc.

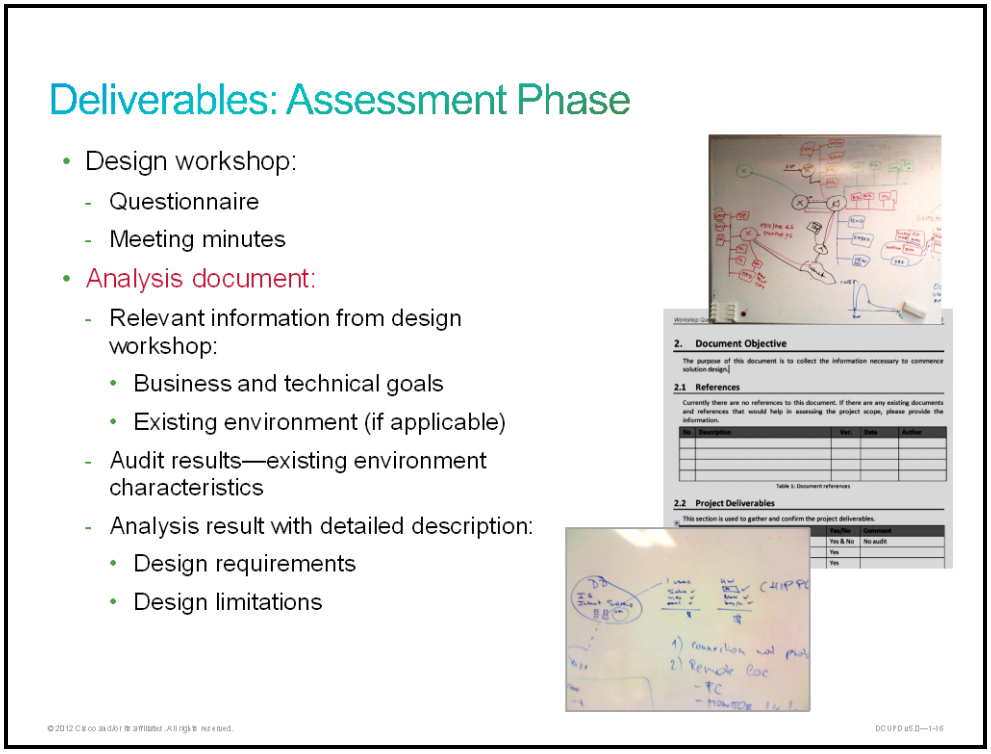
https://t.me/learningnets

certcollecion.net

- All performance requirements
- All security requirements
- All critical features required in order to provide this service or application, including those that are not yet implemented
- All aspects of solution management features like service fulfillment (service provisioning and configuration management), service assurance (fault management and performance management), as well as billing and accounting

It is also advisable that the CRD include the high-level timelines of the project so that the solution designer can plan accordingly.

The CRD thus clearly defines what the customer wants from the solution and is also the basis and input information for the assessment phase.



Each phase of the design process should result in documents that are necessary not only for tracking the efforts of the design team, but also for presenting the results and progress to the customer.

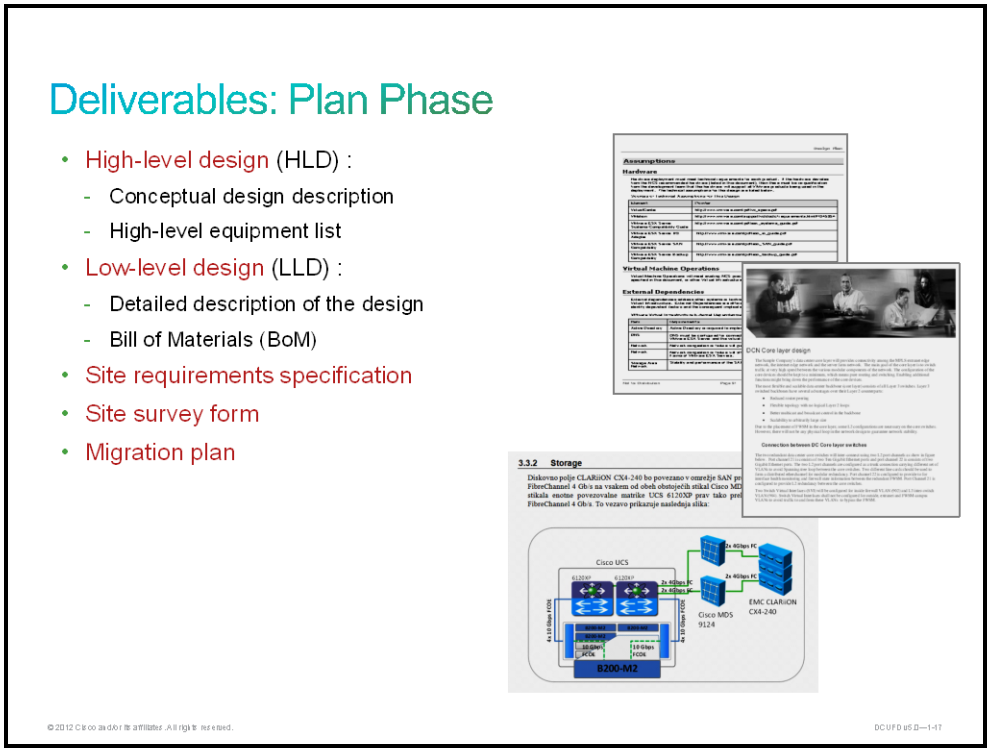
The supporting documentation for the assessment phase can include the following:

- **Questionnaire:** This questionnaire can be distributed to the customer personnel in order to prepare for the design workshop or to provide relevant information in written form when verbal communication is not possible.
- **Meeting minutes:** This document contains the relevant information from the design workshop.

The assessment phase should finally result in the analysis document, which must include all the information that is gathered in the assessment phase.

https://t.me/learningnets

certcollecion.net



The design phase is the most document-intensive phase. It should result in the following documentation:

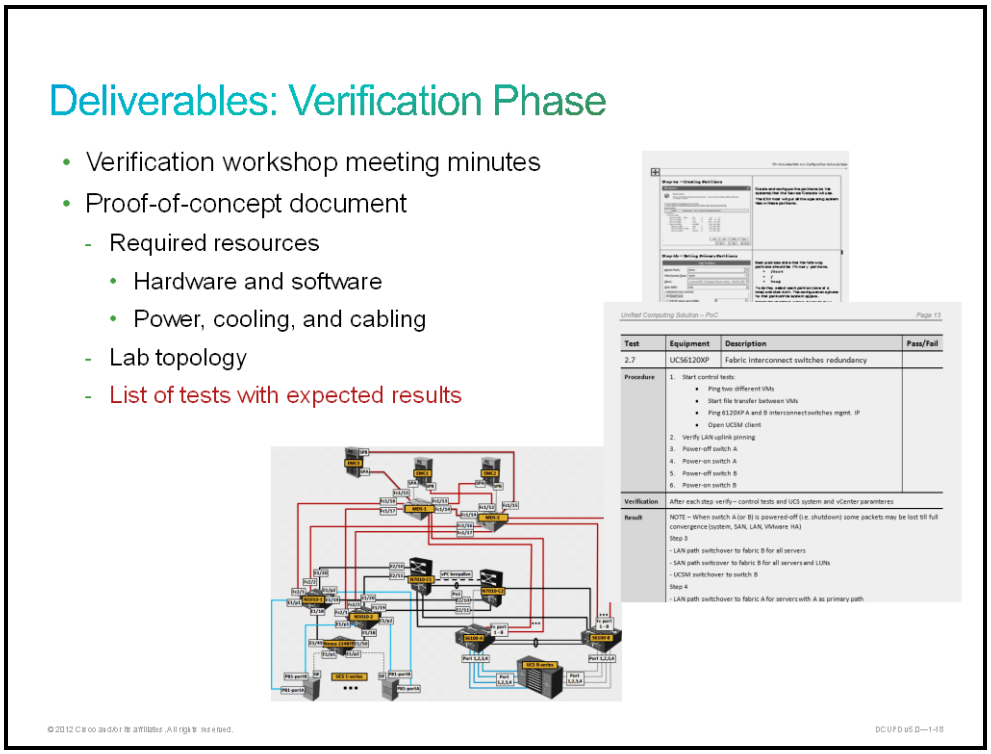
- **High-level design (HLD):** This document describes the conceptual design of the solution, such as the solution components, the equipment to be used (not detailed information), how the high-availability mechanisms work, how the business continuance is achieved, and so on.
- **Low-level design (LLD) (also known as a detailed design):** This document describes the design in detail, such as the detailed list of equipment, the plan of how the devices will be connected physically, the plan of how the devices will be deployed in the racks, as well as information about the relevant configurations, addressing, address pools and naming conventions, resource pools, management IP addressing, service profiles, VLANs, and VSANs.
- **Site requirements specification:** This document (or more than one document when the solution applies to more than one facility) will specify the equipment environmental characteristics, such as power, cooling capacity, weight, and cabling.
- **Site survey form:** This document (or more than one document when the solution applies to more than one facility) is used by the engineers or technicians to conduct the survey of a facility in order to determine the environmental specifications.
- **Migration plan:** This document is necessary when the project is a migration and it must have at least the following sections:
  - **Required resources:** Specifies the resources that are necessary to conduct the migration. These resources can include, for example, extra space on the storage, extra Ethernet ports to connect new equipment before the old equipment is decommissioned, or extra staff or even external specialists.
  - **Migration procedures:** Specifies the actions for conducting the migration (in the correct order) with verification tests and expected results.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

— **Rollback procedures:** Specifies the actions that are necessary to revert to a previous state if there are problems during the migration.



Because the first step of the verification phase is the verification workshop, meeting minutes should be taken in order to track the workshop.

If the customer confirms that the solution design is approved, the customer must sign off on the solution.

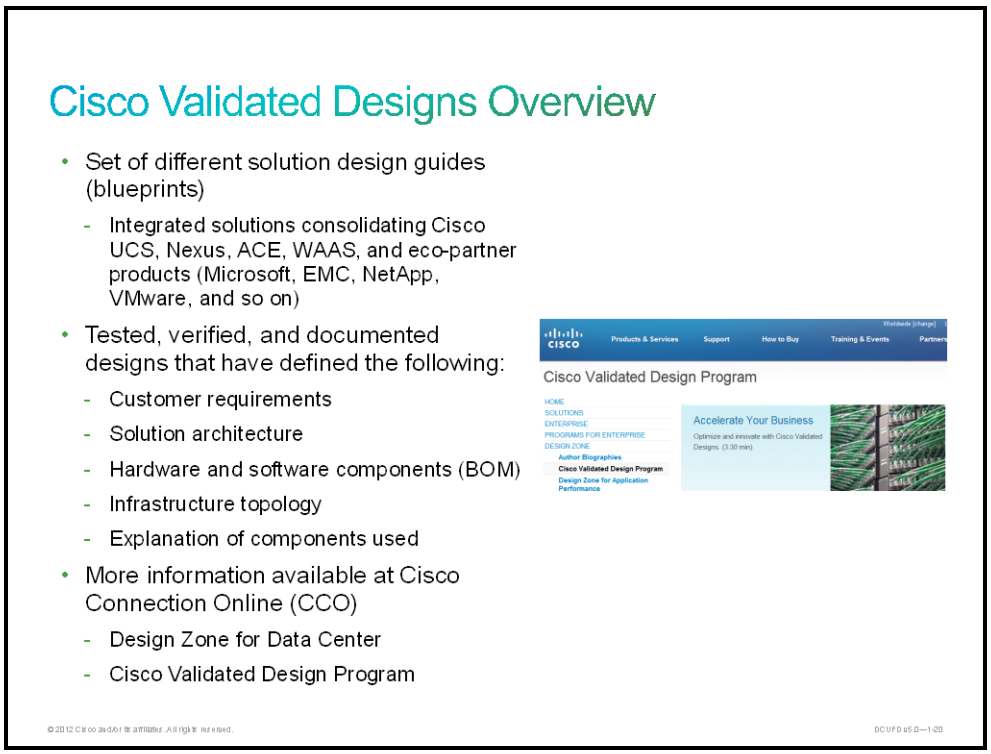
Second, when a proof of concept is conducted, the proof-of-concept document should be produced. The document is a subset of the detailed design document and it is for the equipment that will be used in the proof of concept. Apart from that, the document must specify what resources are required to conduct the proof of concept (not only the equipment but also the environmental requirements), and it should list the tests and the expected results with which the solution is verified.

https://t.me/learningnets

certcollecion.net

Cisco Validated Designs

This topic describes Cisco Validated Designs.



Cisco Validated Designs consist of systems and solutions that are designed, tested, and documented to facilitate and improve customer deployments. These designs incorporate a wide range of technologies and products into a portfolio of solutions that have been developed to address the business needs of our customers. Cisco Validated Designs are organized by solution areas.

Cisco UCS-based validated designs are blueprints that incorporate not only Cisco UCS but also other Cisco Data Center products and technologies along with applications of various eco-partners (Microsoft, EMC, NetApp, VMware, and so on).

The individual blueprint covers the following aspects:

- Solution requirements from an application standpoint
- Overall solution architecture with all the components that fit together
- Required hardware and software BOM
- Topology layout
- Description of the components used and their functionalities

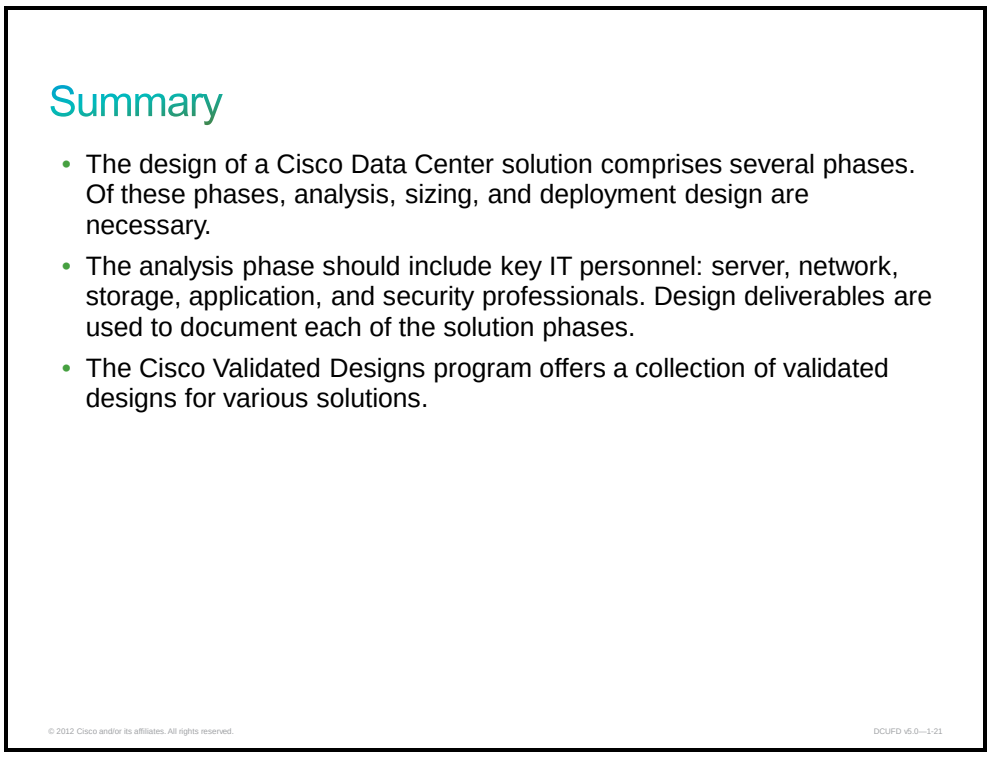
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

Summary

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



https://t.me/learningnets

certcollecion.net

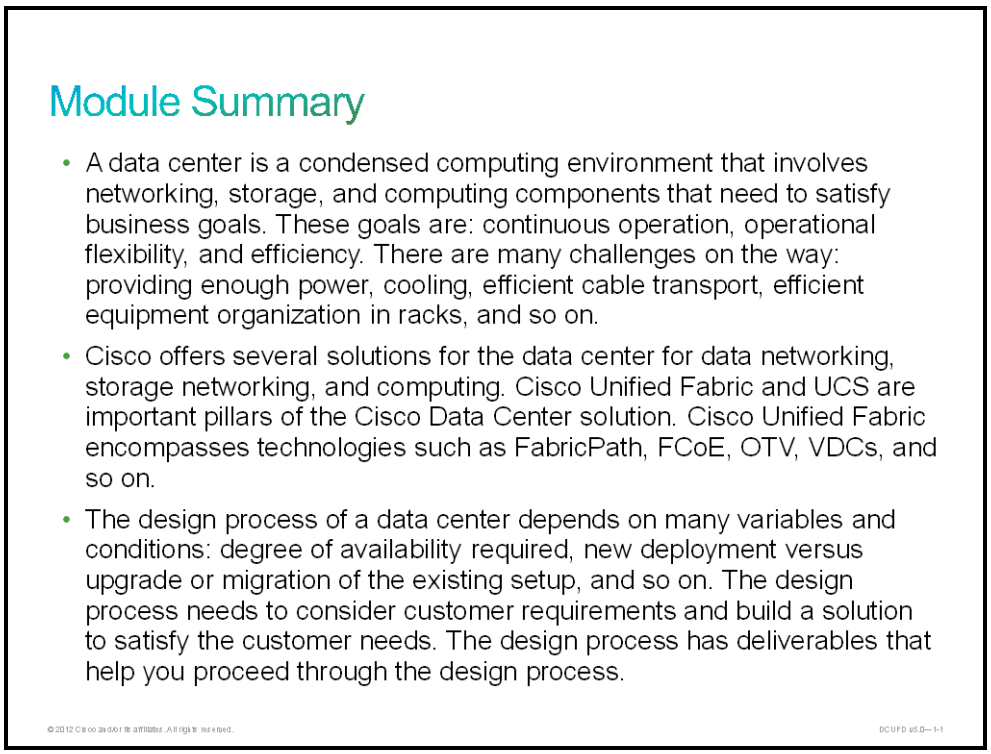
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

Module Summary

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



In this module, you learned which technologies the data center encompasses, and which aspects are covered by Cisco solutions.

Data centers are very complex environments that require collaboration of experts of various technology areas (applications, server and networking hardware, and storage). Interdisciplinary knowledge needs to be gathered, including power delivery, cooling, construction, physical access security, surveillance, accounting, regulatory compliance, and so on.

https://t.me/learningnets

certcollecion.net

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

Module Self-Check

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

Q1) Which two options are important business drivers for data centers? (Choose two.) (Source: Defining the Data Center)

A) global availability
B) global warming
C) reduced communication latency
D) fast application deployment

Q2) Which two options are the main operational limitations of data centers? (Choose two.) (Source: Defining the Data Center)

A) server consolidation
B) power and cooling
C) rack space
D) rack weight

Q3) Where do you install cables in data center rooms? (Source: Identifying the Cisco Data Center Solution)

Q4) What is the thermal control model called? (Source: Identifying the Cisco Data Center Solution)

A) thermomix
B) hot aisle, cold aisle
C) hotspot
D) British thermal Unit

Q5) Which three options are phases of the data center design process? (Choose three.) (Source: Designing the Cisco Data Center Solution)

A) assessment phase
B) spin-out phase
C) plan phase
D) verification phase

Q6) Which three cloud deployments are based on the NIST classification? (Choose three.) (Source: Defining the Data Center)

A) virtual cloud
B) private cloud
C) open cloud
D) public cloud
E) cloud as a service

Q7) Which two mechanisms allow virtualization of the network and IP services? (Choose two.) (Source: Defining the Data Center)

A) VRF
B) security context
C) service profile
D) hypervisor

https://t.me/learningnets

certcollecion.net

Q8) Identify the three important Cisco technologies in the data center. (Choose three.) (Source: Identifying the Cisco Data Center Solution)

A) FEX-Link
B) OTV
C) VLAN
D) VMotion
E) VDC

Q9) What are the three capabilities of the Cisco MDS 9500 platform? (Choose three.) (Source: Identifying the Cisco Data Center Solution)

A) virtualized NAS
B) FCoE
C) Fibre Channel
D) FCIP
E) serial attached SCSI

Q10) Where can you find design best practices and design recommendations when designing data center networks? (Source: Designing the Cisco Data Center Solution)

A) Cisco Best Practices Program
B) Cisco Validated Design Program
C) Cisco Advanced Services
D) Cisco.com

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

Module Self-Check Answer Key

Q1) A, D

Q2) B, C

Q3) Cables are installed under the floor or under the ceiling, depending of the design of the room.

Q4) B

Q5) A, C, D

Q6) B, C, D

Q7) A, B

Q8) A, B, E

Q9) B, C, D

Q10) B

https://t.me/learningnets

certcollecion.net

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

Module 2

Data Center Technologies

Overview

In this module, you will learn about modern data center technologies. The technologies include Layer 2 multipathing, multilink aggregation, and virtualization. These technologies allow for optimum use of data center resources—links that are all utilized and devices that are virtualized with increased utilization efficiency.

Module Objectives

Upon completing this module, you will be able to provide a comprehensive and detailed overview of technologies that are used in data centers, and describe scalability implications and their possible use in cloud environments. This ability includes being able to meet these objectives:

- Describe and design Layer 2 and Layer 3 switched networks, as well as provide for high availability and customer separation
- Identify and design data center component virtualization technologies, present the limitations, and outline best practices and validated designs
- Design data centers using multipathing technologies, such as vPC, MEC, and Cisco FabricPath, all without using STP

https://t.me/learningnets

certcollecion.net

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

Lesson 1

Designing Layer 2 and Layer 3 Switching

Overview

This lesson presents various technologies that are essential for data centers. These technologies include packet switching and routing, hardware-based switching, and scalable routing protocols that are used in data centers.

Objectives

Upon completing this lesson, you will be able to describe and design Layer 2 and Layer 3 switched networks, as well as provide for high availability and customer separation. This ability includes being able to meet these objectives:

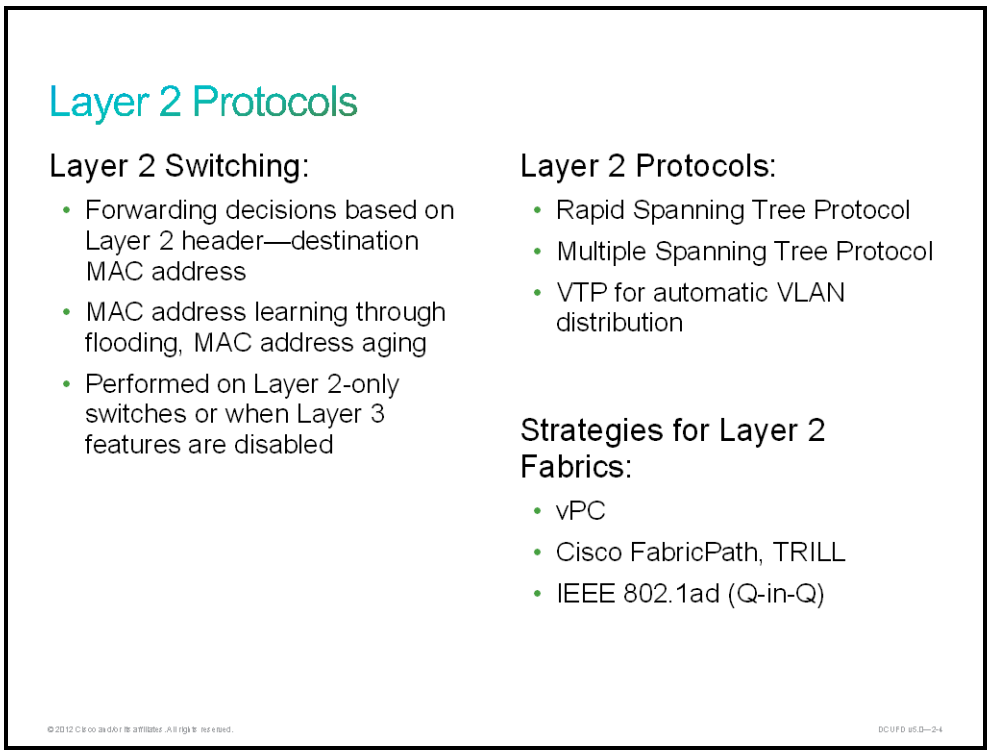
- Understand and explain hardware-forwarding architectures
- Describe IP addressing considerations and IP routing technologies

https://t.me/learningnets

certcollecion.net

Forwarding Architectures

This topic describes hardware-forwarding architectures.



Historically, Layer 2 and Layer 3 devices had very different ways of forwarding packets. Now, most processes are hardware-assisted and there is no performance difference when forwarding packets on Layer 2 or Layer 3.

Layer 2 Packet Forwarding

Forwarding of packets on Layer 2 is referred to as switching. Packets are forwarded exclusively on the information that is present in the Layer 2 packet header. The switching decision is made based on the destination MAC address of the frame.

The control plane operation consists of MAC address learning and aging. MAC addresses are learned as the “conversation” occurs between two hosts, and age out after inactivity. If a MAC address is not known at the moment (that is, it has aged out), the switch “floods” the frame by sending it out of all ports. This causes communication overhead and can be a burden in large networks.

Spanning Tree Protocol (STP) is also part of the control plane and ensures proper operation of the network. It creates a loop-free topology to ensure that broadcast traffic is not replicated all over the network repeatedly.

Layer 2 Protocols

Protocols that are used in Layer 2 networks include variants of STP. The most common ones are Rapid Spanning Tree Protocol (RSTP) and Multiple Spanning Tree Protocol (MSTP). Automatic VLAN distribution can be achieved by using VLAN Trunking Protocol (VTP).

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

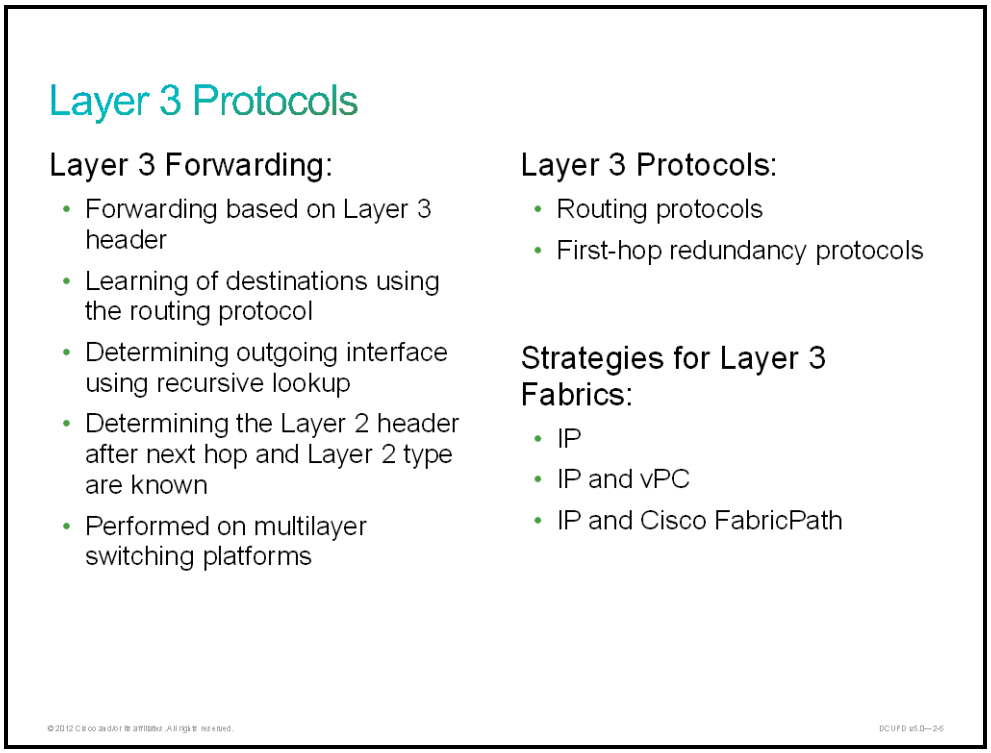
certcollecion.net

Strategies for Layer 2 Fabrics

Several strategies exist to implement Layer 2 fabrics, depending on the goals and equipment. With virtual port channels (vPCs), you can use links between switches in a more optimal way and reduce oversubscription by eliminating spanning tree blocked ports.

Cisco FabricPath allows for even better scalability and load distribution by introducing a new Layer 2 fabric.

When you are close to the threshold of 4096 VLANs, Q-in-Q makes it possible to transport more VLANs by adding another VLAN tag, which can be used to identify the customer to which the VLANs belong.



The primary Layer 3 protocol is IP.

Layer 3 Forwarding

Forwarding of packets on Layer 3 is called IP routing. The router makes a decision based on the IP addresses in the packet header.

Historically, routers had been slower by orders of magnitude compared to switches. Now, the same hardware performs packet forwarding for both Layer 2 and Layer 3, and packets are forwarded at the same rate. This is true for most data center switches.

The control plane operation is much more complex than in pure Layer 2 frame forwarding. There are several stages that build routing information.

The preferred routing protocol builds its internal topology table and selects the best paths to all advertised destinations. These best paths are installed in the global (or virtual routing and forwarding [VRF]) routing table. Typically, the routes are installed as pairs of destination networks and outgoing interfaces, or as pairs of destination networks and next-hop IP addresses.

The router then examines the routing table and, for networks that are installed as pairs of destination networks and IP addresses, performs recursive lookups to determine the outgoing interface for the next-hop IP address.

https://t.me/learningnets

certcollecion.net

When the process is complete and the adjacency table is generated, all information is encoded into hardware, where distributed switching is performed in the data plane.

Note Many routing and switching platforms offer distributed forwarding because of their modular architecture. In these cases, the data plane information is encoded in every module, enhancing the performance of the entire platform.

Layer 3 Protocols

Layer 3 protocols that are used in data center networks include routing protocols and First Hop Redundancy Protocol (FHRP). Routing protocols provide exchange of connectivity information. FHRP provides for default gateway redundancy.

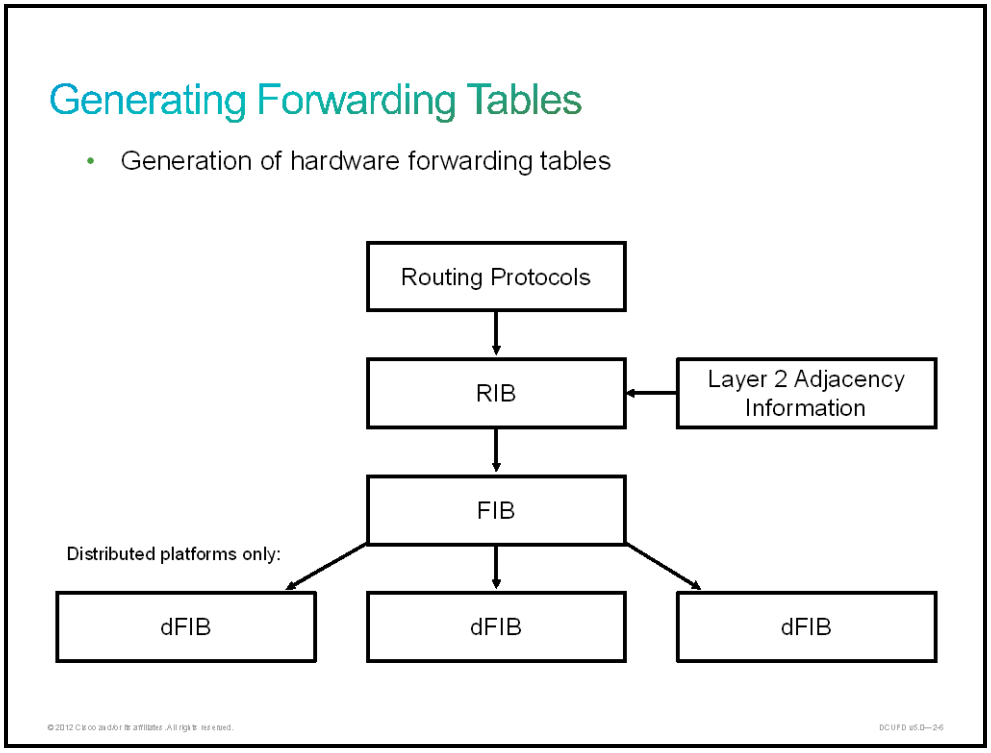
Strategies for Layer 3 Fabrics

Large Layer 3 fabrics can be achieved by using multiple links between aggregation and core switches, a routing protocol, and Equal-Cost Multipath (ECMP) as a load-balancing technology that allows for simultaneous use of multiple links. The load-balancing method depends on what the equipment supports.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net



Packet forwarding is performed based on the forwarding tables that are downloaded into the hardware.

The routing protocols select best routes based on their internal tables. These routes are entered into the IP routing table, which is the routing information base (RIB).

The forwarding information base (FIB) is generated from the RIB and additional Layer 2 adjacency information. The generation of entries in the FIB table is not packet-triggered; it is change-triggered. When something changes in the RIB, the change is also reflected in the FIB.

Because the FIB table contains the complete IP switching table, the router can make definitive decisions based on the information in the FIB. Whenever a router receives a packet, and its destination is not in the FIB, the packet is dropped.

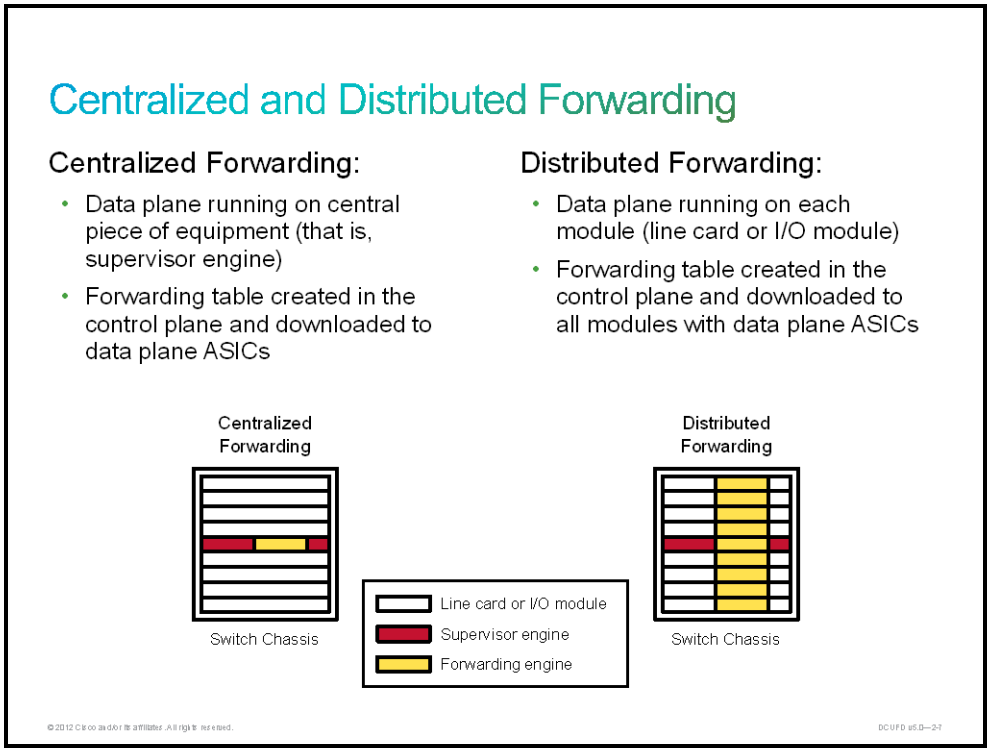
The adjacency table is derived from the Address Resolution Protocol (ARP) cache, but instead of holding only the destination MAC address, it holds the whole Layer 2 header.

On distributed platforms, the FIB is copied to all forwarding engines on the line cards and is known as the distributed FIB (dFIB) table.

Note	Examples of distributed platforms are the Cisco Nexus 7000 Switch and the Cisco Catalyst 6500 Series with line cards that feature the Distributed Forwarding Card daughterboards.

https://t.me/learningnets

certcollecion.net



There are two types of forwarding that occur on network devices: centralized and distributed.

Centralized Forwarding

Centralized forwarding takes place on nonmodular devices or on devices that are modular but have only one (central) forwarding engine.

Packet forwarding is a function of the data plane. The forwarding engine processes the packets and, based on the data that is computed by the control plane, it decides which outgoing interface will be used to forward the packet.

A centralized forwarding engine usually resides on the supervisor module or on the main board of the switch. The forwarding engine includes all required logic to process access control lists (ACLs), quality of service (QoS), packet rewriting, and so on. The memory that is used for these processes is called ternary content addressable memory (TCAM).

Examples of devices that use centralized forwarding include Cisco Catalyst 4500 Series Switches and Cisco Catalyst 6500 Series Switches with forwarding on the supervisor engine (Policy Feature Cards [PFCs] only, and no distributed forwarding cards [DFCs]).

Distributed Forwarding

Distributed forwarding is performed on modular devices.

Control plane information is precomputed on the supervisor engine. These forwarding tables are then downloaded to each forwarding engine on the line cards in the system. Each line card is fully autonomous in its switching decisions. Forwarding information on line cards is synchronized whenever there is a change in the network topology.

On distributed platforms, forwarding does not need to be interrupted when the control is changed from one supervisor to the other if they are in the same chassis.

Examples of devices that use distributed forwarding include Cisco Nexus 7000 Series Switches and Cisco Catalyst 6500 Series Switches with DFCs installed on line cards.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

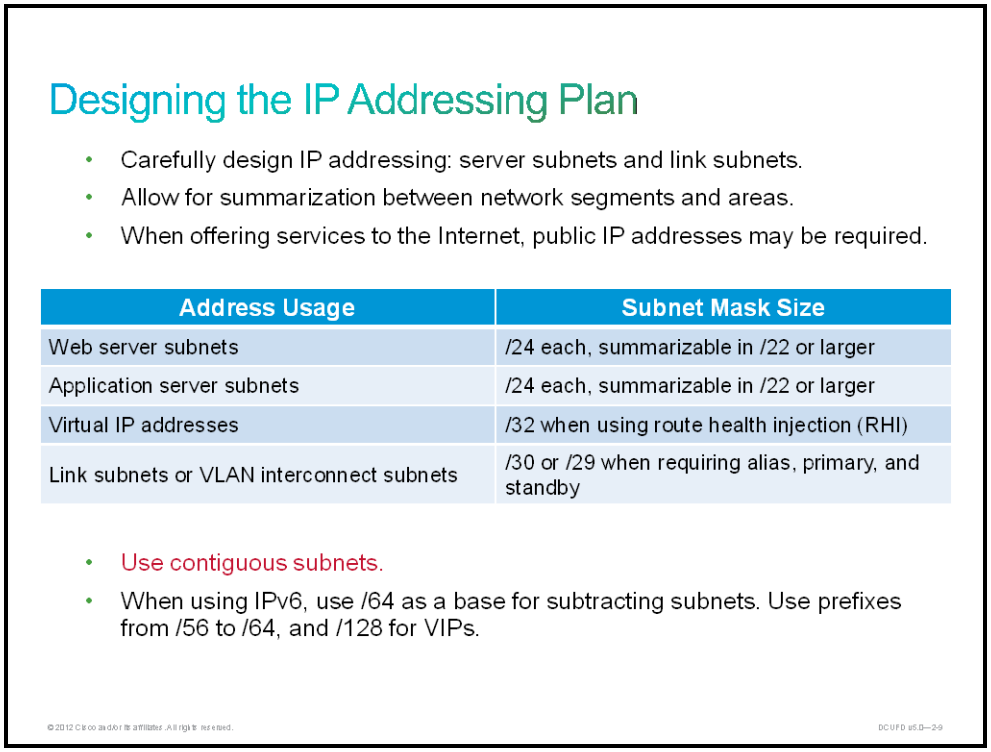
Note	Distributed forwarding on the Cisco Nexus 7000 Series Switches and on the Cisco Catalyst 6500 Series Switches is similar but not identical. The Cisco Nexus 7000 Series Switch does not perform any hardware-assisted switching on the supervisor engine, while the Cisco Catalyst 6500 Series Switch performs centralized switching for line cards that do not support distributed forwarding.

https://t.me/learningnets

certcollecion.net

IP Addressing and Routing

This topic describes IP addressing considerations and IP routing technologies.



IP addressing considerations are similar for both IP version 4 (IPv4) and IP version 6 (IPv6).

The IP addressing plan must be carefully designed with summarization in mind.

When designing a data center network, you must design the IP addressing plan as well. The IP addressing plan must be easy to use and easy to summarize:

- Define clear ranges for server subnets, and clear ranges for link subnets or “link VLANs.”
- Create an addressing plan that allows you to follow the network topology and to accommodate routing between different servers. If it is a large data center, use summarization to simplify routing decisions and troubleshooting.
- Incorporate public IP addresses if you are using them to offer services to the Internet.
- The IPv6 addressing and subnetting logic should be related enough to IPv4 subnet numbers to simplify troubleshooting.

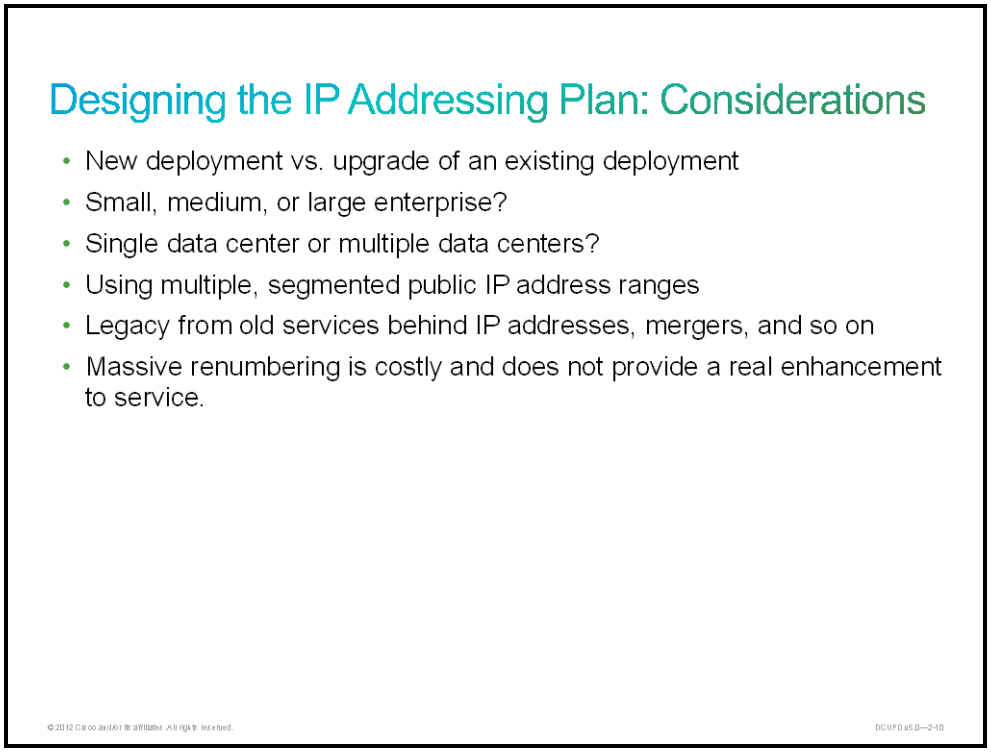
You also need a management network with a provisioned IP subnet. Typically, a /24 should be sufficient. If your data center has more devices, you may need a /23. Alternatively, you can segment the management subnet into multiple subnets: one for network infrastructure, another one for server virtualization infrastructure, and so on.

Note If your data center is very large and you need very large server subnets to accommodate moveable virtual machines, span multiple sites, and so on, you may need to provision a larger subnet than /24 for those servers.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

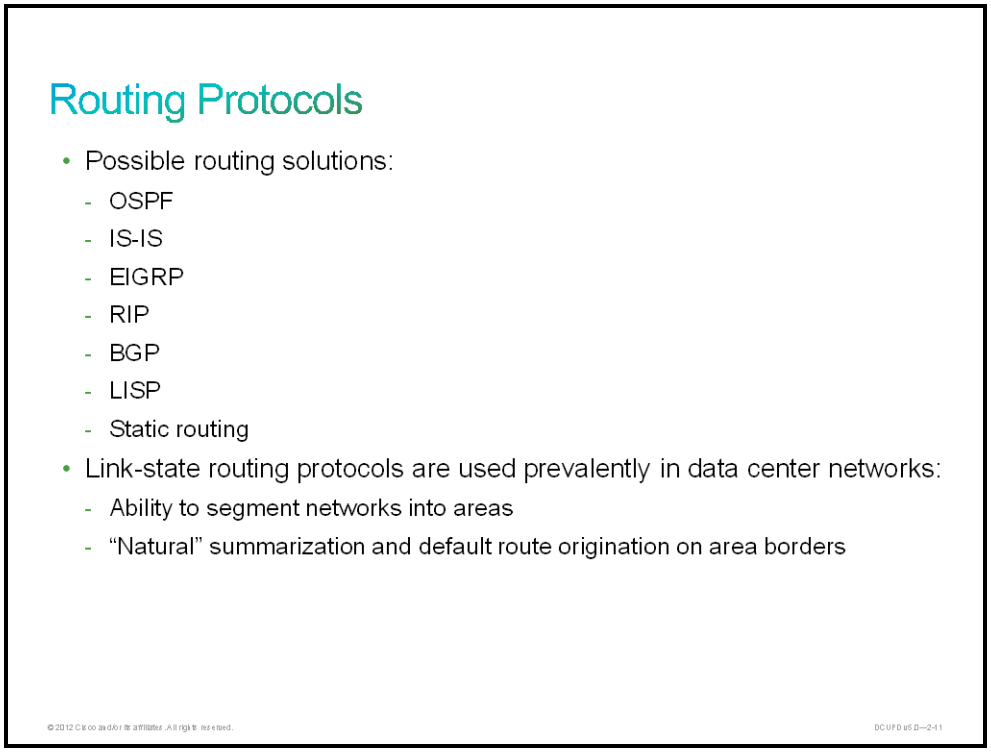


There are some additional considerations for IP addressing:

- Are you designing an addressing scheme for a new deployment or for an existing deployment? New scenarios generally allow you to assign IP addressing that can be summarized between contiguous subnets and so on, while adding a subnet to an existing network will likely not allow you to perform summarization.
- Do you need to design an IP addressing scheme for a small, a medium, or a large data center? Do you have enough IP address space in the address plan that includes the campus network?
- Do you need to provision IP address space for an additional (secondary) data center? A good practice is to keep IP addressing similar, with the second octet defining to which data center the subnet address belongs. Example: 10.132.0.0/16 for the primary data center subnets, and 10.133.0.0/16 for the secondary data center. These two can be summarized as well.
- Do you need to work with public IP addresses? You likely hold several discontiguous blocks of IP addresses of different sizes. These usually cannot be summarized.
- How did the IP addressing change in the past? Was the data center active before the companies or data centers merged or were acquired? Typically, the IP addressing remains until the existing services are in use.
- Renumbering of IP addresses is costly and requires a lot of work, not only in terms of planning, but also involving configuration work on all possible servers that are running various operating systems, on appliances, and so on. Renumbering does not enhance the service in any way.

https://t.me/learningnets

certcollecion.net



The routing protocols that are used in data center networks usually match the ones that are used in campus networks. Most commonly, the routing protocol is an interior gateway protocol (IGP), such as Open Shortest Path First (OSPF), Intermediate System-to-Intermediate System (IS-IS), or Enhanced Interior Gateway Routing Protocol (EIGRP). By using the same routing protocol, the routing domain is kept homogenous and you can take advantage of routing protocol features, such as summarization, and built-in default route origination.

The OSPF routing protocol is the most common choice. OSPF is a multiarea routing protocol that allows you to segment the enterprise network into multiple parts, which are called OSPF areas. The data center network typically uses one area.

The IS-IS protocol is used most often in service provider data centers when service providers are using IS-IS as their backbone IGP. Enterprise IS-IS deployments are not very common. The functionality is similar to OSPF in that they are both link-state routing protocols.

The EIGRP routing protocol may be used in data center networks as well. It has slightly better scalability and convergence speed, but the protocol design is different from OSPF or IS-IS. EIGRP is a hybrid routing protocol and does not feature automatic summarization on area borders.

Protocols such as Routing Information Protocol (RIP) and Border Gateway Protocol (BGP) are less common in data centers, but they are sometimes used for a specific purpose. RIP is very lightweight on processor resources, while BGP uses TCP as transport and can be used over service modules and appliances in routed mode. BGP can be used when multicast traffic cannot be transported over devices. Examples of such equipment are firewalls and application and server load balancers. By using BGP, dynamic routing is retained, and routing updates are transported by TCP sessions

Locator/Identity Separation Protocol (LISP) is used to optimally transport traffic when the location of the destination address can change.

Static routing is typically used for administratively defining a traffic path or for when equipment does not support dynamic routing. Examples of such equipment are firewalls, and application and server load balancers.

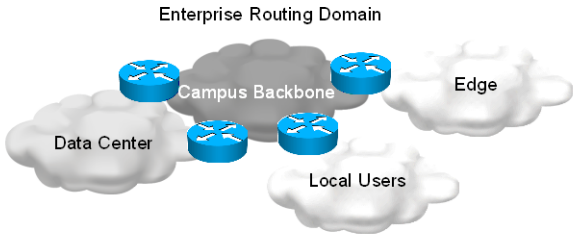
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

Routing Design Principles in Data Centers

- Use a protocol that offers fast convergence.
- Take advantage of network segmentation and summarization.
- Tune the hello intervals and dead timers; throttle Shortest Path First (SPF) calculation for link-state protocols
- Use authentication support.
- Use a protocol that supports IPv6 for future application readiness.
- Use a protocol that is supported on devices like firewalls.



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

There are guidelines regarding how to design the routing protocol configuration for data center routing.

The primary rule is to use a routing protocol that can provide fast convergence. Typically, link-down events immediately trigger new path recalculations.

Detection of a neighbor loss that is not directly connected (that is, a router over a Layer 2 switch) can take more time because routing protocols need the dead timer to expire. To speed up the convergence, you can tune the hello and dead time intervals.

Router authentication should also be used because it prevents injection of rogue routes by nonauthorized routers.

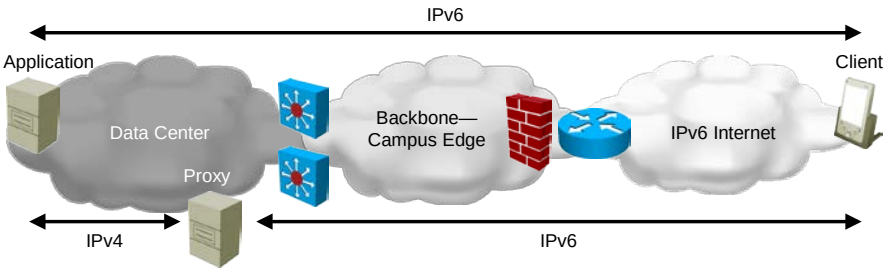
IPv6 support is available for many routing protocols. IPv6 readiness helps with migrating the applications to IPv6.

https://t.me/learningnets

certcollecion.net

Scalable IP Addressing and Routing: IPv6

- End-to-end IPv6 preferably used when serving content to IPv6-only endpoints
- Need to update and test the applications, if they are truly protocol-agnostic
- Thorough checking for compatibility with service modules and appliances
- Alternative: Involve a proxy or translator from IPv6 to IPv4 on data center edge as an interim solution if applications or equipment do not fully support IPv6



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

The IPv6 protocol offers services to IPv6 endpoints.

IPv6 is used in networks where IPv4 addressing is either unobtainable or impractical because of its scalability limitations.

IPv6 offers a larger address space for a large deployment of client devices. Mobile service providers are taking advantage of IPv6, and to many IPv6 clients you must serve IPv6 content.

IPv6 is well-supported on most networking equipment and appliances. Before deploying an IPv6 service, compatibility testing should be performed if your data center has equipment that is not from Cisco.

There are no major differences in network and routing protocol design between IPv6 and IPv4. You segment the network in the same way and you apply the same logic when designing routing protocol deployment.

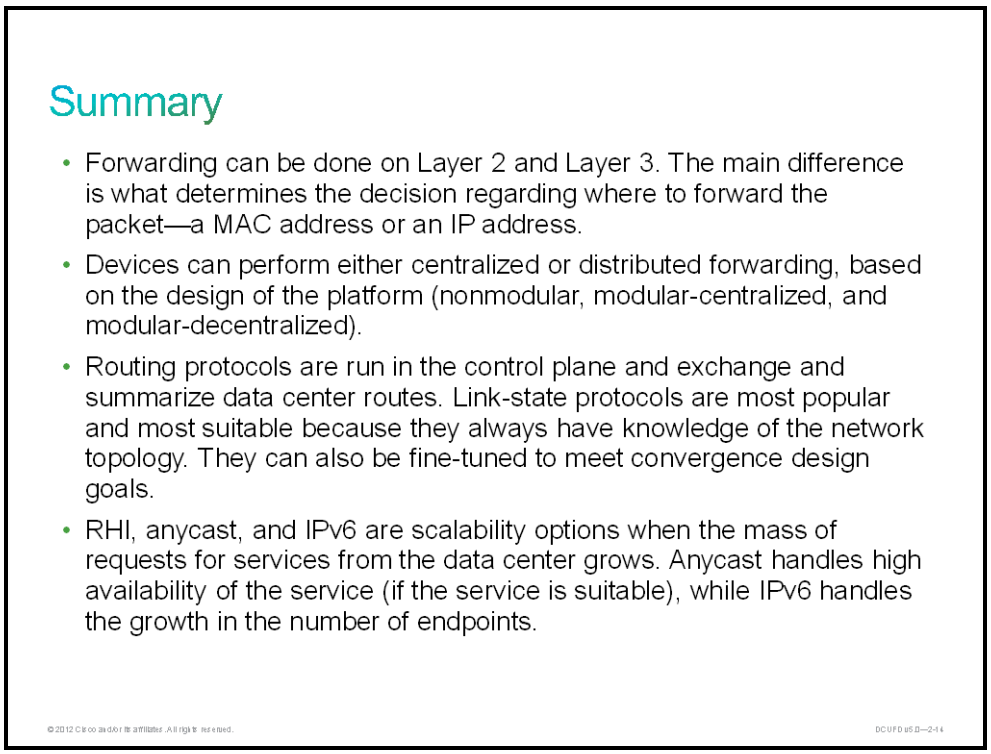
© 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

Summary

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



https://t.me/learningnets

certcollecion.net

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

Lesson 2

Virtualizing Data Center Components

Overview

This lesson describes the virtualization technologies and concepts that are used on various equipment in Cisco Data Center networks.

Objectives

Upon completing this lesson, you will be able to identify and design data center component virtualization technologies, present the limitations, and outline best practices and validated designs. This ability includes being able to meet these objectives:

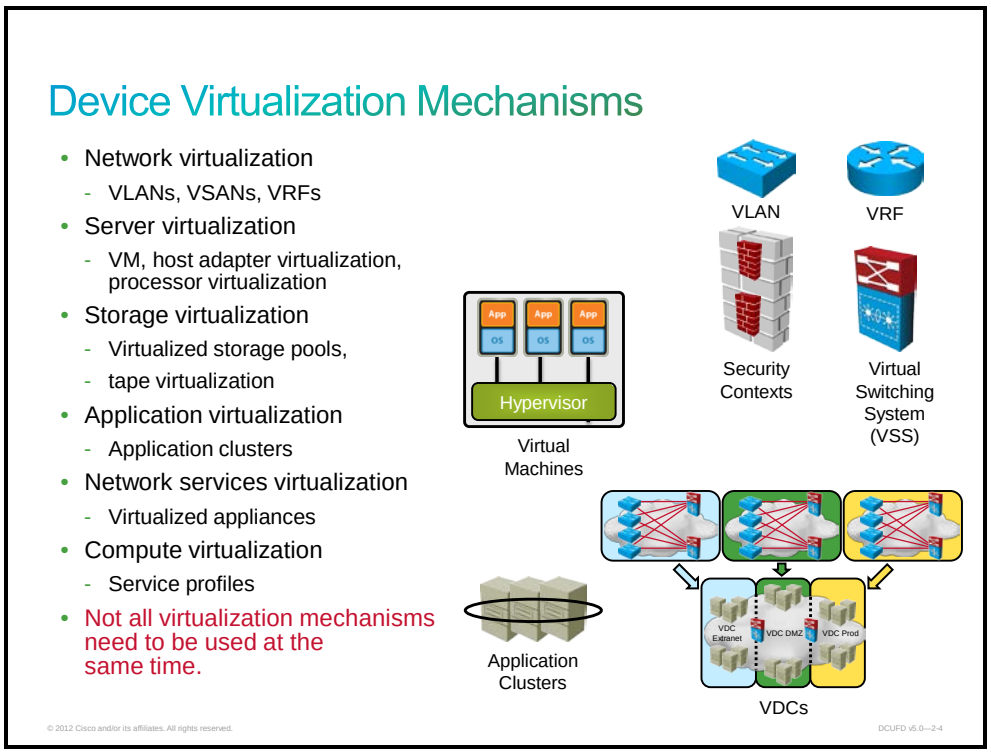
- Identify device virtualization mechanisms
- Design virtualized solutions using VDCs
- Design virtualized services using contexts on firewalling and load-balancing devices
- Design virtualized services using virtual appliances

https://t.me/learningnets

certcollecion.net

Device Virtualization Mechanisms

This topic describes device virtualization mechanisms.



Virtualization delivers tremendous flexibility to build and design data center solutions. Diverse networking needs of different enterprises might require separation of a single user group or a separation of data center resources from the rest of the network. Separation tasks become complex when it is not possible to confine specific users or resources to specific areas in the network. When separation occurs, the physical positioning no longer addresses the problem.

Note	Not all virtualization mechanisms are used at the same time and not all virtualizations are needed for a specific data center or business case. It is up to the customer to choose what virtualization to implement and possibly migrate in stages.

Network Virtualization

Network virtualization can address the problem of separation. Network virtualization also provides other types of benefits such as increasing network availability, better security, consolidation of multiple networks, segmentation of networks, and increased network availability. Examples of network virtualization are VLANs and VSANs in Fibre Channel SANs. VLAN virtualizes Layer 2 segments, making them independent of the physical topology. This virtualization allows you to connect two servers to the same physical switch, even though they participate in different logical broadcast domains (VLANs). A similar concept represents VSAN in Fibre Channel SANs.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

Server Virtualization

Server virtualization enables physical consolidation of servers on the common physical infrastructure. Deployment of another virtual server is easy because there is no need to buy a new adapter and no need to buy a new server. For a virtual server to be enabled, you only need to activate and properly configure software. Server virtualization, therefore, simplifies server deployment, reduces the cost of management, and increases server utilization. VMware and Microsoft are examples of companies that support server virtualization technologies.

Device Virtualization

The Cisco Nexus 7000 supports device virtualization or Cisco Nexus Operating System (NX-OS) virtualization. The virtual device context (VDC) represents the ability of the switch to enable multiple virtual and independent switches on the common physical switch to participate in data center networks. This feature provides various benefits to the application services, such as higher service availability, fault isolation, separation of logical networking infrastructure based on traffic service types, and flexible and scalable data center design.

Storage Virtualization

Storage virtualization is the ability to pool storage on diverse and independent devices into a single view. Features such as copy services, data migration, and multiprotocol and multivendor integration can benefit from storage virtualization.

Application Virtualization

The web-based application must be available anywhere and at any time, and it should be able to utilize unused remote server CPU resources, which implies an extended Layer 2 domain. Application virtualization enables VMware VMotion and efficient resource utilization.

Network Services Virtualization

Network services are no longer available only as standalone physical devices, but are increasingly available as virtual appliances. You can easily deploy a virtual appliance to facilitate deployment of a new application or a new customer.

Compute Virtualization

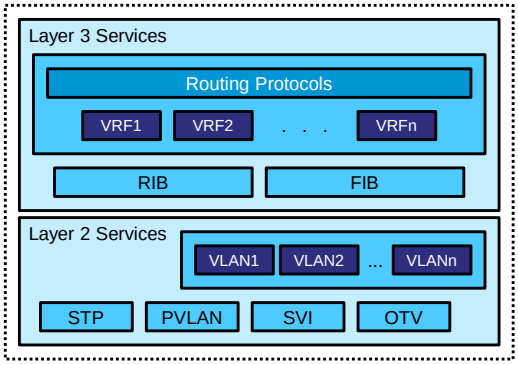
Cisco Unified Computing System (UCS) uses “service profiles” that are used as a computing virtualization mechanism. The service profiles define the personality of the server and can be applied on any hardware component that supports the abstracted hardware that is configured in the service profile. For example, if you configure a service profile with two network interface cards (NICs), the service profile can be applied to any physical server with two NICs or more.

https://t.me/learningnets

certcollecion.net

Network Devices Virtualization Overview

- Layer 2 services virtualization:
  - VLANs
  - VSANs
  - vPC
  - Cisco FabricPath
  - Cisco OTV
- Layer 3 services virtualization:
  - VRF



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

Examples of network virtualization are VLANs, which virtualize Layer 2 segments, making them independent of the physical topology.

When using unified fabric, VSANs provide a similar degree of virtualization on the SAN level. Additionally, all fabric services are started for a created VSAN.

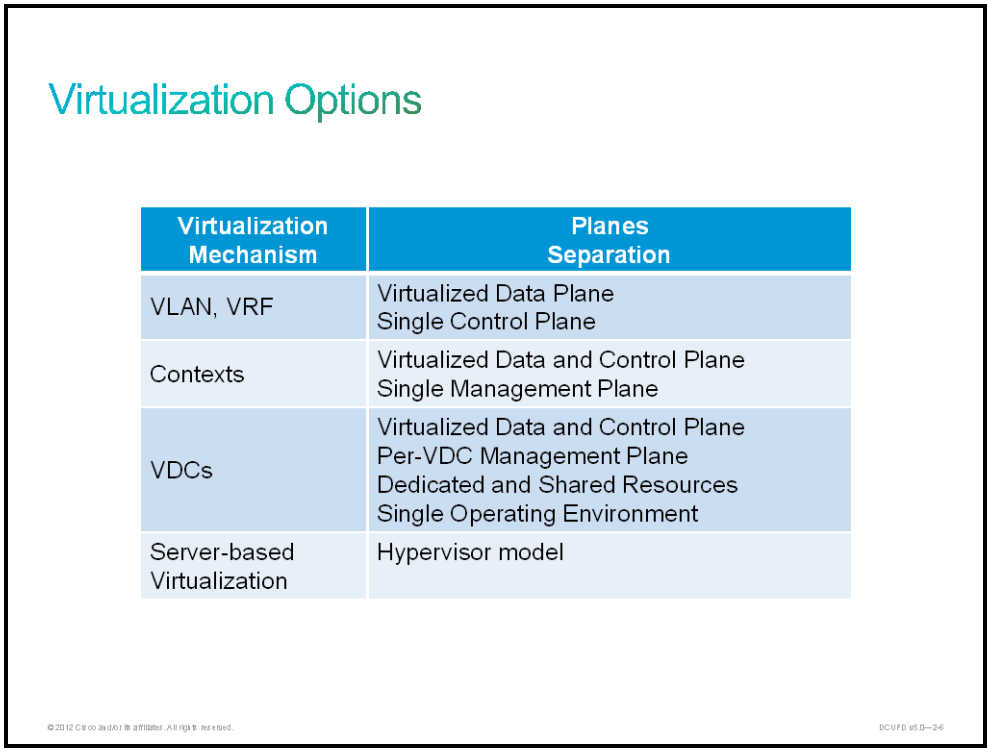
The virtual port channel (vPC) and Cisco FabricPath are examples of fabric virtualization. The vPC virtualizes the control plane in such a way that Spanning Tree Protocol (STP) on the neighbor switch is not aware that it is connected to two different switches. It receives a uniform bridge protocol data unit (BPDU).

VRF is an example of virtualization on Layer 3 that allows multiple instances of the routing table to co-exist within the same router at the same time.

Cisco Overlay Transport Virtualization (OTV) is an example of a Layer 2 extension technology that extends the same VLAN across any IP-based connectivity to multiple sites.

https://t.me/learningnets

certcollecion.net



A device can be virtualized in various ways. Each way is defined by the level of fault containment and management separation that is provided.

Several virtualization mechanisms provide separation between data, control, and management planes.

These are the main elements that are associated with virtualization:

- **Control plane:** The ability to create multiple independent instances of the control plane elements, enabling the creation of multiple logical topologies and fault domains.
- **Data (or forwarding) plane:** Forwarding tables and other databases that can be partitioned to provide data segregation.
- **Management plane:** Well-delineated management environments that can be provided independently for each virtual device.
- **Software partitioning:** Modular software processes grouped into partitions and dedicated to specific virtual devices, therefore creating well-defined fault domains.
- **Hardware components:** Hardware components partitioned and dedicated to specific virtual devices, allowing predictable allocation of hardware resources.

Switches currently provide a limited level of virtualization that uses virtual routing and forwarding (VRF) and VLANs. This level of virtualization does not partition all the various components and elements of a switch, meaning that one infrastructure can ultimately affect all VRFs and VLANs that are used.

VDC Shared and Dedicated Resources

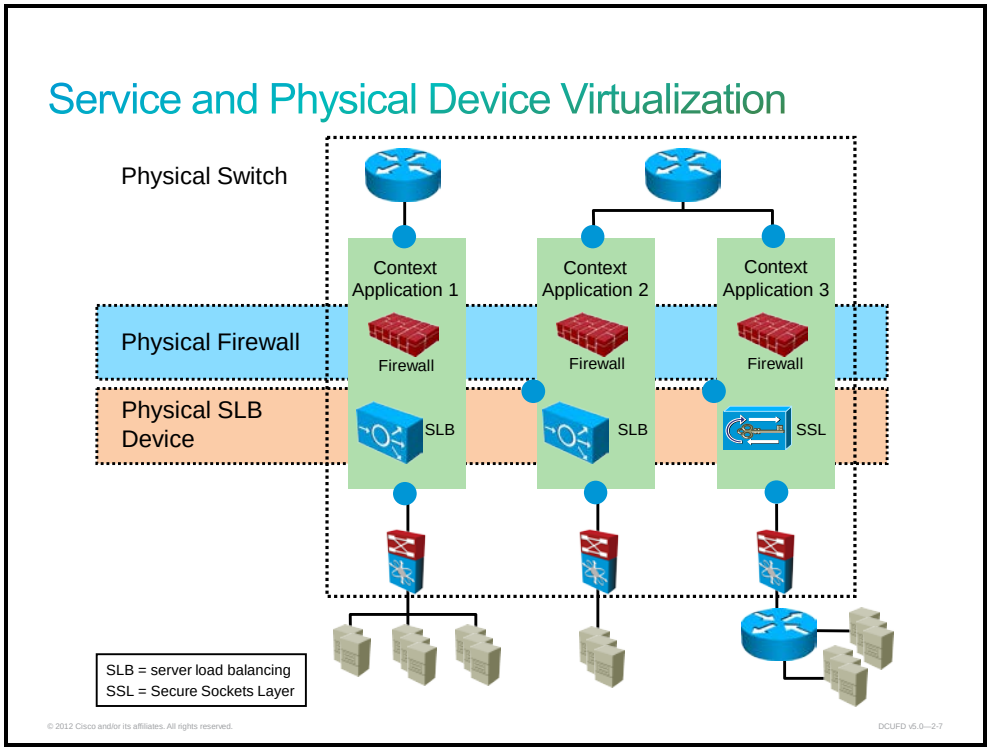
The Cisco Nexus 7000 VDCs use shared and dedicated resources.

- Dedicated resources are the resources that are used exclusively in one VDC, such as physical interfaces, ternary content addressable memory (TCAM) table space, and so on.
- Shared resources are the supervisor engine, the management interface, fabric modules, and other common hardware.

https://t.me/learningnets

certcollecion.net

There is one exception: An interface can be shared between VDCs if the port is running unified fabric. In this mode, data traffic is managed by the data VDC, where Fibre Channel over Ethernet (FCoE) traffic is managed by the storage VDC.



This figure depicts one physical service module that has been logically partitioned into several virtual service modules, and a physical switch that has been logically partitioned into several VDCs.

This partitioning reduces the number of physical devices that must be deployed and managed, but still provides the same functionality that each device could provide.

The figure shows how the physical devices (horizontal) are divided into multiple contexts, serving various applications (vertical) by using contexts, VLANs, and VRFs as virtualization means.

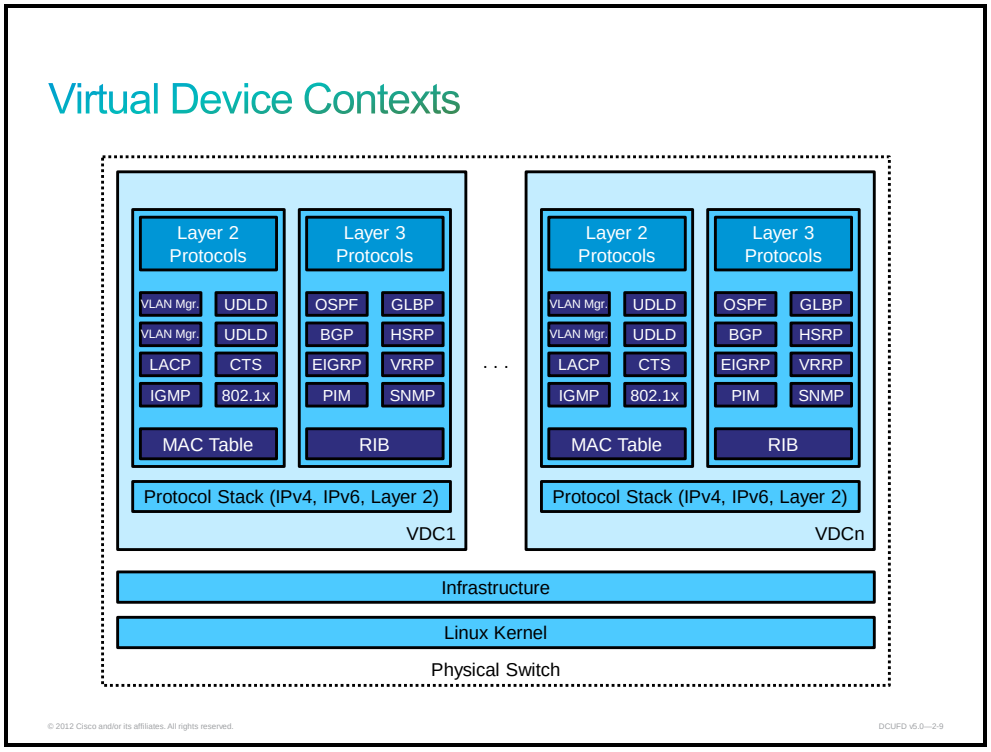
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

Virtual Device Contexts

This topic describes how to design virtualized solutions using VDCs.



Cisco Nexus 7000 Series Switches use a number of virtualization technologies that are already present in Cisco IOS Software. At Layer 2, you have VLANs. At Layer 3, you have VRFs. These two features are used to virtualize the Layer 3 forwarding and routing tables. The Cisco Nexus 7000 Switch then extends this virtualization concept to VDCs that virtualize the device itself by presenting the physical switch as multiple logical devices, each independent of each other.

Within each VDC there is a set of unique and independent VLANs and VRFs, with physical ports being assigned to each VDC. This arrangement also allows the hardware data plane to be virtualized, along with a separate management domain that can manage the VDC, therefore allowing the management plane to be virtualized as well.

In its default state, the switch control plane runs as a single device context called “VDC 1” that will run approximately 80 processes. Some of these processes will have other threads spawned, resulting in as many as 250 processes actively running on the system at any given time. This collection of processes constitutes what is seen as the control plane for a single physical device (that is, one with no other VDC that is enabled). VDC 1 is always active, always enabled, and can never be deleted. Even if no other VDC is created, support for virtualization through VRFs and VLANs is still available.

The Cisco Nexus 7000 supports multiple VDCs. The creation of additional VDCs takes these processes and replicates them for each device context that is created. When this occurs, the duplication of VRF names and VLAN IDs is possible, because each VDC represents its own logical or virtual switch context, with its own set of processes.

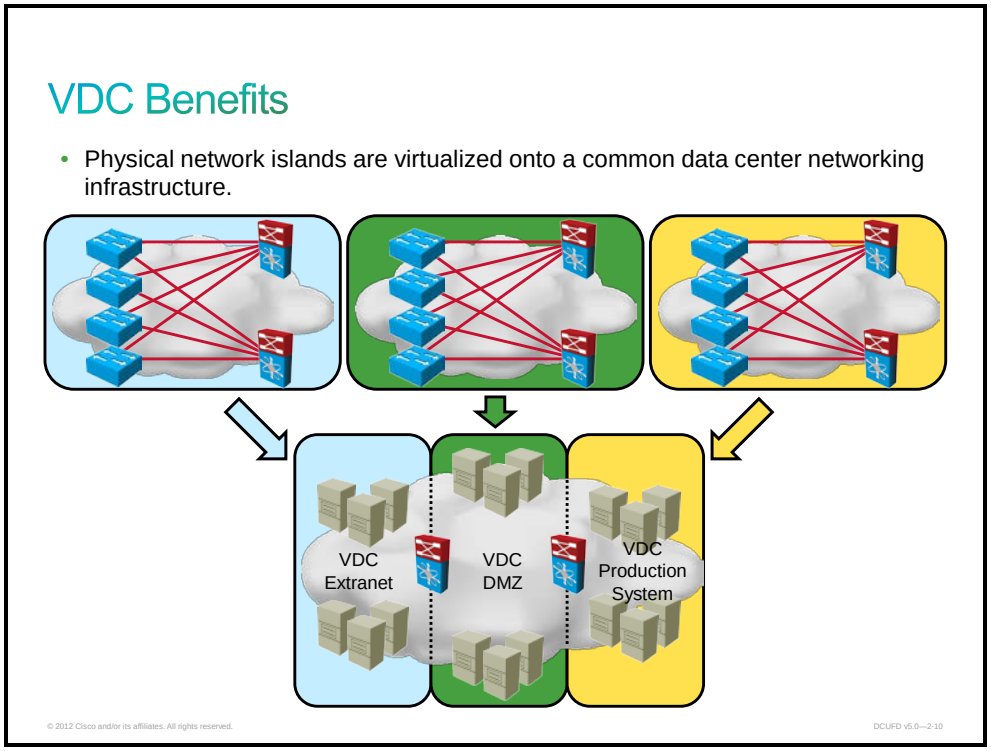
Note	Storage (FCoE) connectivity requires deployment in its own VDC. In that VDC, the Cisco Nexus 7000 Switch is a full Fibre Channel Forwarder (FCF) switch.

https://t.me/learningnets

certcollecion.net

Note You need to connect servers using unified fabric (FCoE) in a dedicated VDC, which is *not* the storage VDC. Unified fabric interfaces are then shared between the storage VDC and the "server access" VDC. Neither of these can be the default VDC.

Note Currently, Cisco FabricPath and fabric extenders (FEXs) cannot be used within the same VDC. You need to isolate the Cisco FabricPath cloud into one VDC, isolate the FEXs into another VDC, and link them using a cable.



The use of VDCs provides numerous benefits:

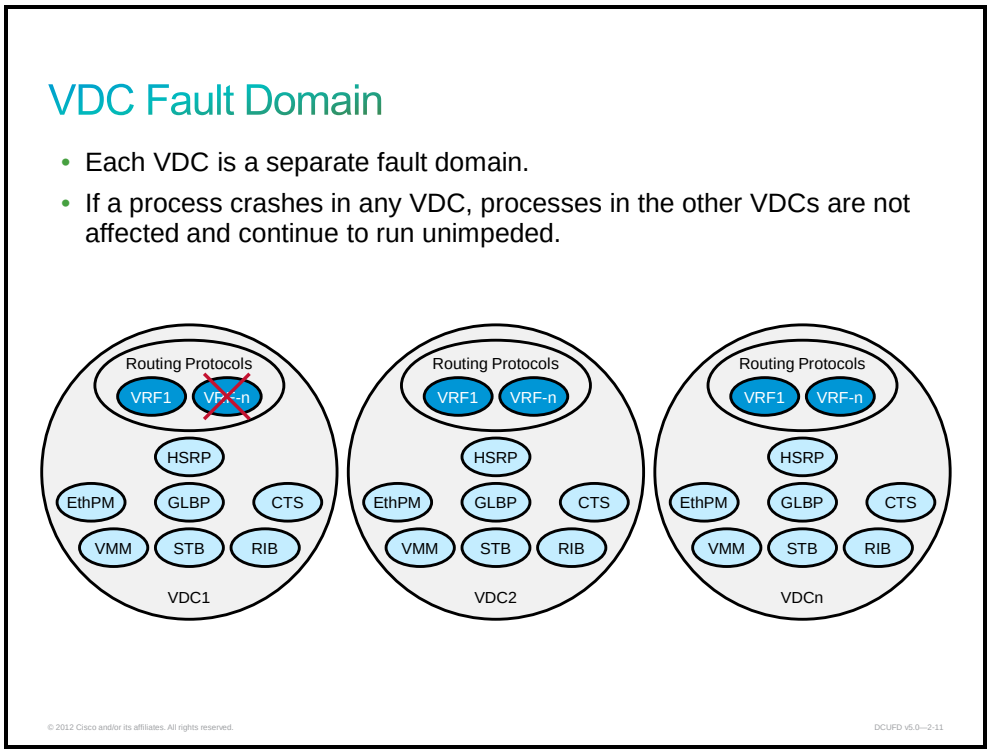
- Offering a secure network partition between different user department traffic
- Providing departments with the ability to administer and maintain their own configuration
- Providing a device context for testing new configurations or connectivity options without affecting production systems
- Consolidating multiple department switch platforms to a single physical platform while still maintaining independence from the operating system, administration, and traffic perspective
- Using a device context for network administrator and operator training purposes

If you want to send traffic from one VDC to another on the same switch, you need physical wire connectivity between them. This makes the VDC technology acceptable for various environments with compliance requirements like the Payment Card Industry (PCI), Health Insurance Portability and Accountability Act (HIPAA), or governmental regulations.

https://t.me/learningnets

certcollecion.net

Note The Cisco Nexus 7000 VDC feature has been certified by various authorities as offering sufficient security for most demanding usages.
NSS Labs certified the use of the Cisco Nexus 7000 VDC feature for PCI-compliant environments.
The Federal Information Processing Standards (FIP-140-2) certification was completed in 2011.
In the same year, the Cisco Nexus 7000 was also awarded Common Criteria Evaluation and Validation Scheme certification #10349 with EAL4 conformance.



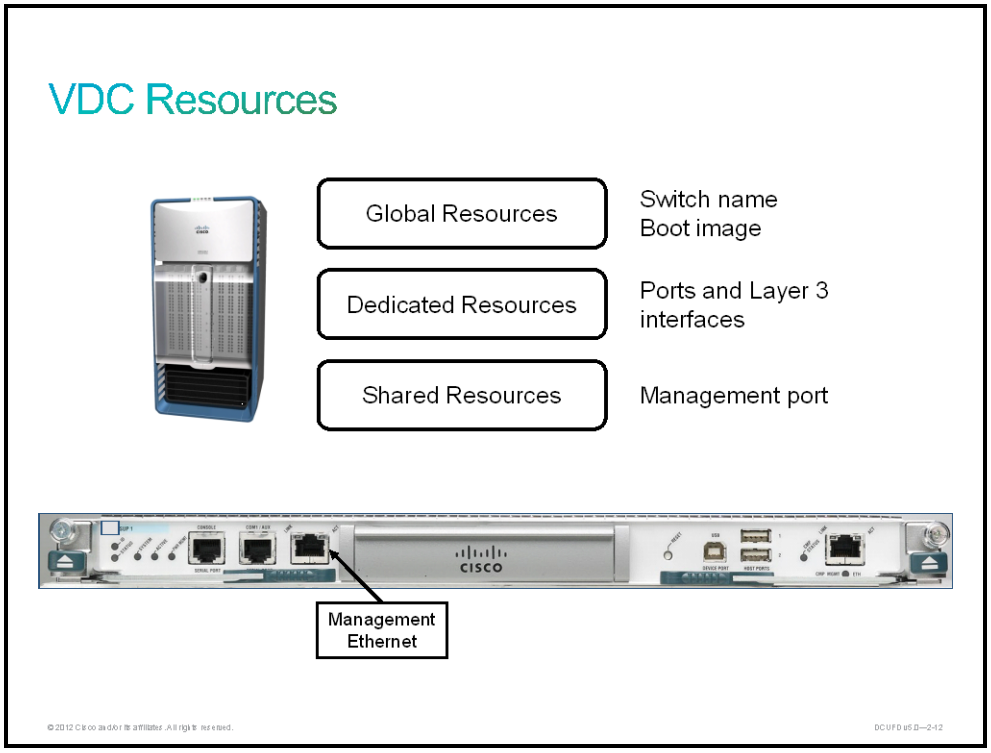
When multiple VDCs are created in a physical switch, the architecture of the VDC provides a means to prevent failures within any VDC from affecting another. For example, if a spanning tree recalculation is started in one VDC, it does not affect the spanning tree domains of other VDCs in the same physical chassis. The same isolation occurs for other processes, such as the Open Shortest Path First (OSPF) process: if it crashes in one VDC, that crash is isolated from other VDCs.

Process isolation within a VDC is important for fault isolation and serves as a major benefit for organizations that implement the VDC concept.

In addition, fault isolation is enhanced with the ability to provide per-VDC debug commands and per-VDC logging of messages from syslog. These features provide administrators with the ability to locate problems within their own VDC.

https://t.me/learningnets

certcollecion.net



There are three types of VDC resources:

- **Global resources:** Resources that can only be allocated, set, or configured globally for all VDCs, such as boot image configuration, the switch name, and in-band span session.
- **Dedicated resources:** Resources that are allocated to a particular VDC, such as Layer 2 and Layer 3 ports.
- **Shared resources:** Resources that are shared between VDCs, such as the out-of-band (OOB) Ethernet management port.

An example of a global resource is the boot string that specifies the version of software that should be used on booting up the device.

An example of a shared resource on the switch is that there is only one OOB Ethernet management port.

Note	If multiple VDCs are configured and accessible from the management port, they must share it, and they cannot be allocated to a VDC like other regular ports. The management interface does not support IEEE 802.1Q, and the management interfaces of the VDCs should be configured for the management VRF and be on the same IP subnet.

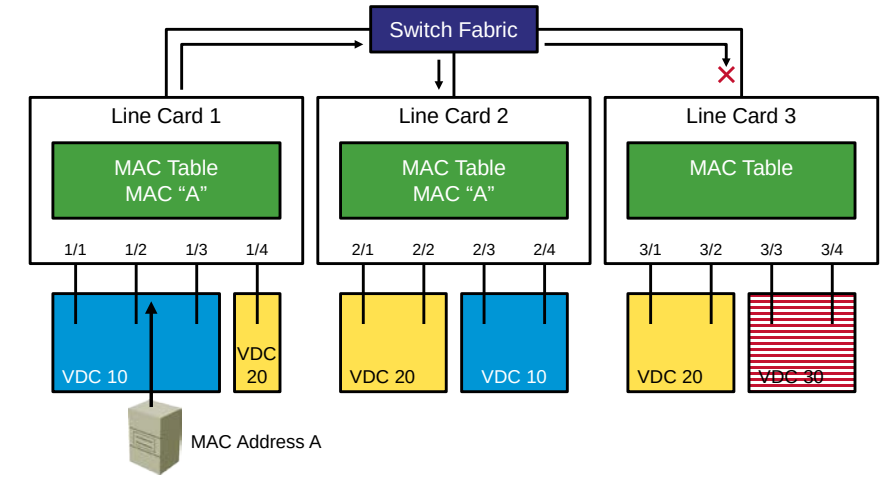
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

VDC Resource Utilization: Layer 2

- Interfaces on I/O modules are allocated to VDCs.
- I/O modules have a copy of the FIB table only for the VDCs that use ports on that I/O module.



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

When using VDCs, you must manually allocate interfaces to the VDC. The interface ceases to belong to the default VDC, and now belongs to the assigned VDC.

Note	The I/O modules have interfaces arranged into groups of ports that share a common ASIC. The ports in the same port group need to be assigned to the same VDC or are added automatically. Refer to the documentation regarding I/O modules for distribution of ports in port groups.

The forwarding engine on each line card is responsible for Layer 2 address learning and maintains a local copy of the Layer 2 forwarding table. The MAC address table on each line supports 128,000 MAC addresses, and when a new MAC address is learned by a line card, a copy is forwarded to other line cards, enabling the Layer 2 address learning process to be synchronized across all line cards.

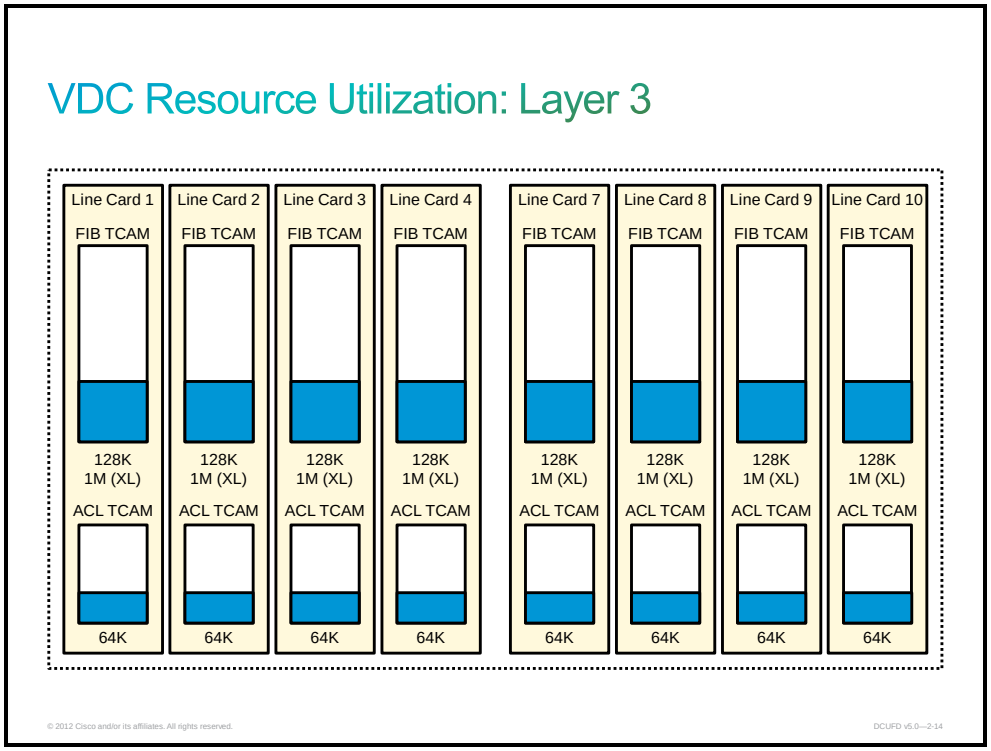
Layer 2 learning is a VDC local process and has a direct effect on the addresses that are placed on a line card. Here is an example:

1. On line card 1, MAC address A is learned from port 1/2.
2. The address is installed in the local Layer 2 forwarding table of line card 1.
3. The MAC address is then forwarded to line cards 2 and 3.
4. Line card 3 has no ports that belong to VDC 10, so it does not install any MAC addresses that are learned from that VDC.
5. Line card 2 does have a local port in VDC 10, so it installs the MAC address A into its local forwarding tables.

Note	Every I/O module has all the forwarding information bases (FIBs) for every VDC that has ports on that I/O module.

https://t.me/learningnets

certcollecion.net



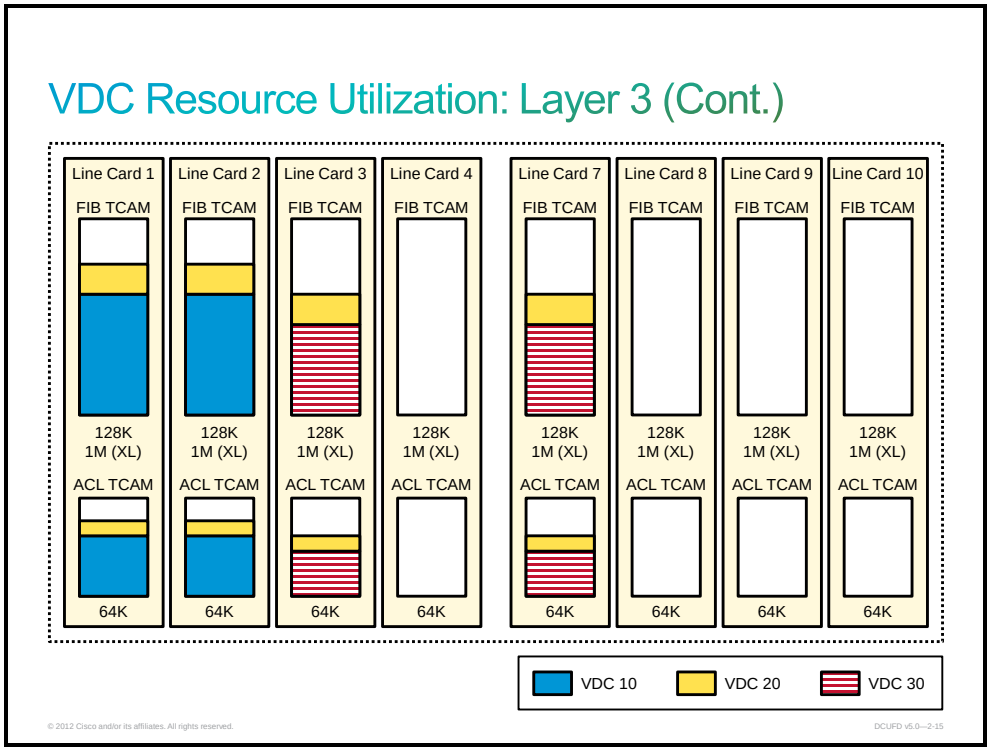
The forwarding engine on each line card supports 128,000 entries in the FIB (to store forwarding prefixes), 64,000 access control lists (ACLs), and 512,000 ingress and 512,000 egress NetFlow entries.

When the default VDC is the only active VDC, learned routes and ACLs are loaded onto each line card TCAM table so that the line card has the necessary local information to make an informed forwarding decision. This can be seen in the figure, where the routes for the default VDC are present in the FIB and ACL TCAMs.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net



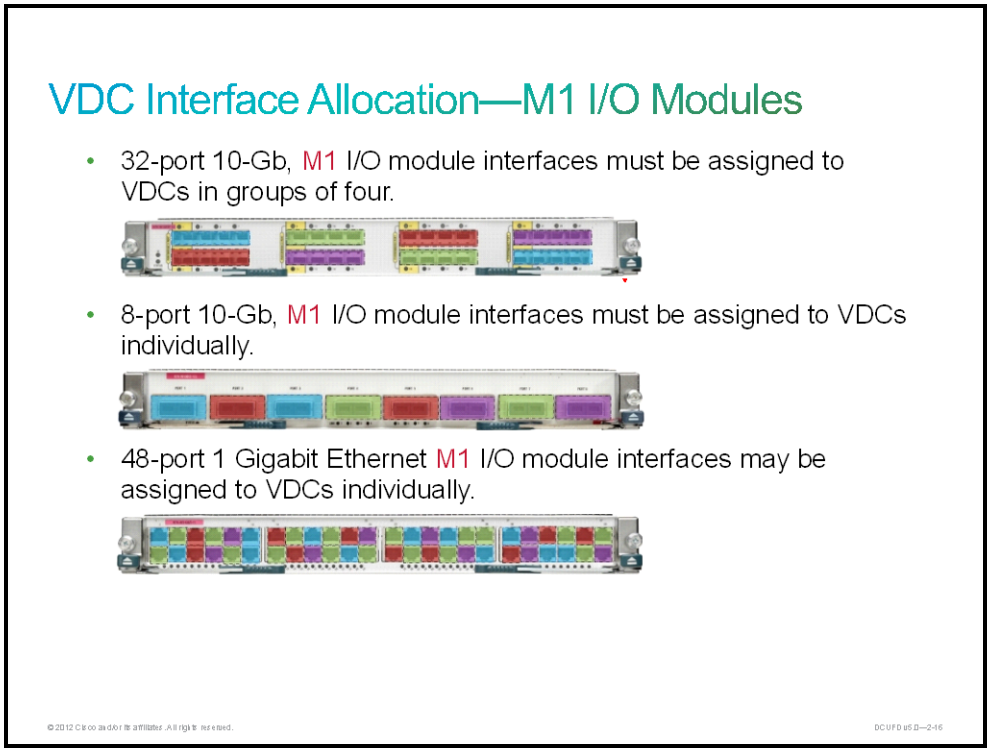
The effect of allocating a subset of ports to a given VDC results in the FIB and ACL TCAMs for the respective line cards being primed with the forwarding information and ACLs for that VDC. Using the previous example, a total of 180,000 forwarding entries are installed in a switch that, without VDCs, would have a system limit of 128,000 forwarding entries. Likewise, a total of 100,000 Cisco Application Control Engine (ACE) Module devices have been installed where a single VDC would only allow 64,000 access control entries.

More importantly, the FIB and ACL TCAM space on line cards 4, 8, 9, and 10 are free for use by additional VDCs that might be created, allowing resources to be extended beyond the system limits.

As with the TCAMs for FIB and ACLs, the use of the NetFlow TCAM is also more granular when multiple VDCs are active. When a flow is identified, a flow record is created on the local NetFlow TCAM that is resident on that line card. Both ingress and egress NetFlow are performed on the ingress line card, so it is the NetFlow TCAM of the ingress line card where the flow is stored. The collection and export of flows is always performed on a per-VDC basis. No flow in VDC 10 is exported to a collector that is part of VDC 20. After the flow is created in a NetFlow TCAM on line card 2, it is not to be replicated to NetFlow TCAMs on other line cards that are part of the same VDC; therefore, the use of the TCAM is optimized.

https://t.me/learningnets

certcollecion.net



When designing the interconnections between the switches and the VDCs, the ports on the I/O modules can be allocated to different VDCs following the hardware port groups.

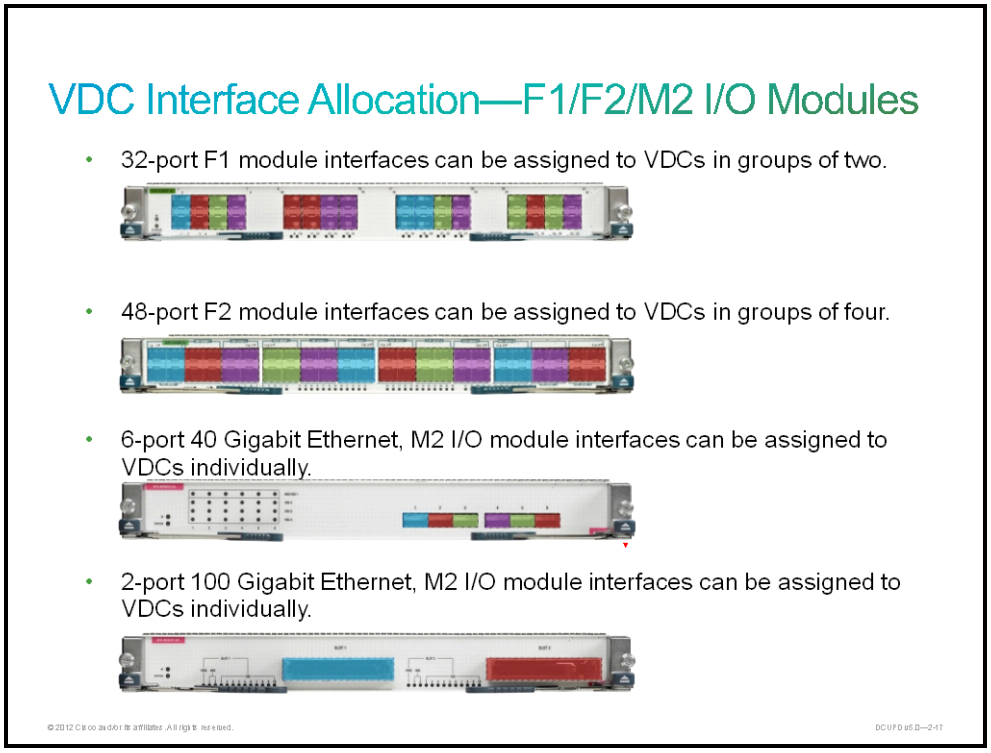
The port group layout depends on the type of the I/O module that is used.

- **32-port 10-Gb M1 and M1-XL I/O module:** The ports can be allocated to a VDC in port groups of four ports.
- **8-port 10-Gb M1-XL I/O module:** The ports can be allocated to a VDC on a per-port basis.
- **48-port 10/100/1000 M1 and M1-XL I/O module:** The ports can be allocated to a VDC on a per-port basis.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net



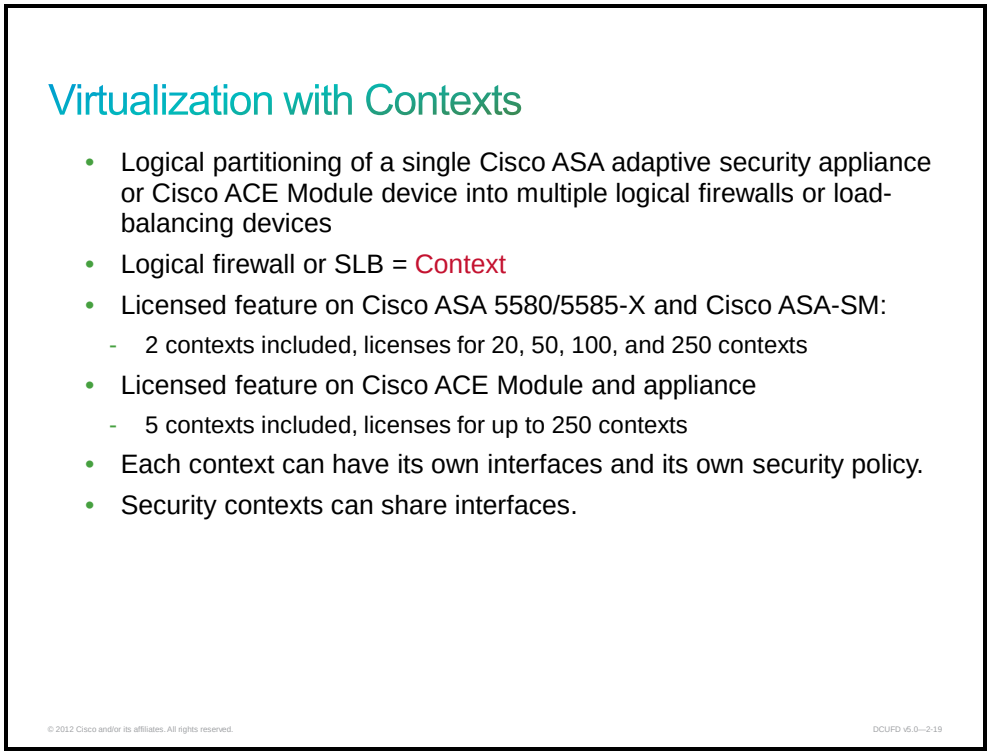
- **32-port 10-Gb F1 I/O module:** The ports can be allocated to a VDC in groups of two adjacent ports.
- **48-port 10-Gb F2 I/O module:** The ports can be allocated to a VDC in groups of four adjacent ports. All F2 I/O modules must be allocated in their own VDC. They cannot run in the same VDC as M1 and F1 I/O modules.
- **6-port 40 Gigabit Ethernet M2 I/O module:** The ports can be allocated to VDCs individually.
- **2-port 100 Gigabit Ethernet M2 I/O module:** The ports can be allocated to VDCs individually.

https://t.me/learningnets

certcollecion.net

Virtualization with Contexts

This topic describes how to design virtualized services using contexts on firewalling and load-balancing devices.



Context Virtualization Concept

Virtual firewalling presents logical partitioning of a single physical Cisco ASA adaptive security appliance, Cisco Catalyst 6500 Series Firewall Services Module (FWSM), into multiple logical firewalls. A logical firewall is called “security context” or “virtual firewall.” Similarly, the Cisco ACE Module and appliance can be partitioned in multiple contexts to accommodate multiple applications.

Security contexts allow administrators to separate and secure data center silos while providing easy management using a single system. They lower overall management and support costs by hosting multiple virtual firewalls in a single device.

Security Contexts Overview

The Cisco ASA adaptive security appliance, ASA Service Module, and the Cisco FWSM can be partitioned into multiple virtual firewalls known as security contexts. By default, two security contexts can be created on one Cisco FWSM. You need a license to deploy 20, 50, 100, and 250 concurrent security contexts.

A system configuration file controls the options that affect the entire module and defines the interfaces that are accessible from each security context.

The system configuration file can also be used to configure resource allocation parameters to control the amount of system resources that are allocated to a context.

Controlling resources enables multiple demilitarized zones (DMZs) and service differentiation classes (such as gold, silver, or bronze) per context for different data center segments.

Each individual security context has its own security policies, interfaces, and administrators.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

Each context has a separate configuration file that contains most of the definition statements that are found in a standalone Cisco FWSM configuration file. This configuration file controls the policies for the individual context, including items such as IP addressing, Network Address Translation (NAT) and Port Address Translation (PAT) definitions, authentication, authorization, and accounting (AAA) definitions, traffic control access control lists (ACLs), and interface security levels.

A virtual management interface enables management and administration for each security context and its data, and a global management interface provides configuration in real time for the entire system.

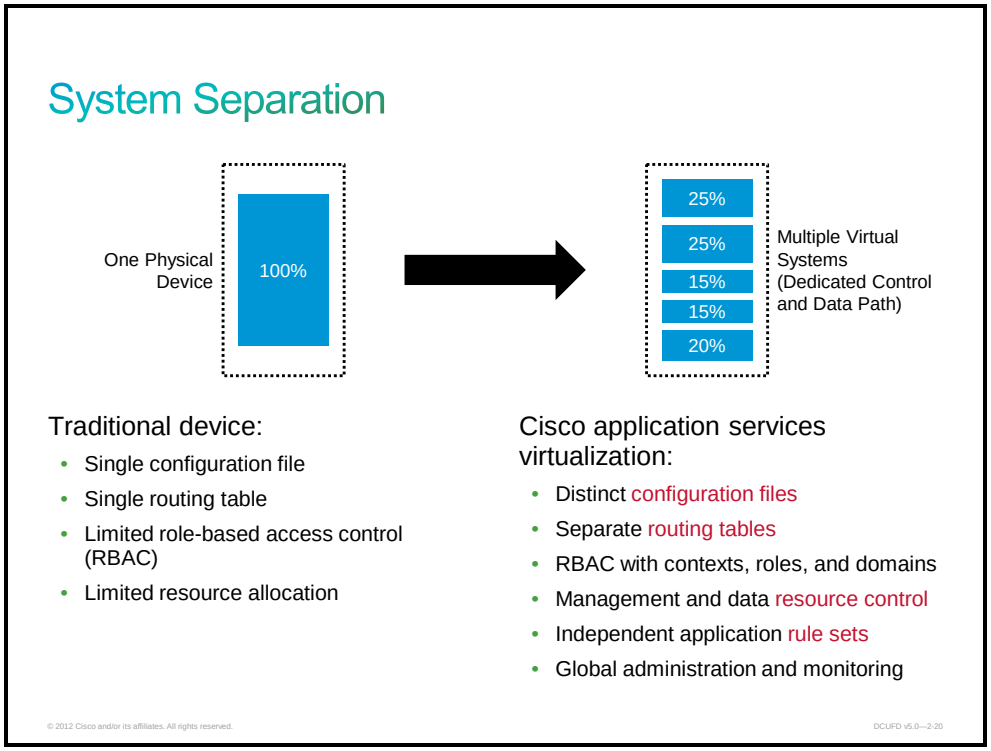
Note	Interfaces can be dedicated to a single context or shared among many contexts.
Note	On firewalls, some features such as OSPF and RIP routing are not supported in multiple context mode. On the Cisco ACE appliance, only static routing is supported.

Each security context on a multimode Cisco ASA adaptive security appliance or Cisco FWSM has its own configuration that identifies the security policy, interfaces, and almost all the options that you can configure on a single-mode firewall. Administrators can configure each context separately, even while having access to their own context only. When different security contexts connect to the same network—for example, the Internet—you can also use one physical interface that is shared across all security contexts.

Note	You can independently set the mode of each context to be either routed or transparent in Cisco FWSM and ASA Service Module.

https://t.me/learningnets

certcollecion.net

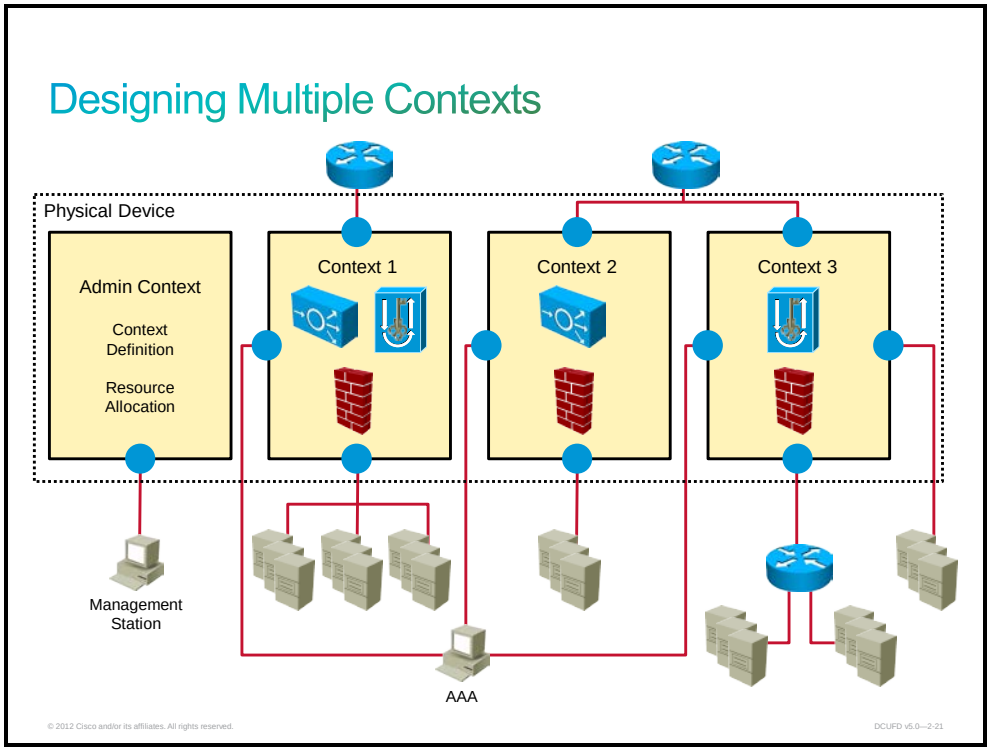


The Cisco ACE Module also supports the creation of virtual Cisco ACE Module images called “contexts.” Each context has its own configuration file and operational data, providing complete isolation from other contexts on both the control and data levels. Hardware resources are shared among the contexts on a percentage basis.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net



Network resources can be dedicated to a single context or shared between contexts, as shown in the figure.

By default, a context named “Admin” is created by the Cisco ACE Module. This context cannot be removed or renamed. Additional contexts, and the resources to be allocated to each context, are defined in the configuration of the Admin context.

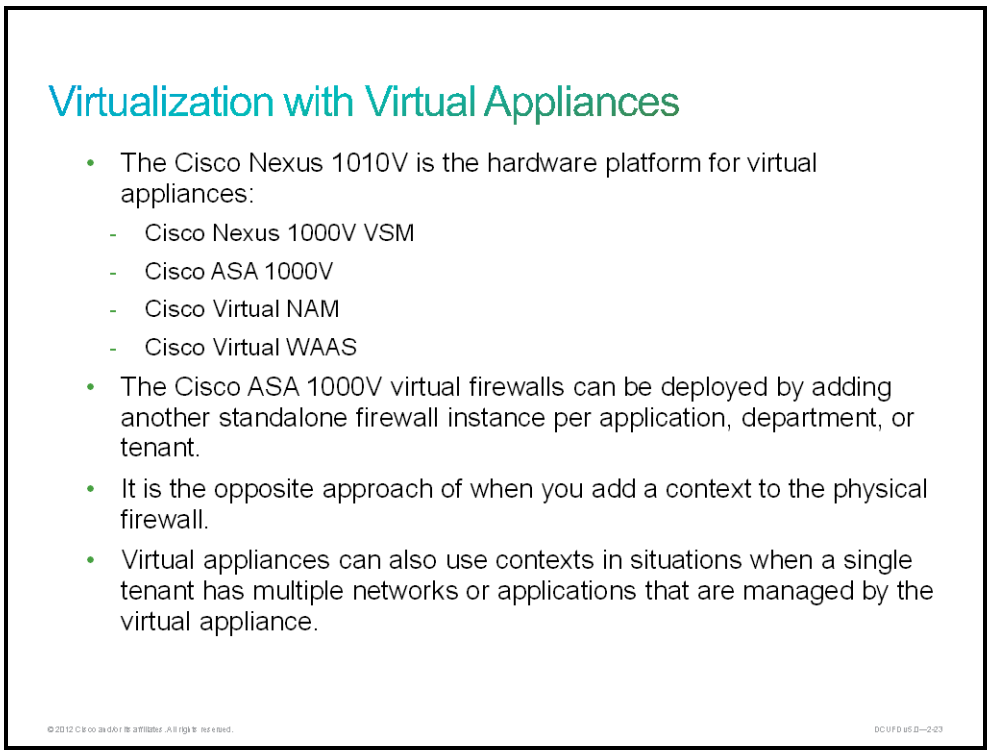
The number of contexts that can be configured is controlled by licensing on the Cisco ACE Module. The base code allows 5 contexts to be configured, and licenses are available that expand the virtualization that is possible to 250 contexts per Cisco ACE Module or 20 contexts per Cisco ACE appliance. The Admin context does not count toward the licensed limit on the number of contexts.

https://t.me/learningnets

certcollecion.net

Virtualization with Virtual Appliances

This topic describes how to design virtualized services using virtual appliances.



The classic data center design approach is to deploy physical devices and virtualizing them in order to achieve the required flexibility and resource utilization. Physical devices are then segmented into multiple contexts, or VDCs, and so on.

Cisco Nexus 1010V is the hardware platform that runs virtual devices. Originally running only the Cisco Nexus 1000V Virtual Supervisor Module (VSM), the appliance is able to run virtual service modules such as Cisco ASA 1000V, virtual NAM, virtual WAAS, and so on.

Using this platform, you can deploy a virtual firewall for a new application, customer, or department, with the full functionality of a physical firewall. In this case, you have services that would be difficult to obtain on a physical device—in the example of firewalls, a combination of routing, multiple contexts, VPN connectivity, and so on.

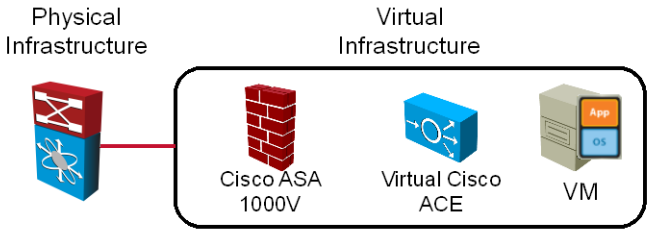
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

Virtual Services

- Traffic handled by virtual infrastructure:
  - Virtual switch (Cisco Nexus 1000V)
  - Virtual appliances
- Virtual appliances in single context mode
- Higher deployment flexibility—one chain of virtual appliances per application, department, or tenant
- Lower data throughput compared to physical appliances in multiple context mode



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

Virtual services are provided by virtual appliances. These appliances reside in the “compute” cloud within the server infrastructure in the virtual infrastructure. The primary component is the virtual switch, which resides in the virtualized host. The switch then forwards the traffic to multiple virtual appliances, which are chained by using VLANs into the correct sequence, to the virtual machine that is running the application.

The benefit of this approach is greater deployment flexibility compared to physical devices. You can simply deploy (or clone) a new virtual appliance and hand over the management of the appliance to the appropriate team or to the customer.

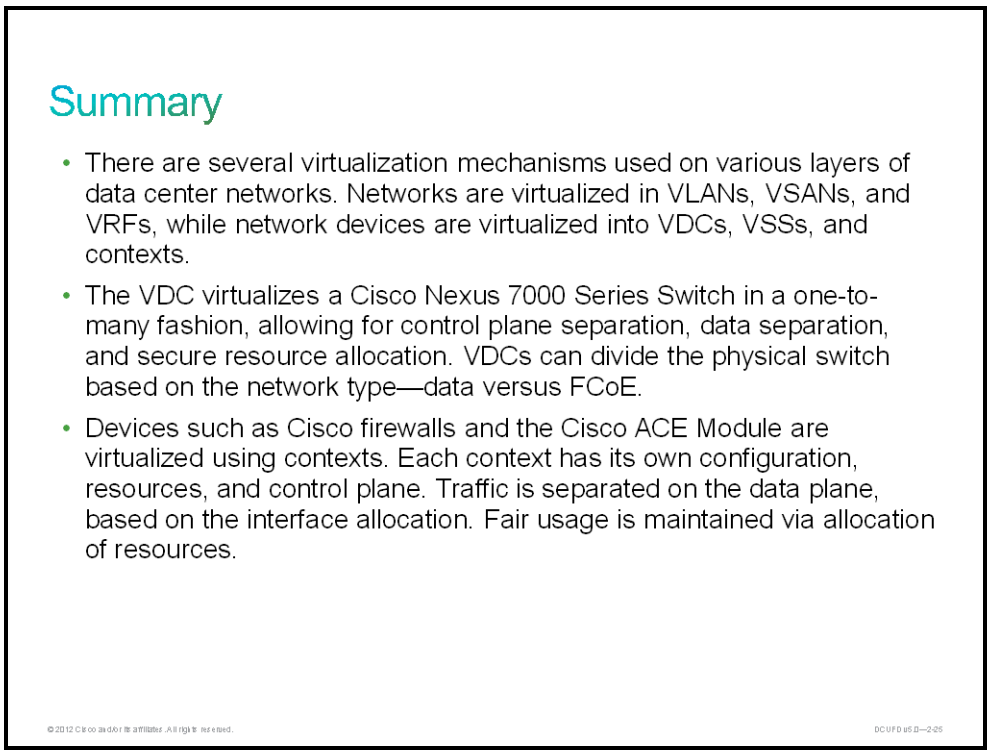
The drawback of this approach is lower data throughput, because traffic is typically switched in software on the virtualized host. Physical appliances have specialized hardware (network processors) that is purpose-built to inspect and forward data traffic. Virtualized devices run on general-purpose hardware.

https://t.me/learningnets

certcollecion.net

Summary

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



 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

Lesson 3

Designing Layer 2 Multipathing Technologies

Overview

This lesson describes multipathing technologies that are used in modern data centers with Cisco equipment. Multipathing technologies are available for both Layer 2 and Layer 3 forwarding.

Objectives

Upon completing this lesson, you will be able to design data centers using multipathing technologies, such as virtual port channel (vPC), Multichassis EtherChannel (MEC), and Cisco FabricPath, all without using Spanning Tree Protocol (STP). This ability includes being able to meet these objectives:

- Explain link virtualization technologies that allow for scaling of the network
- Design solutions using vPCs and MEC
- Design solutions using Cisco FabricPath

https://t.me/learningnets

certcollecion.net

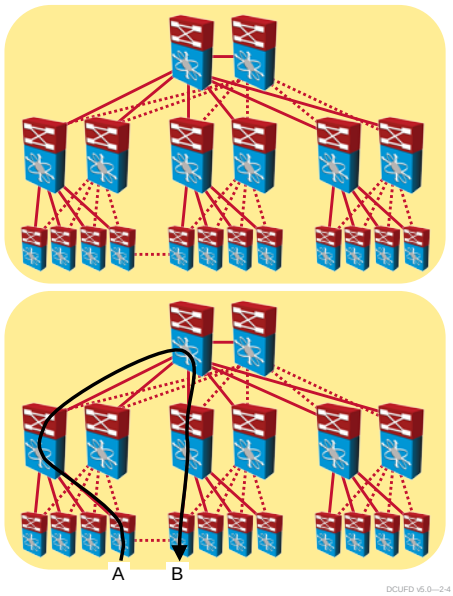
Network Scaling Technologies

This topic describes link virtualization technologies that allow for scaling of the network.

Traditional Networks Scalability Limitations

Traditional networks have scalability limitations:

- Only one physical or logical link between any two switches
- Suboptimal paths between two switches introduced by the tree topology
- North-South traffic flows across only one link
  - Only 50% of bandwidth available for client-server connections
- East-West traffic must be switched across aggregation or even core switch
  - Same bandwidth constraint for extended clusters or VM mobility



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

When traditional networks are implemented using Layer 2, there are scalability limitations that are primarily introduced because of usage of STP.

Generally, a good solution to this limitation is to segment the network to divide it into several Layer 3 domains. While this solution is proven and recommended, it is not trivial to implement— it requires at least an IP addressing plan and IP routing.

STP is not scalable enough to have larger Layer 2 domains. By building a tree topology, it blocks all other links to switches that could create a Layer 2 loop or provide an alternate path.

STP limits the upstream traffic (“North-South”) to only 50 percent of the bandwidth that is available toward the upstream switch.

Traffic between servers (“East-West”), if they are not on the same access switch, suffers the same limitation. Examples of such traffic are clustered or distributed applications, and virtual machine mobility (such as VMware VMotion).

There are two technologies that provide a solution for the limitations of STP:

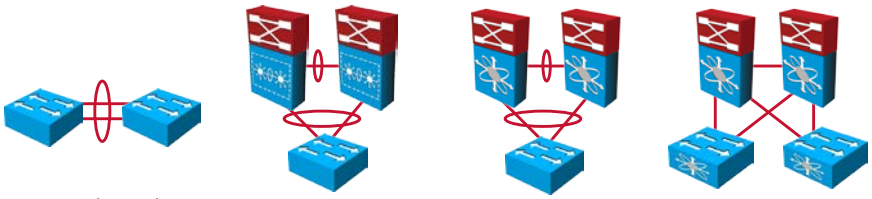
1. Multilink aggregation by using MEC or vPC to overcome the blocked bandwidth limitation imposed by STP
2. Cisco FabricPath, for even greater scalability

https://t.me/learningnets

certcollecion.net

Network Scaling Technologies

- Layer 2 multipathing technologies are used to scale bandwidth of links:
  - Traditional EtherChannel or PortChannel
- Mechanisms to overcome the STP issue of a blocked link:
  - MEC
  - vPC
- Technologies for even greater scalability:
  - Cisco FabricPath
  - TRILL



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

Link virtualization technologies are used to scale links between data center switches and other equipment.

Multipathing technologies virtualize links so that a link is presented as a single logical link to the control planes, but, on a lower level, it is a bundled link of several physical links. An example of this technology is EtherChannel or port channel.

To add more bandwidth between a pair of generic switches, the suitable technology is EtherChannel (or port channel).

To link an access switch with multiple aggregation switches and without having STP blocking one link, you can use, depending on equipment, either of the following:

- MEC on Cisco Catalyst 6500 Virtual Switching System (VSS)
- vPC on the Cisco Nexus family of switches

Not all platforms support all types of EtherChannels and port channels.

Another technology that uses link virtualization is fabric extender (FEX). FEXs provide additional ports that are attached to a remote, lightweight chassis that does not perform control plane operations. All interfaces are managed through the managing switch and presented through the interface that attaches the FEX.

Note The interfaces on the FEX are presented as logical interfaces on the managing switch.

Both MEC and vPC technologies scale to up to two upstream switches. These technologies provide a robust upstream (aggregation) layer, consisting of a single switch pair or multiple pairs of switches. However, for traffic that needs to flow between several aggregation blocks, the traffic paths are deterministic and need to travel through the data center core.

Cisco FabricPath is a technology that allows you to wire access and aggregation switches in a fabric, making the path selection process more flexible and more redundant. This technology might utilize the links better than vPC or MEC.

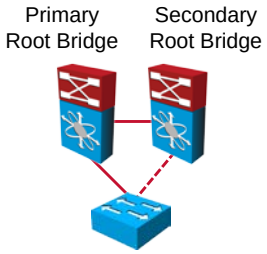
https://t.me/learningnets

certcollecion.net

Cisco FabricPath can accommodate “East-West” traffic across several links and across multiple topologies, without traffic leaving the aggregation layer.

Scaling STP

- STP is used for compatibility or as a failsafe mechanism for vPC.
- Rapid PVST+
  - Easy deployment
  - Every VLAN has its own instance, generated BPDUs, and configured primary and secondary root bridges.
- MST
  - Better scalability than Rapid PVST+
  - Every instance has its primary and secondary root bridge, and generates BPDUs. There is less CPU processing.
  - VLANs assigned to instances
  - Typically, two instances are configured per aggregation block.



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

STP is still used in modern data centers. The use cases are when a piece of equipment does not support any Layer 2 multipathing mechanism, or as a failsafe mechanism when using vPC.

There are two options for STP deployment: Rapid Per VLAN Spanning Tree Plus (Rapid PVST+), and Multiple Spanning Tree (MST).

For Rapid PVST+, the switch starts an instance of STP for every created VLAN. You can configure primary and secondary root bridges for every VLAN, and the switch will generate and process bridge protocol data units (BPDUs) for every VLAN that it has defined. When you have many VLANs that are defined, the CPU load is significant, especially when failures occur in the network.

MST is much more scalable because you manually define STP instances, and primary and secondary root bridges. The switch then generates and processes BPDUs *per instance*, which is much fewer than *per VLAN*. VLANs are then assigned to those instances.

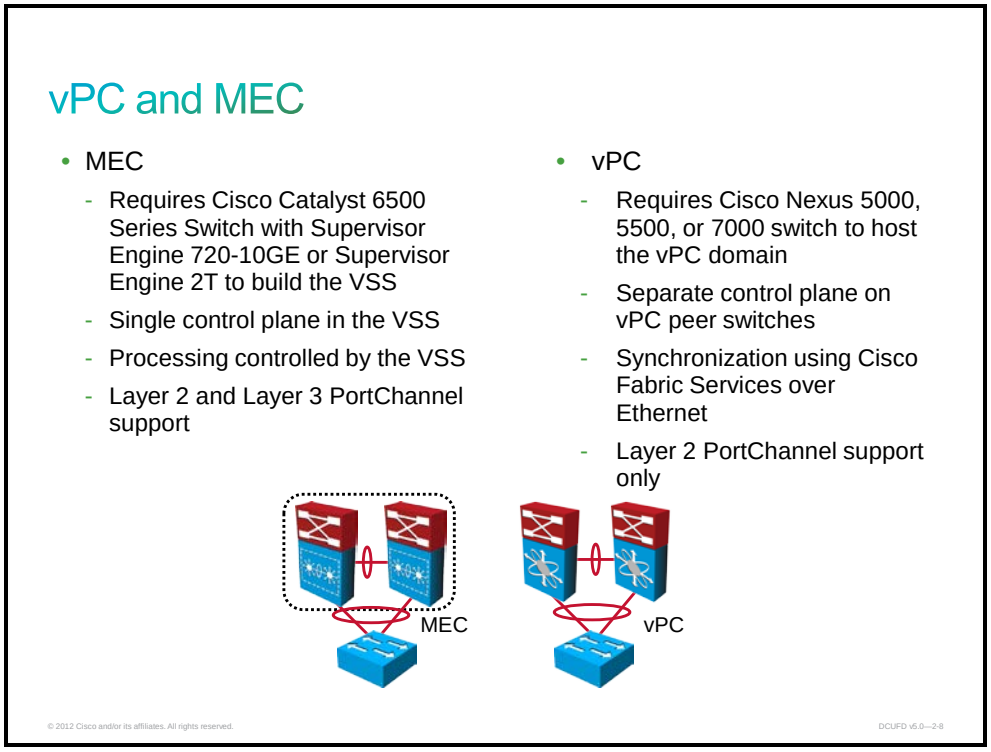
Typically, two instances of MST are sufficient per aggregation block.

https://t.me/learningnets

certcollecion.net

vPC and MEC

This topic describes how to design solutions using vPCs and MEC.



MEC and vPC are technologies that allow you to terminate a single port channel on two remote devices. Remote devices run a control protocol that synchronizes the state of the port channel and maintains it.

MEC is used where Cisco Catalyst 6500 Series Switches are bonded in VSS configuration. vPC is used where Cisco Nexus Series switches are used. From the perspective of the downstream switch, both technologies look the same, but there are fundamental differences regarding how the control plane works on the aggregation (or upstream) devices.

MEC and VSS

The Cisco Catalyst VSS unifies a pair of two Cisco Catalyst 6500 Series Switches with the Supervisor Engine 720-10GE, or Supervisor Engine 2T, into a single logical system, using a single control plane. The control plane takes care of how the multichassis port channel is managed and synchronized. There is one single control plane, running on the active supervisor engine.

vPC

The vPC functions differently. The aggregation Cisco Nexus Series Switches (5000, 5500, 7000) function as two separate switches, each one with its own control plane. The vPC is managed using a common entity on both switches—the vPC domain. This vPC domain uses a special protocol to synchronize and maintain the port channels—the Cisco Fabric Services over Ethernet.

https://t.me/learningnets

certcollecion.net

Primary Differences

This table lists the differences between VSS and vPC from the control plane perspective.

When VSS is created, two Cisco Catalyst 6500 Series Switches form a single virtual switch with a single control plane. The virtual switch builds a single routing instance. From the perspective of all other devices, the virtual switch is one network device, a single dynamic routing protocol neighbor. On VSS, only one configuration is maintained.

A Cisco Nexus switch with a configured vPC has an independent control plane. That means that there are two independent routing instances. vPC member devices have independent configurations.

VSS versus vPC

	VSS on Cisco Catalyst 6500 Series Router	vPC on Cisco Nexus 7000 Series Switch
Control plane	Single logical node	Two independent nodes
Control plane protocols	Single instance	Independent instances
Layer 3 port channel support	Yes	No
High availability	Interchassis	Intrachassis, per process
EtherChannel	Static, PAgP, PAgP+, LACP	Static, LACP
Configuration	Configuration on one device	Two configurations to manage

Note Unlike MEC, vPC does not support Layer 3 port channels. Routing protocol adjacencies cannot be formed over a vPC. A dedicated Layer 3 link must be used for this purpose.

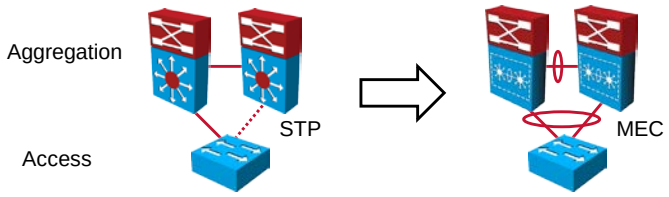
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

MEC and vPC Benefits

- Existing infrastructure can be reused—only aggregation switches might require an upgrade.
- Builds loop-free networks; STP is run as a failsafe mechanism.
- Condition: physical cabling must be of “looped triangle” design.
- Benefits from regular EtherChannels: added resiliency, easy scaling by adding links, optimized bandwidth utilization, and improved convergence.
- Depending on the platform, the amount of active links can be up to 8 (of 16 configured), or 16 (of 32 configured).



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

One of the important benefits of deploying MEC or vPC is that existing infrastructure can be reused, even without rewiring. Switches like the Cisco Catalyst 6500 can require a supervisor engine upgrade to be able to form a VSS, but, as a benefit, the amount of oversubscription between aggregation and access is reduced by half.

The primary loop avoidance mechanism is provided by MEC or vPC control protocols. STP is still in operation and is running as a failsafe mechanism.

To be able to upgrade or migrate to MEC or vPC, access switches must be connected using the “looped triangle” cabling design, as shown in the figure.

Link Aggregation Control Protocol (LACP) is the protocol that allows for dynamic port channel negotiation and allows up to 16 interfaces into a port channel.

Note On Cisco Nexus 7000 Series switches, a maximum of eight interfaces can be active, and a maximum of eight interfaces can be placed in a standby state on the Cisco Nexus 7000 M Series modules. When using Cisco Nexus 7000 F-Series modules, up to 16 active interfaces are supported in a port channel.

Port Channel Load-Balancing Algorithms

Ethernet port channels provide load balancing based on the following criteria:

- For Layer 2 frames, it uses the source and destination MAC address.
- For Layer 3 frames, it uses the source and destination MAC address and the source and destination IP address.
- For Layer 4 frames, it uses the source and destination MAC address, the source and destination IP address, and the source and destination port address.

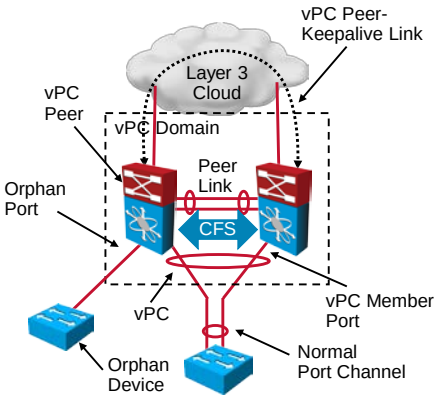
Note You can select the criterion in the configuration.

https://t.me/learningnets

certcollecion.net

vPC Architecture Components

- vPC Peers: A pair of vPC-enabled switches
- vPC Domain: A pair of vPC peers and associated vPC components
- vPC Peer-Keepalive Link: Routed link carrying heartbeat packets for active/active detection
- vPC Peer Link: Carries control traffic between vPC peer devices
- vPC: Combined port channel between the vPC peers and a port channel-capable downstream device
- vPC Member Port: One of a set of ports that form a vPC



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

The vPC architecture consists of the following components:

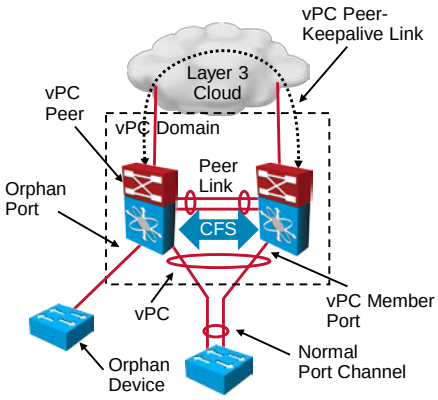
- **vPC peers:** The core of the vPC architecture is a pair of Cisco Nexus switches. This pair of switches acts as a single logical switch, which allows other devices to connect to the two chassis using MEC.
- **vPC domain:** The vPC domain includes both vPC peer devices, the vPC peer-keepalive link, the vPC peer link, and all of the port channels in the vPC domain that are connected to the downstream devices. A numerical vPC domain ID identifies the vPC. You can have only one vPC domain ID on each virtual device context (VDC).
- **vPC peer-keepalive link:** The peer-keepalive link is a logical link that often runs over an out-of-band (OOB) network. It provides a Layer 3 communication path that is used as a secondary test to determine whether the remote peer is operating properly. No data or synchronization traffic is sent over the vPC peer-keepalive link; only IP packets that indicate that the originating switch is operating and running a vPC are transmitted. The peer-keepalive status is used to determine the status of the vPC peer when the vPC peer link goes down. In this scenario, it helps the vPC switch to determine whether the peer link itself has failed, or if the vPC peer has failed entirely.
- **vPC peer link:** This link is used to synchronize states between the vPC peer devices. Both ends must be on 10 Gigabit Ethernet interfaces. This link is used to create the illusion of a single control plane by forwarding BPDUs and LACP packets to the primary vPC switch from the secondary vPC switch.
- **vPC:** A vPC is a MEC, a Layer 2 port channel that spans the two vPC peer switches. The downstream device that is connected on the vPC sees the vPC peer switches as a single logical switch. The downstream device does not need to support vPC itself. It connects to the vPC peer switches using a regular port channel, which can either be statically configured or negotiated through LACP.
- **vPC member port:** This is a port on one of the vPC peers that is a member of one of the vPCs that is configured on the vPC peers.

https://t.me/learningnets

certcollecion.net

vPC Architecture Components (Cont.)

- vPC VLAN: VLAN carried over the peer link and across the vPC
- Non-vPC VLAN: STP VLAN not carried over the peer link
- Orphan Device: A device that is connected to a vPC peer using a non-vPC link
- Orphan Port: Port on a vPC peer that connects to an orphan device
  - The term “orphan port” is also used for a vPC member port that connects to a device that has lost connectivity to the other vPC peer.
- Cisco Fabric Services: A protocol that is used for state synchronization and configuration validation between vPC peer devices

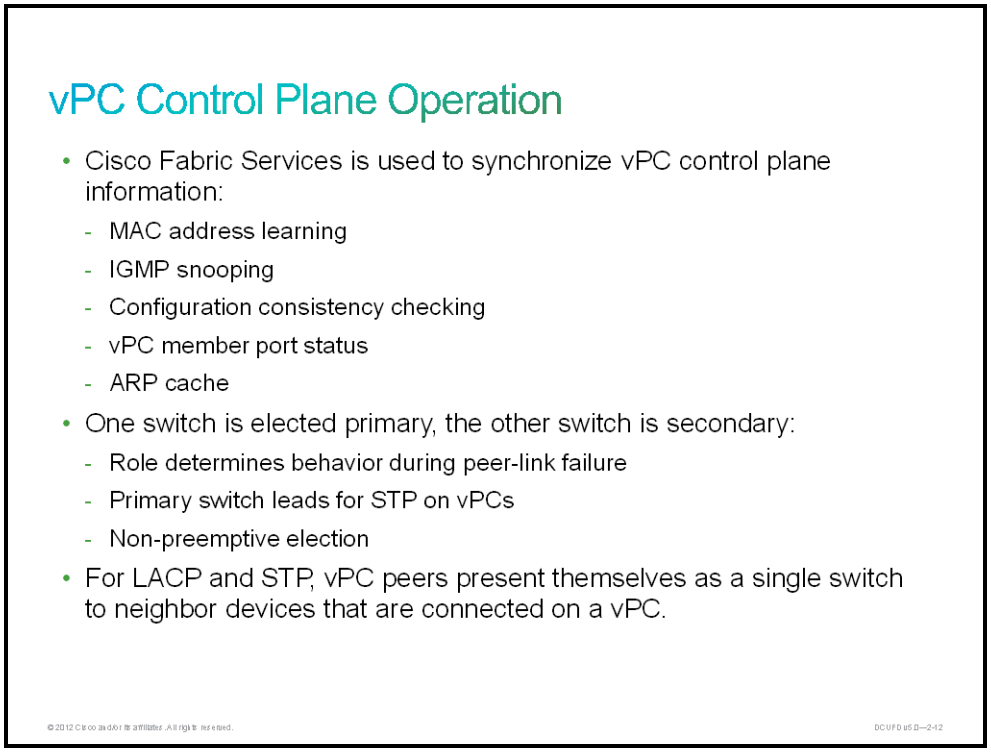


© 2012 Cisco and/or its affiliates. All rights reserved.

- **vPC VLAN:** This is one of the VLANs that is carried over the peer link and is used to communicate via vPC with a peer device.
- **Non-vPC VLAN:** This is one of the STP VLANs that is not carried over the peer link.
- **Orphan device:** This term refers to any device that is connected to a vPC domain using regular links instead of connecting through a vPC. A device that is connected to one vPC peer is considered an orphan device. VLANs that are configured on orphan devices cross the peer link.
- **Orphan port:** This term refers to a switch port that is connected to an orphan device. The term is also used for vPC ports whose members are all connected to a single vPC peer. This situation can occur if a device that is connected to a vPC loses all its connections to one of the vPC peers. An orphan port is a non-vPC interface on a switch where other ports in the same VLAN are configured as vPC interfaces.
- **Cisco Fabric Services:** The Cisco Fabric Services protocol is a reliable messaging protocol that is designed to support rapid stateful configuration message passing and synchronization. The vPC peers use the Cisco Fabric Services protocol to synchronize data plane information and implement necessary configuration checks. vPC peers must synchronize the Layer 2 Forwarding (L2F) table between the vPC peers. This way, if one vPC peer learns a new MAC address, that MAC address is also programmed on the L2F table of the other peer device. The Cisco Fabric Services protocol travels on the peer link and does not require any configuration by the user. To help ensure that the peer link communication for the Cisco Fabric Services over Ethernet protocol is always available, spanning tree keeps the peer-link ports always forwarding. The Cisco Fabric Services over Ethernet protocol is also used to perform compatibility checks to validate the compatibility of vPC member ports to form the channel, to synchronize the Internet Group Management Protocol (IGMP) snooping status, to monitor the status of the vPC member ports, and to synchronize the Address Resolution Protocol (ARP) table.

https://t.me/learningnets

certcollecion.net



Cisco Fabric Services is used as the primary control plane protocol for vPC. It performs several functions:

- vPC peers must synchronize the Layer 2 MAC address table between the vPC peers. If one vPC peer learns a new MAC address on a vPC, that MAC address is also programmed on the L2F table of the other peer device for that same vPC.
  This MAC address learning mechanism replaces the regular switch MAC address learning mechanism and prevents traffic from being forwarded across the vPC peer link unnecessarily.
- The synchronization of IGMP snooping information is performed by Cisco Fabric Services. L2F of multicast traffic with vPC is based on modified IGMP snooping behavior that synchronizes the IGMP entries between the vPC peers. In a vPC implementation, IGMP traffic that is entering a vPC peer switch through a vPC triggers hardware programming for the multicast entry on both vPC member devices.
- Cisco Fabric Services is also used to communicate essential configuration information to ensure configuration consistency between the peer switches. Similar to regular port channels, vPCs are subject to consistency checks and compatibility checks. During a compatibility check, one vPC peer conveys configuration information to the other vPC peer to verify that vPC member ports can form a port channel. In addition to compatibility checks for the individual vPCs, Cisco Fabric Services is also used to perform consistency checks for a set of switchwide parameters that must be configured consistently on both peer switches.
- Cisco Fabric Services is used to track the vPC status on the peer. When all vPC member ports on one of the vPC peer switches go down, Cisco Fabric Services is used to notify the vPC peer switch that its ports have become orphan ports and that traffic that is received on the peer link for that vPC should now be forwarded to the vPC.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

- Starting from Cisco Nexus Operating System (NX-OS) Software version 5.0(2) for the Cisco Nexus 5000 Switches and Cisco NX-OS Software version 4.2(6) for the Cisco Nexus 7000 Switches, Layer 3 vPC peers synchronize their respective ARP tables. This feature is transparently enabled and helps ensure faster convergence time upon reload of a vPC switch. When two switches are reconnected after a failure, they use Cisco Fabric Services to perform bulk synchronization of the ARP table.

Between the pair of vPC peer switches, an election is held to determine a primary and secondary vPC device. This election is non-preemptive. The vPC primary or secondary role is chiefly a control plane role that determines which of the two switches will be responsible for the generation and processing of spanning-tree BPDUs for the vPCs.

Note Starting from Cisco NX-OS Software version 4.2(6) for the Cisco Nexus 7000 Switches, the vPC peer-switch option can be implemented, which allows both the primary and secondary vPC device to generate BPDUs for vPCs independently. The two switches use the same spanning-tree bridge ID to ensure that devices that are connected on a vPC still see the vPC peers as a single logical switch.

Both switches actively participate in traffic forwarding for the vPCs. However, the primary and secondary roles are also important in certain failure scenarios, most notably in a peer-link failure. When the vPC peer link fails, but the vPC peer switches determine through the peer-keepalive mechanism that the peer switch is still operational, the operational secondary switch suspends all vPC member ports. The secondary device also shuts down all switch virtual interfaces (SVIs) that are associated with any VLANs that are configured as allowed VLANs for the vPC peer link.

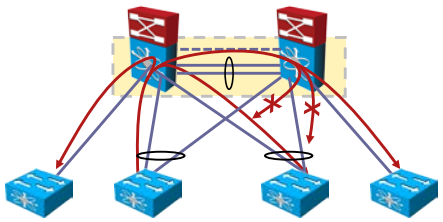
For LACP and STP, the two vPC peer switches present themselves as a single logical switch to devices that are connected on a vPC. For LACP, this is accomplished by generating the LACP system ID from a reserved pool of MAC addresses, combined with the vPC domain ID. For STP, the behavior depends on the use of the peer-switch option. If the peer-switch option is not used, the primary vPC is responsible for generating and processing BPDUs and uses its own bridge ID for the BPDUs. The secondary vPC relays BPDU messages, but does not generate BPDUs itself for the vPCs. When the peer-switch option is used, both the primary and secondary switches send and process BPDUs. However, they use the same bridge ID to present themselves as a single switch to devices that are connected on a vPC.

https://t.me/learningnets

certcollecion.net

vPC Data Plane Operation

- The vPC peer link carries the following traffic only:
  - vPC control traffic
  - Flooded traffic (broadcast, multicast, unknown unicast)
  - Traffic for orphan ports
- Regular switch MAC address learning is replaced with Cisco Fabric Services-based MAC address learning for vPCs:
  - Non-vPC ports use regular MAC address learning
- Frames that enter a vPC peer switch from the peer link cannot exit the switch on a vPC member port.



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

vPC is designed to limit the use of the peer link specifically to switch management traffic and the occasional traffic flow from a failed network port. The peer link does not carry regular traffic for vPCs. It carries only the traffic that needs to be flooded, such as broadcast, multicast, and unknown unicast traffic. It also carries traffic for orphan ports.

One of the most important forwarding rules for vPC is that a frame that enters the vPC peer switch from the peer link cannot exit the switch from a vPC member port. This principle prevents frames that are received on a vPC from being flooded back onto the same vPC by the other peer switch. The exception to this rule is traffic that is destined for an orphaned vPC member port.

© 2012 Cisco Systems, Inc.

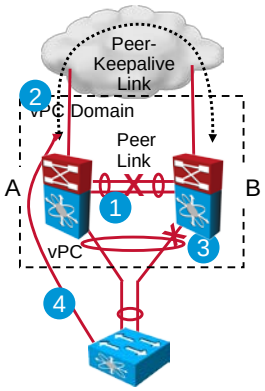
https://t.me/learningnets

certcollecion.net

vPC Failure Scenarios

vPC Peer-Link Failure

1. The vPC peer link on Switch A fails.
2. The software checks the status of the remote vPC on Peer B using the peer-keepalive link.
3. If the vPC on Peer B is up, the secondary vPC on Peer B disables all vPC ports on its device to prevent loops and black-holing or flooding traffic.
4. The data then forwards down the remaining active links of the port channel.



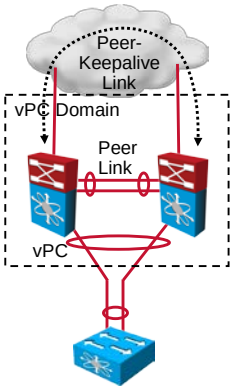
© 2012 Cisco and/or its affiliates. All rights reserved.

If the vPC peer link fails, the software checks the status of the remote vPC peer device using the peer-keepalive link, which is a link between vPC peer devices that ensures that both devices are up. If the vPC peer device is up, the secondary vPC device disables all vPC ports on its device, to prevent loops and disappearing or flooding traffic. The data then forwards down the remaining active links of the port channel.

vPC Failure Scenarios (Cont.)

vPC Peer Failure—Peer-Keepalive Link

- The software learns of a vPC peer device failure when the keepalive messages are not returned over the peer-keepalive link.
- Use a separate link (vPC peer-keepalive link) to send configurable keepalive messages between the vPC peer devices. The keepalive messages on the vPC peer-keepalive link determine whether a failure is on the vPC peer link only or on the vPC peer device. The keepalive messages are used only when all the links in the peer link fail.



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

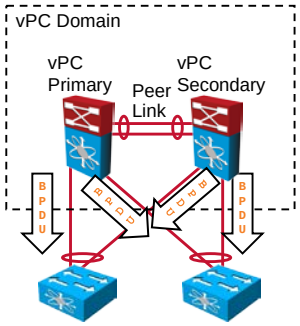
This figure shows vPC peer-keepalive link usage.

https://t.me/learningnets

certcollecion.net

vPC Peer Switch

- A pair of Cisco Nexus 7000 Series devices appears as a single STP root in the Layer 2 topology.
- STP BPDUs are sent on both the vPC legs to avoid issues that are related to STP BPDU timeout on the downstream switches, which can cause traffic disruption.
- It eliminates the recommendation to increase STP hello time on the vPC-pair switches.



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

The vPC peer-switch feature was added to Cisco NX-OS Release 5.0(2) to address performance concerns around STP convergence. This feature allows a pair of Cisco Nexus 7000 Series devices to appear as a single STP root in the Layer 2 topology. This feature eliminates the need to pin the STP root to the vPC primary switch and improves vPC convergence if the vPC primary switch fails.

To avoid loops, the vPC peer link is excluded from the STP computation. In vPC peer-switch mode, STP BPDUs are sent from both vPC peer devices to avoid issues that are related to STP BPDU timeout on the downstream switches, which can cause traffic disruption.

This feature can be used with these topologies:

- The pure peer-switch topology in which the devices all belong to the vPC
- The hybrid peer-switch topology in which there is a mixture of vPC and non-vPC devices in the configuration

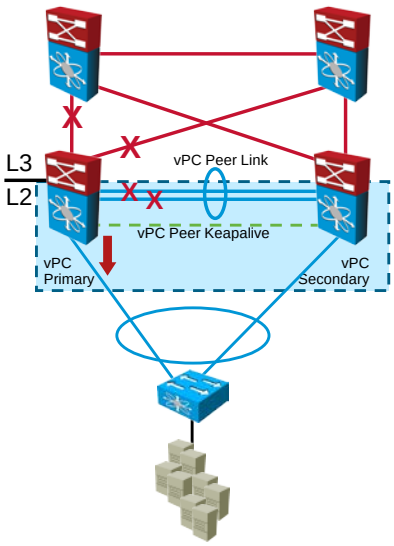
© 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

vPC Object Tracking

- Providing flexible behavior under failover conditions
- Tracking state of links of a vPC peer device
- Peer link and core interfaces can be tracked as a list of Boolean objects.
- vPC object tracking suspends vPCs on the impaired device so that traffic can be diverted over the remaining vPC peer.



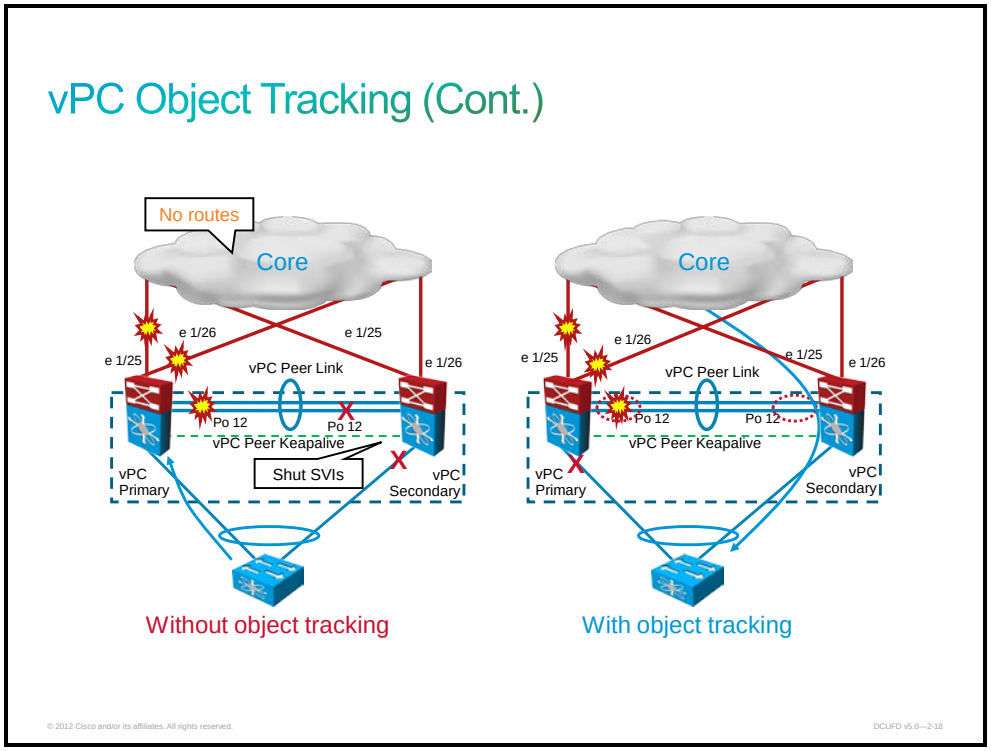
© 2012 Cisco and/or its affiliates. All rights reserved.

Use this configuration to avoid dropping traffic if a particular module goes down, because when all the tracked objects on the track list go down, the system does the following:

- Stops the vPC primary peer device from sending peer-keepalive messages, which forces the vPC secondary peer device to take over
- Brings down all the downstream vPCs on that vPC peer device, which forces all the traffic to be rerouted in the access switch toward the other vPC peer device

https://t.me/learningnets

certcollecion.net



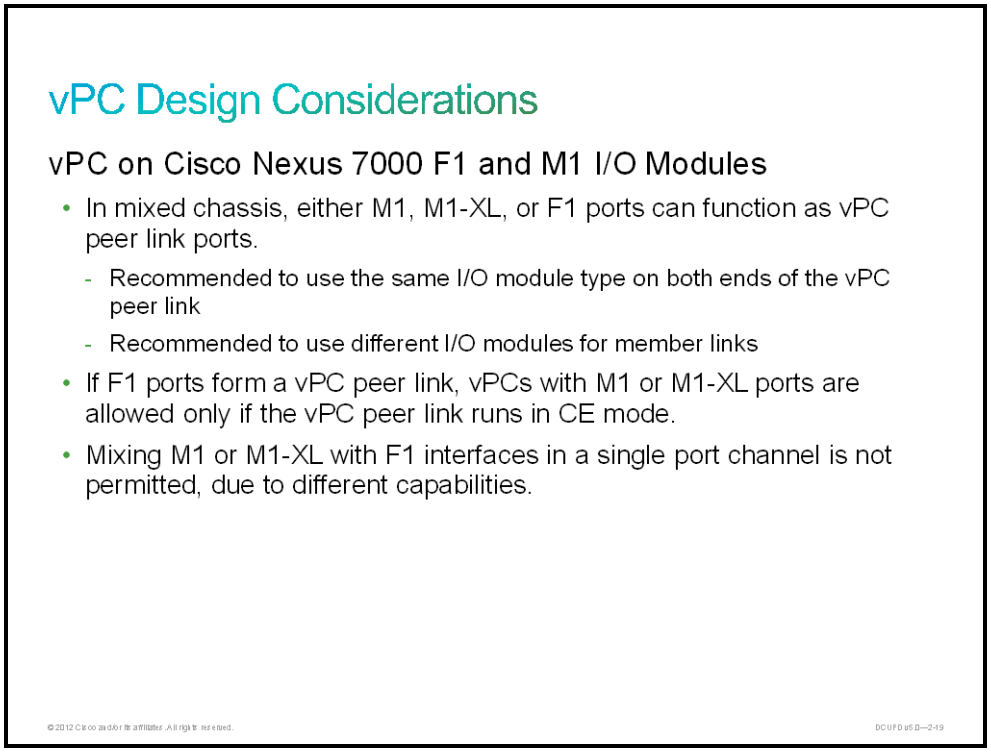
After you configure this feature, and if the module fails, the system automatically suspends all the vPC links on the primary vPC peer device and stops the peer-keepalive messages. This action forces the vPC secondary device to take over the primary role and all the vPC traffic goes to this new vPC primary device until the system stabilizes.

Create a track list that contains all the links to the core and all the vPC peer links as its object. Enable tracking for the specified vPC domain for this track list. Apply this same configuration to the other vPC peer device.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net



When implementing a vPC on Cisco Nexus 7000 Series switches that are populated with F1 and M1 I/O modules, there are some design issues to consider:

- In mixed chassis, M1, M1-XL, or F1 ports can function as vPC peer-link ports.
- You must use ports from the same module type on each side of the vPC peer link (all M1 or all F1 ports on each side of the vPC peer link).
- It is recommended that you use multiple I/O modules for member links.
- If F1 ports form the vPC peer link, vPCs with M1 or M1-XL ports are allowed only if the vPC peer link runs in Classical Ethernet (CE) mode.
- Mixing M1 or M1-XL and F1 interfaces in a single port channel is not allowed due to different capabilities.

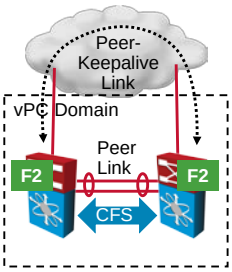
https://t.me/learningnets

certcollecion.net

vPC Design Considerations (Cont.)

vPC on Cisco Nexus 7000 F2 I/O Modules

- vPC remains the same except for the following:
  - Peer Link
    - The peer link on F2 I/O modules needs identical F2 modules on both sides.
  - Multicast
    - F2 vPC cannot support a dual DR for Layer 3 multicast.



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

When implementing vPC on Cisco Nexus 7000 switches that are populated with F2 I/O modules, vPC remains the same except for the following:

- **Peer link:** The vPC peer link on F2 I/O modules needs identical F2 modules on both sides.
- **Multicast:** F2 vPC cannot support a dual designated router (DR) for Layer 3 multicast.

This table lists the support that is available for vPC peer link and vPC interfaces for Cisco Nexus 7000 I/O modules.

Support for vPC Peer Link and vPC Interfaces for Cisco Nexus 7000 I/O Modules

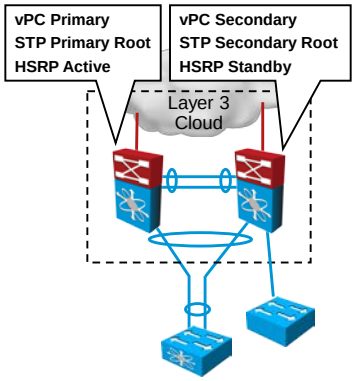
I/O Module	vPC Peer-Link	vPC Interfaces
N7K-M108X2-12L	Yes	Yes
N7K-M132XP-12 N7K-M132XP-12L	Yes	Yes
N7K-M148GT-11 N7K-M148GT-11L	No	Yes
N7K-M148GS-11 N7K-M148GS-11L	No	Yes
N7K-F132XP-15 N7K-F248XP-25	Yes	Yes

https://t.me/learningnets

certcollecion.net

Spanning Tree Design Recommendations

- Configure aggregation vPC peers as root and secondary root.
  - If the vPC peer-switch is implemented, both vPC peers will behave as a single STP root.
- Align the STP primary root, HSRP active router, and Protocol Independent Multicast (PIM) DR with the vPC primary peer.
- Bridge assurance is enabled by default on the vPC peer link.
- Do not enable loop guard and bridge assurance on the vPC. They are disabled by default.
- Enable STP port type “edge” and port type “edge trunk” on host ports.
- Enable STP BPDU-guard globally.
- Disable STP channel-misconfig guard if it is supported by the access switches.



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

You must manually configure the following features to conform to the primary and secondary mapping of each of the vPC peer devices:

- **STP Root:** Configure the primary vPC peer device as the STP primary root device and configure the vPC secondary device to be the STP secondary root device.
- **Hot Standby Router Protocol (HSRP) Active/Standby:** If you want to use HSRP and VLAN interfaces on the vPC peer devices, configure the primary vPC peer device with the HSRP active highest priority. Configure the secondary device to be the HSRP standby.
- **vPC Peer-Gateway** The vPC peer-gateway capability allows a vPC switch to act as the active gateway for packets that are addressed to the router MAC address of the vPC peer. This feature enables local forwarding of such packets without the need to cross the vPC peer link. In this scenario, the feature optimizes use of the peer link and avoids potential traffic loss.

  Configuring the peer-gateway feature must be done on both primary and secondary vPC peers and is nondisruptive to the operations of the device or to the vPC traffic. The vPC peer-gateway feature can be configured globally under the vPC domain submode.
- **vPC Peer-Switch:** The vPC peer-switch feature was added to Cisco NX-OS Release 5.0(2) to address performance concerns around STP convergence. This feature allows a pair of Cisco Nexus 7000 Series devices to appear as a single STP root in the Layer 2 topology. This feature eliminates the need to pin the STP root to the vPC primary switch and improves vPC convergence if the vPC primary switch fails.

  To avoid loops, the vPC peer link is excluded from the STP computation. In vPC peer-switch mode, STP BPDUs are sent from both vPC peer devices to avoid issues that are related to STP BPDU timeout on the downstream switches, which can cause traffic disruption.

Note	Layer 3 adjacencies cannot be formed over a vPC or over a vPC peer link because vPC is a Layer 2-only connection. To bring up a routing protocol adjacency with a peer switch, provision an additional Layer 3 link.

https://t.me/learningnets

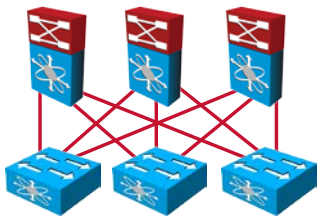
certcollecion.net

Cisco FabricPath

This topic describes how to design solutions using Cisco FabricPath.

Why Cisco FabricPath?

- Cisco FabricPath allows easy scaling of Layer 2 domains in data centers.
- Why Layer 2 domains?
  - Application and protocol requirements
  - Simple implementation
  - Easy server provisioning
  - Allows VM mobility
- Cisco FabricPath provides the following:
  - "Plug-and-play" implementation
  - No addressing required
  - Scalable domain with automatic multipathing
  - Load balancing
  - Redundancy



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

Cisco FabricPath is a technology that provides additional scalability and simplification of an Ethernet network. It also provides more efficient forwarding and eliminates the need for STP.

Cisco FabricPath is an innovative technology that is supported in Cisco NX-OS and brings the benefits of Layer 3 forwarding to Layer 2 networks.

In modern data centers, there is demand for Layer 2 domains to grow larger. Large Layer 2 domains allow for easy server provisioning and virtual machine mobility. They also accommodate requirements of clustered applications.

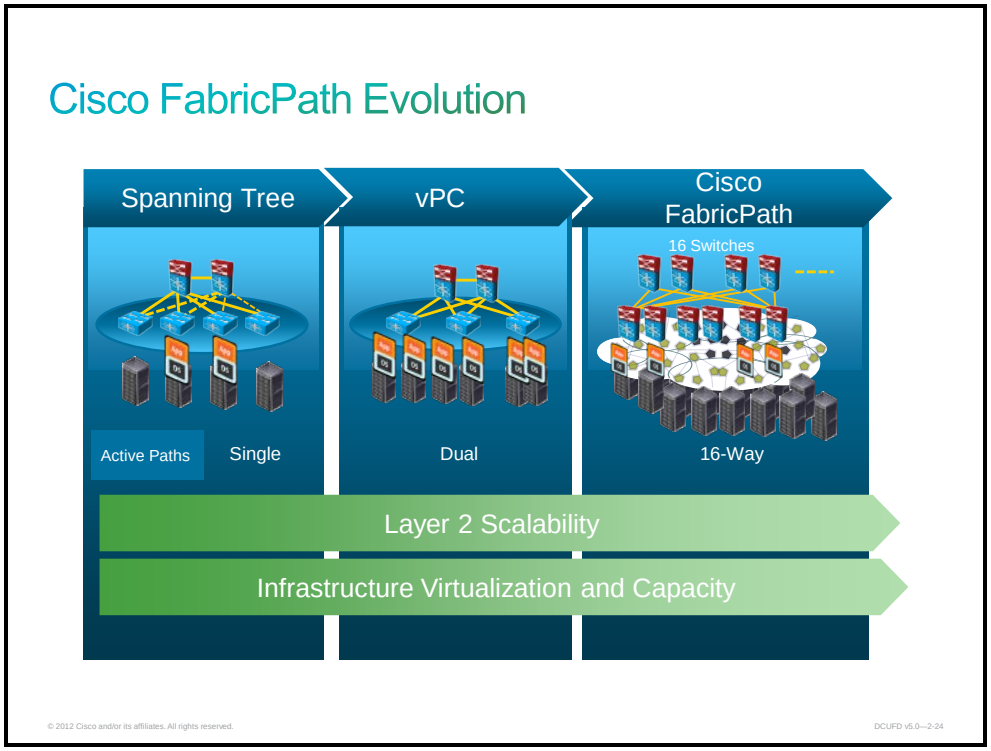
Layer 2 domains are easy to implement and do not require any addressing to be preconfigured. However, there were limitations due to volumes of broadcast traffic and overutilization of MAC table ternary content addressable memory (TCAM) resources.

Cisco FabricPath is a technology that scales easily, prevents Layer 2 loops, protects Layer 2 TCAM resources, and provides for automatic load balancing and redundancy.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net



Current data center designs are a compromise between the flexibility that is provided by Layer 2 and the scaling that is offered by Layer 3.

The first generation of Layer 2 networks was run by STP. Its role was to provide a stable network, to block links that would form Layer 2 loops, and to unblock them in case there was a change in the topology. However, there were drawbacks to networks that were based on STP:

- **Limited scale:** Layer 2 provides flexibility but cannot scale. Bridging domains are therefore restricted to small areas, strictly delimited by Layer 3 boundaries.
- **Suboptimal performance:** Traffic forwarding within a bridged domain is constrained by spanning-tree rules, limiting bandwidth and enforcing inefficient paths between devices.
- **Complex operation:** Layer 3 segmentation makes data center designs static and prevents them from matching the business agility that is required by the latest virtualization technologies. Any change to the original plan is complicated and configuration is intensive and disruptive.

The second-generation data center provides the ability to use all links in the LAN topology by taking advantage of technologies such as vPCs.

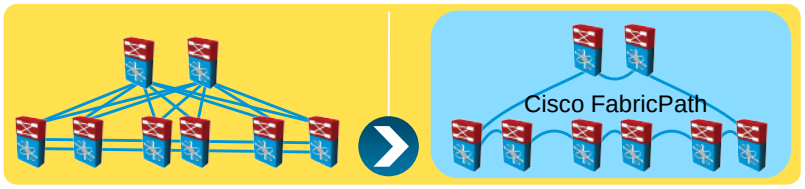
Cisco FabricPath technology on the Cisco Nexus 7000 Series switches and on the Cisco Nexus 5500 Series switches introduces new capabilities and design options that allow network operators to create Ethernet fabrics that increase bandwidth availability, provide design flexibility, and simplify and reduce the costs of network and application deployment and operation.

https://t.me/learningnets

certcollecion.net

Cisco FabricPath—An Ethernet Fabric

- Connect a group of switches using an arbitrary topology and aggregate them into a fabric.
- An open protocol, based on Layer 3 technology, that provides fabric-wide intelligence and ties the elements together.



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

Cisco FabricPath is an innovative Cisco NX-OS feature that is designed to bring the stability and performance of routing to Layer 2. It brings the benefits of Layer 3 routing to Layer 2 switched networks to build a highly resilient and scalable Layer 2 fabric.

Cisco FabricPath is simple to configure. The only necessary configuration consists of distinguishing the core ports (which link the switches) from the edge ports (where end devices are attached). There is no need to tune any parameter for an optimal configuration, and switch addresses are assigned automatically.

Cisco FabricPath uses a control protocol in addition to the powerful Intermediate System-to-Intermediate System (IS-IS) routing protocol, an industry standard that provides fast convergence and that has been proven to scale up to the largest service provider environments.

A single control protocol is used for unicast forwarding, multicast forwarding, and VLAN pruning. The Cisco FabricPath solution requires less combined configuration than an equivalent STP-based network, further reducing the overall management cost.

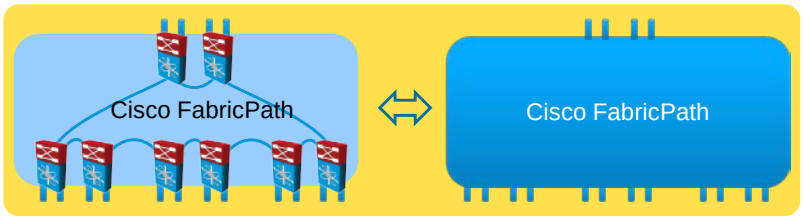
© 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

Cisco FabricPath—An Ethernet Fabric (Cont.)

- Externally, from a CE perspective, a Cisco FabricPath fabric looks like a single switch.
  - Presents itself as an STP root bridge
- Internally, a protocol adds fabric-wide intelligence and ties the elements together. This protocol provides the following in a plug-and-play fashion:
  - Optimal, low latency connectivity that is any-to-any
  - High bandwidth, high resiliency
  - Open management and troubleshooting
- Cisco FabricPath provides additional capabilities in terms of scalability and Layer 3 integration.



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

Cisco FabricPath delivers the foundation for building a scalable fabric—a network that itself looks like a single virtual switch from the perspective of its users. This property is achieved by providing optimal bandwidth between any two ports, regardless of their physical locations.

Also, because Cisco FabricPath does not suffer from the scaling restrictions of traditional transparent bridging, a particular VLAN can be extended across the whole fabric, reinforcing the perception of a single virtual switch.

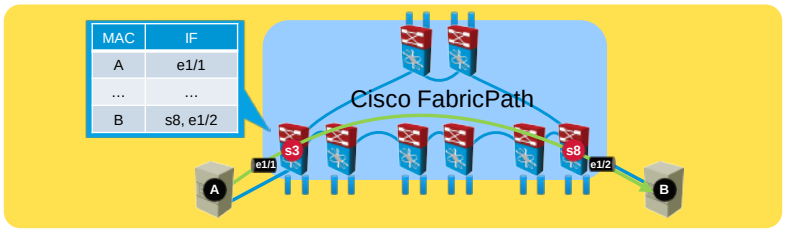
Cisco FabricPath takes control as soon as an Ethernet frame transitions from an Ethernet network (referred to as Classical Ethernet) to a Cisco FabricPath fabric. Ethernet bridging rules do not dictate the topology and the forwarding principles in a Cisco FabricPath fabric. The frame is encapsulated with a Cisco FabricPath header, which consists of routable source and destination addresses. These addresses are simply the address of the switch on which the frame was received and the address of the destination switch to which the frame is heading. From there on, the frame is routed until it reaches the remote switch, where it is de-encapsulated and delivered in its original Ethernet format.

https://t.me/learningnets

certcollecion.net

Cisco FabricPath—An Ethernet Fabric (Cont.)

- Shortest path, any-to-any.
- Single address lookup at the ingress edge identifies the exit port across the fabric.
- Traffic is then switched using the shortest path available.
- Reliable Layer 2 and Layer 3 connectivity, any-to-any.
  (Layer 2 as if it was within the same switch, with no STP inside.)

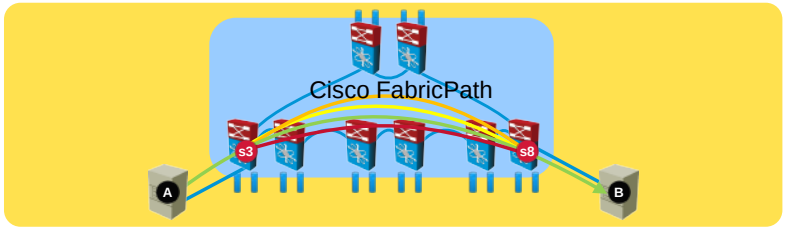


© 2012 Cisco and/or its affiliates. All rights reserved.

Frames are forwarded along the shortest path to their destination, reducing the latency of the exchanges between end stations when compared to a spanning tree-based solution.

Cisco FabricPath—An Ethernet Fabric (Cont.)

- ECMP
- Multipathing (up to 256 links active between any two devices)
- Traffic is redistributed across remaining links in case of failure, providing fast convergence.
- Conversational learning—per-port MAC address table only needs to learn the peers that are reached across the fabric.



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

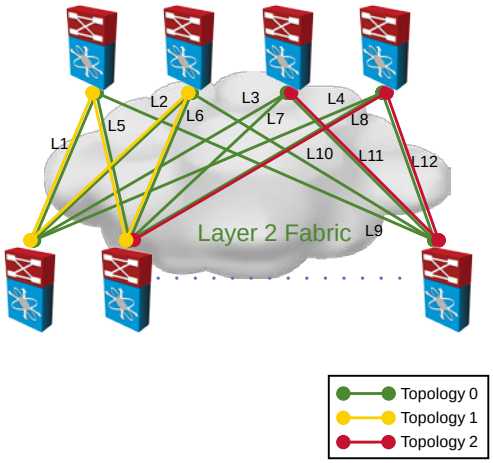
Because Equal-Cost Multipath (ECMP) can be used in the data plane, the network can use all the links that are available between any two devices. The first-generation hardware supporting Cisco FabricPath can perform 16-way ECMP, which, when combined with 16-port 10-Gb/s port channels, represents a potential bandwidth of 2.56 terabits per second (Tb/s) between switches.

https://t.me/learningnets

certcollecion.net

Multitopology Cisco FabricPath

- Topology: A group of links in the fabric. By default, all links are part of Topology 0.
- Other topologies can be created by assigning a subset of the links to them.
- A link can belong to several topologies.
- A VLAN is mapped to a unique topology.
- Topologies can be used for traffic engineering, security, and so on.



L1 to L12 = Layer 1 to Layer 12

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

By default, Cisco FabricPath fabrics have only one logical topology that they use—Topology 0. You can create additional topologies by assigning links to these topologies. Assigning links to multiple topologies allows you to assign traffic to dedicated paths if that is required.

Links are assigned to topologies. You can have a link carrying traffic for multiple topologies.

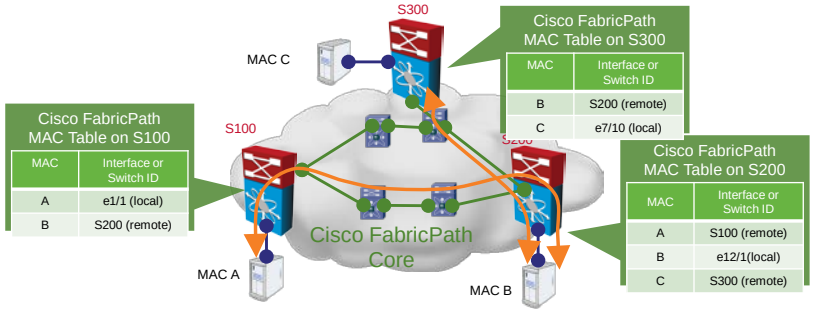
VLANs are assigned to a topology. You can perform traffic engineering (“manual” load balancing) on the fabric.

https://t.me/learningnets

certcollecion.net

Control Plane Operation: Conversational MAC Address Learning

- The MAC learning method is designed to conserve MAC table entries on Cisco FabricPath edge switches.
- Each forwarding engine distinguishes between two types of MAC entry: local and remote.
- The forwarding engine learns the remote MAC only if a bidirectional conversation occurs between the local and remote MAC.



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

Cisco FabricPath Control Plane Operation

Conversational MAC Address Learning

Conversational MAC address learning means that each interface learns only those MAC addresses for interested hosts, rather than all MAC addresses in the domain. Each interface learns only those MAC addresses that are actively speaking with the interface.

In traditional MAC address learning, each host learns the MAC address of every other device on the network. With Cisco FabricPath, not all interfaces have to learn all the MAC addresses on an F-Series module, which greatly reduces the size of the MAC address tables.

Beginning with Cisco NX-OS Release 5.1 and using the N7K-F132XP-15 module, the MAC learning process is optimized. Conversational MAC learning is configured per VLAN. All outer destination address Cisco FabricPath VLANs always use conversational learning. You can configure CE VLANs for conversational learning on this module as well.

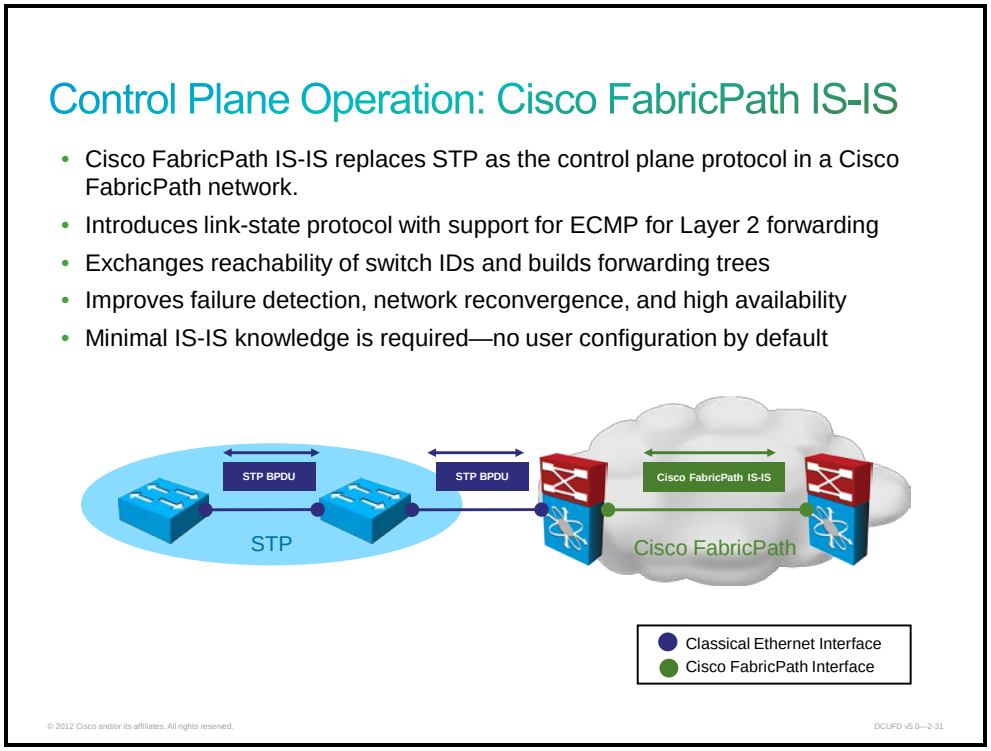
The N7K-F132XP-15 module has 16 forwarding engines. The MAC learning process takes place on only one of them. Each forwarding engine performs MAC address learning independently of the other 15 forwarding engines on the module. An interface only maintains a MAC address table for the MACs that ingress or egress through that forwarding engine. The interface does not have to maintain the MAC address tables on the other 15 forwarding engines on the module.

The Cisco Nexus 5500 Series switches also have Cisco FabricPath support as of Cisco NX-OS version 5.1(3)N1(1).

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net



Cisco FabricPath IS-IS

With Cisco FabricPath, you use the Layer 2 IS-IS protocol for a single control plane that functions for unicast, broadcast, and multicast packets. There is no need to run STP. It is a purely Layer 2 domain. Cisco FabricPath Layer 2 IS-IS is a separate process from Layer 3 IS-IS.

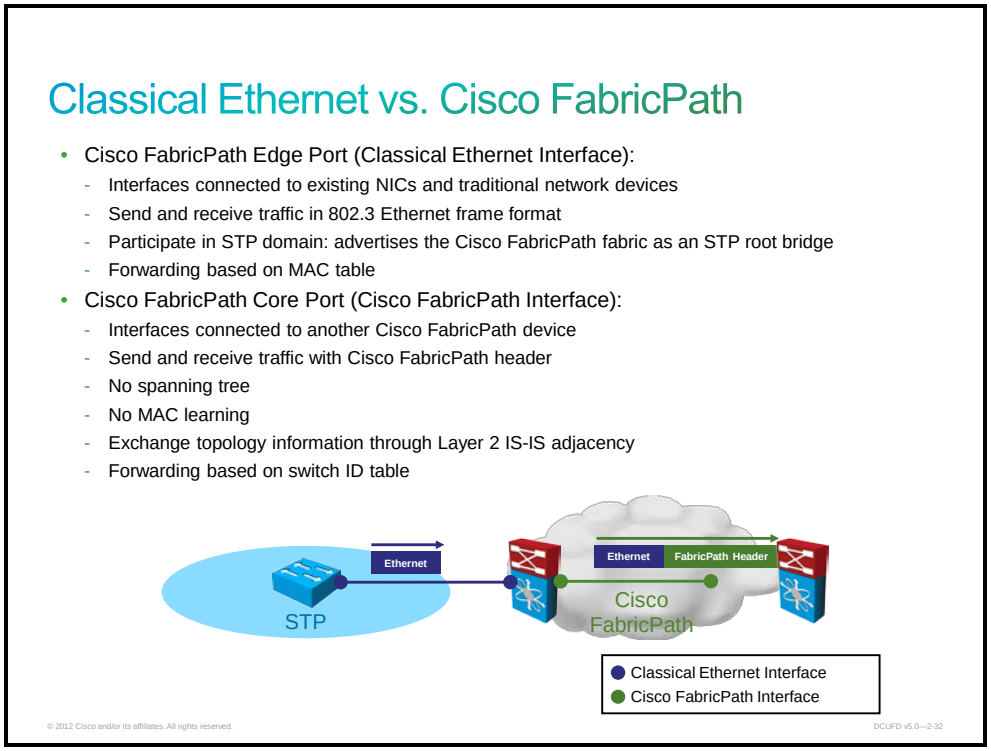
IS-IS provides the following benefits:

- **No IP dependency:** There is no need for IP reachability in order to form adjacency between devices.
- **Easily extensible:** Using custom types, lengths, values (TLVs), IS-IS devices can exchange information about virtually anything.
- **Shortest Path First (SPF) routing:** This provides superior topology building and reconvergence characteristics.

Every switch must have a unique source ID (SID) to participate in the Cisco FabricPath domain. A new switch initially selects a random SID and checks to see if that value is already in use. Although the Cisco FabricPath network automatically verifies that each switch has a unique SID, a configuration command is provided for the network administrator to statically assign a SID to a Cisco FabricPath switch. If you choose to manually configure SIDs, be certain that each switch has a unique value because any switch with a conflicting SID will suspend data plane forwarding on Cisco FabricPath interfaces while the conflict exists.

https://t.me/learningnets

certcollecion.net



Every interface that is involved in Cisco FabricPath switching falls into one of two categories:

- **Cisco FabricPath edge port:** Cisco FabricPath edge ports are interfaces at the edge of the Cisco FabricPath domain. These interfaces run Classical Ethernet and behave exactly like normal Ethernet ports. You can attach any Classical Ethernet device to the Cisco FabricPath fabric by connecting it to a Cisco FabricPath edge port. Cisco FabricPath switches perform MAC address learning on edge ports, and frames that are transmitted on edge ports are standard IEEE 802.3 Ethernet frames. You can configure an edge port as an access port or as an IEEE 802.1Q trunk.

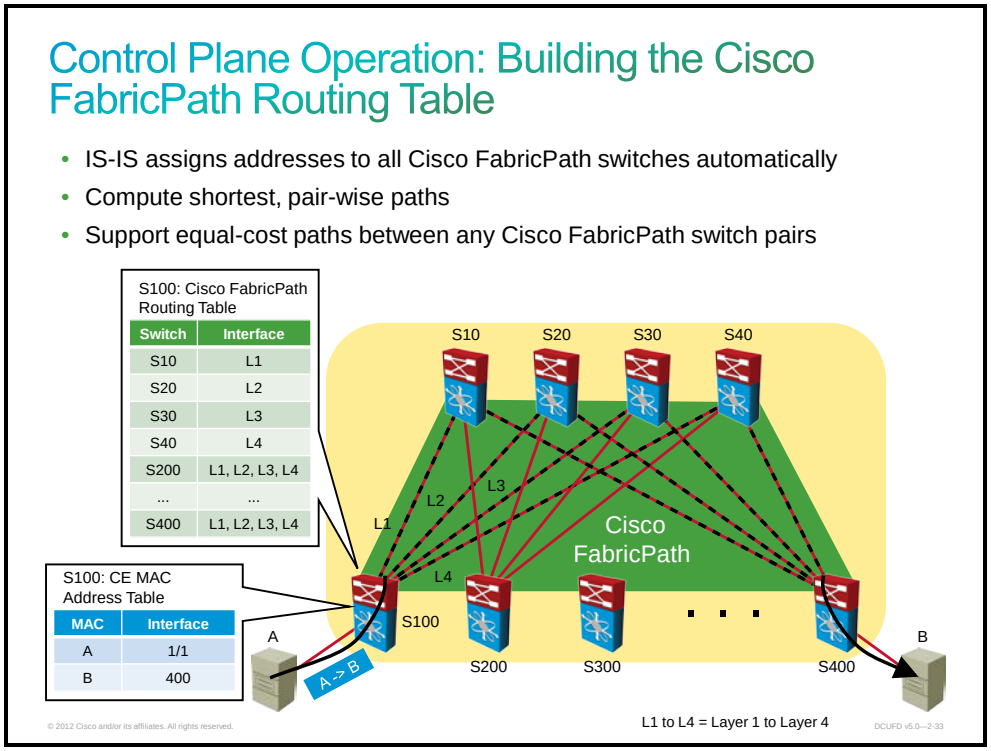
  The whole Cisco FabricPath fabric appears as a spanning-tree root bridge toward the Classical Ethernet cloud. The switch generates a BPDU on the CE interface, putting the Cisco FabricPath cloud at the top of the STP tree.

- **Cisco FabricPath core port:** Cisco FabricPath core ports always forward Ethernet frames that are encapsulated in a Cisco FabricPath header. Generally, no MAC address learning occurs on Cisco FabricPath core ports. Forwarding decisions occur based exclusively on lookups in the switch table. Ethernet frames that are transmitted on a Cisco FabricPath interface always carry an IEEE 802.1Q tag and, therefore, the port can conceptually be considered a trunk port.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net



Building the Cisco FabricPath Routing Table

The protocol used to establish the routed topology is a modified version of IS-IS.

The IS-IS protocol is easily extensible and does not require any IP configuration to discover the topology and determine shortest-path trees. IS-IS automatically assigns addressing and switch names, and computes shortest paths between any two switches in the Cisco FabricPath cloud.

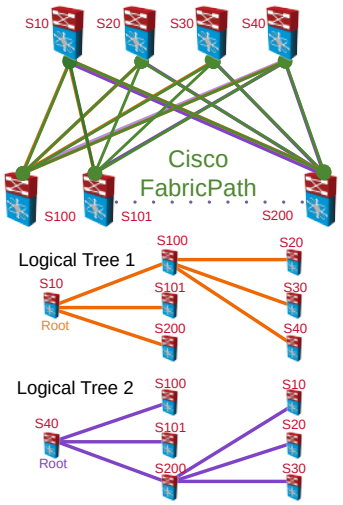
If multiple paths are available between two switches, IS-IS installs both routes in the Cisco FabricPath routing table and performs ECMP between these two switches.

https://t.me/learningnets

certcollecion.net

Control Plane Operation: Multidestination Trees

- Multidestination traffic is constrained to loop-free trees that are touching all Cisco FabricPath switches.
- A root switch is assigned for each multidestination tree in the Cisco FabricPath domain.
- A loop-free tree is built from each root and assigned a network-wide identifier (the FTag).
- Support for multiple multidestination trees provides multipathing for multidestination traffic.



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

Multidestination Trees

Cisco FabricPath introduces a new loop-free broadcast functionality that carries broadcast, unknown unicast, and multicast packets, or multidestination traffic. For each broadcast, unknown unicast, and multicast traffic flow, the system chooses the forwarding path from among multiple system-created paths or trees.

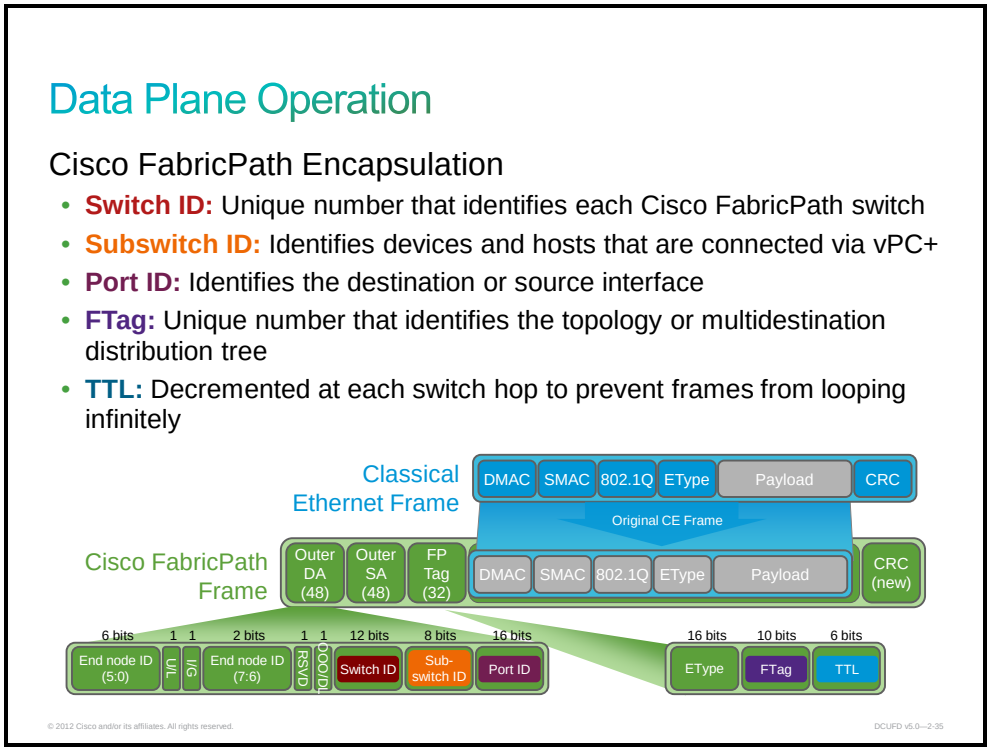
Note	For Cisco NX-OS Release 5.1, the system creates two trees to forward the multidestination traffic for each topology.

For the Cisco FabricPath network, the system creates a broadcast tree that carries broadcast traffic, unknown unicast traffic, and multicast traffic through the Cisco FabricPath network. The system also creates a second tree and all the multicast traffic flows are load-balanced across these two trees for each flow. Each tree is identified in the Cisco FabricPath network by a unique value, or forwarding tag (FTag). Within the Cisco FabricPath network, the system elects a root node that becomes the root for the broadcast tree. That node also identifies another bridge to become the root for the second multidestination tree, which load-balances the multicast traffic.

Note	Cisco FabricPath accommodates multiple topologies. For every topology, multidestination trees can be created, depending on the size of the topology and the links that are available.

https://t.me/learningnets

certcollecion.net



Cisco FabricPath Encapsulation

Cisco FabricPath encapsulation uses a MAC address-in-MAC address encapsulation format. The original Ethernet frame, along with an IEEE 802.1Q tag, is prepended by a 48-bit outer source address, a 48-bit outer destination address, and a 32-bit Cisco FabricPath tag. While the outer source address and destination address may appear as 48-bit MAC addresses, Cisco FabricPath switches that are receiving such frames on a Cisco FabricPath core port parse these fields according to the format shown in this figure.

The fields of the Cisco FabricPath header are described here:

- **End node ID:** As of Cisco NX-OS Release 5.2(1), the end node ID field is not used by the Cisco FabricPath implementation. However, the presence of this field may provide the future capability for an end station that is enabled for Cisco FabricPath to uniquely identify itself, allowing forwarding decisions based on Cisco FabricPath down to the virtual or physical end-station level.
- **Universal/Local (U/L) Bit:** Cisco FabricPath switches set the U/L bit in all unicast outer source address and destination address fields, indicating that the MAC address is locally administered (rather than universally unique). This setting is required because the outer source address and destination address fields are not, in fact, MAC addresses and do not uniquely identify a particular hardware component as a standard MAC address would.
- **Individual/Group (I/G) Bit:** The I/G bit serves the same function in Cisco FabricPath as in standard Ethernet, determining whether the address is an individual address or a group address. All multidestination addresses have this bit set.
- **Out of Order/Does Not Learn (OOO/DL) Bit:** The function of the OOO/DL bit varies depending on whether the bit is set in the outer destination address (OOO) field or the outer source address (DL) field. As of Cisco NX-OS Release 5.2(1), this bit is not used in Cisco FabricPath implementation. However, the presence of this field may provide the future capability for per-packet load sharing when ECMP paths are available.

https://t.me/learningnets

certcollecion.net

- **Switch ID:** Every switch in the Cisco FabricPath domain is assigned a unique 12-bit switch ID. In the outer source address, this field identifies the Cisco FabricPath switch that originated the frame (typically the ingress Cisco FabricPath edge switch). In the outer destination address, this field identifies the destination Cisco FabricPath switch.

  For multidestination frames, the value in the outer destination address is set to a specific value depending on the type of multidestination frame:

  — For multicast frames, this field is populated with the corresponding bits of the destination MAC address field of the original (encapsulated) Ethernet frame.

  — For unknown unicast frames, this field is populated with the corresponding bits of a reserved multicast address (01:0F:FF:C1:01:C0).

  — For frames with a known inner destination MAC address but an unknown source, this field is populated with the corresponding bits of a reserved multicast address (01:0F:FF:C2:02:C0) to facilitate MAC address table updates on Cisco FabricPath edge switches.

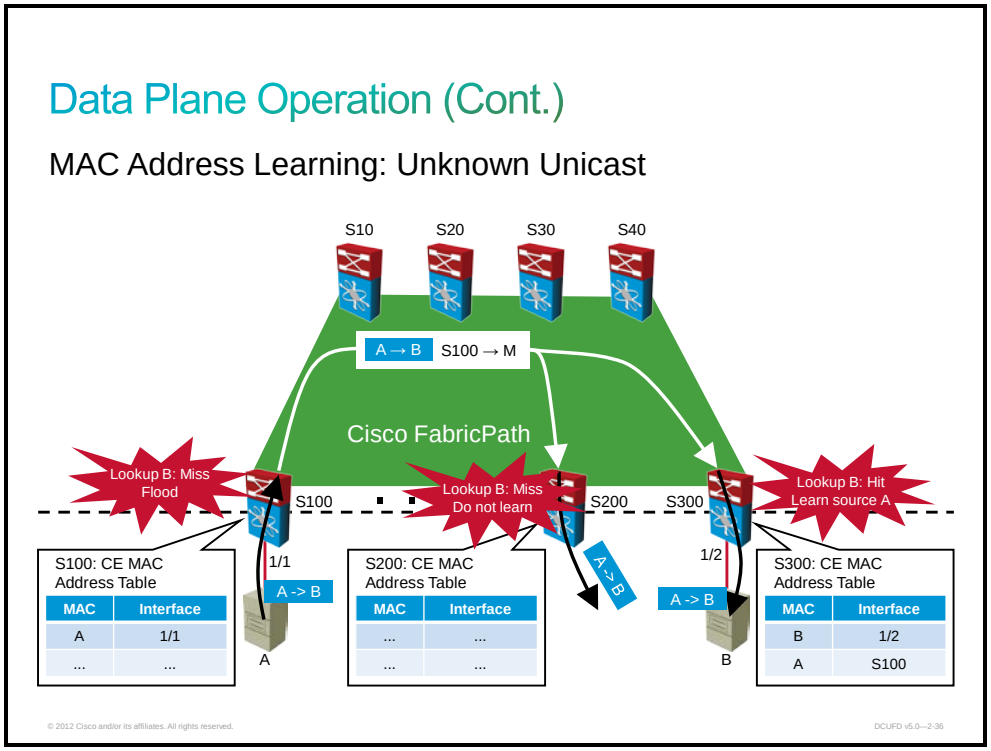
- **Subswitch ID:** The subswitch ID field identifies the source or destination virtual port channel plus (vPC+) interface that is associated with a particular vPC+ switch pair. Cisco FabricPath switches running vPC+ use this field to identify the specific vPC+ port channel on which traffic is to be forwarded. The subswitch ID value is locally significant to each vPC+ switch pair. In the absence of vPC+, this field is set to 0.
- **Port ID:** The port ID, also known as the local identifier (local ID), identifies the specific physical or logical interface on which the frame was sourced or to which it is destined. The value is locally significant to each switch. This field in the outer destination address allows the egress Cisco FabricPath switch to forward the frame to the appropriate edge interface without requiring a MAC address table lookup. For frames that are sourced from or destined to a vPC+ port channel, this field is set to a common value that is shared by both vPC+ peer switches, and the subswitch ID is used to select the outgoing port instead.
- **EtherType (EType):** The EType value for Cisco FabricPath encapsulated frames is 0x8903.
- **FTag:** The function of the FTag depends on whether a particular frame is unicast or multidestination. With unicast frames, the FTag identifies the Cisco FabricPath topology that the frame is traversing. As of Cisco NX-OS Release 5.2(1), only a single topology is supported, and this value is always set to 1. With multidestination frames, the FTag identifies the multidestination forwarding tree that the frame should traverse.

Note	Instead of a loop prevention mechanism, Cisco FabricPath uses the Time to Live (TTL) field in the frame to prevent unlimited frame flooding and, therefore, a loop.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net



Conversational MAC Address Learning

Conversational MAC address learning means that each interface learns only those MAC addresses for interested hosts, rather than all MAC addresses in the domain. Each interface learns only those MAC addresses that are actively speaking with the interface.

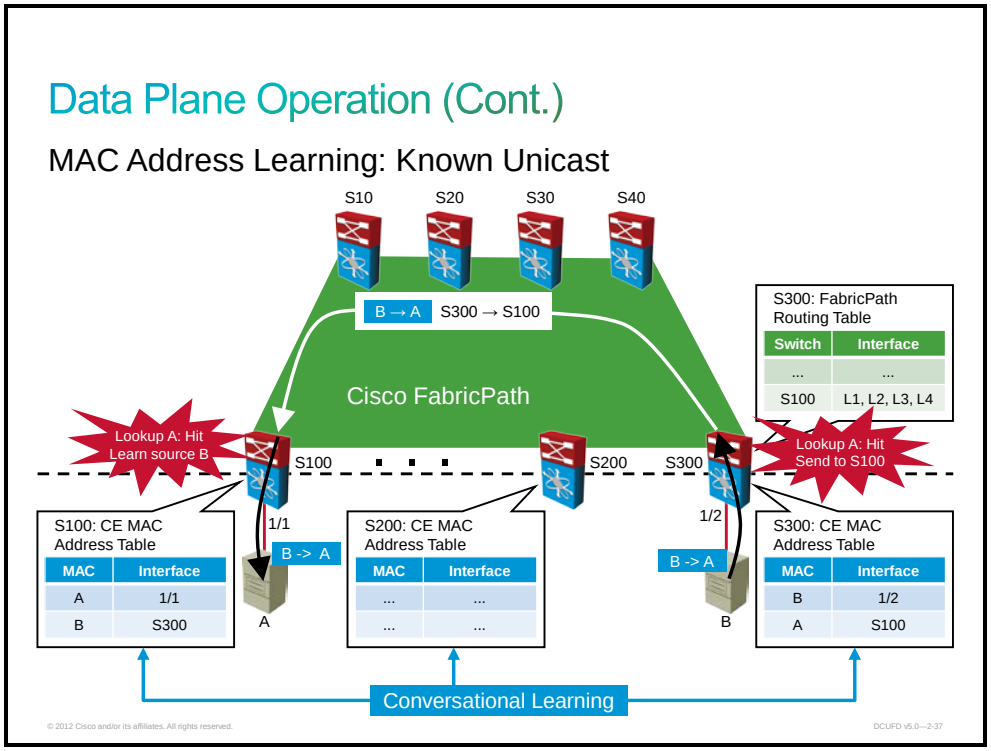
The switch that receives the frame (S100), with an unknown MAC address as the destination, will flood the frame out of all ports in the domain.

The switch that does not have the destination MAC address in its MAC address table (S200) will simply disregard the flooded frame. It will *not* learn the source address. This way, MAC address table resources are conserved.

If the switch has the destination MAC address in its local MAC address table (S300), it will forward the frame and learn the source MAC address of the frame.

https://t.me/learningnets

certcollecion.net



After the switches have learned all relevant pairs of MAC addresses, they forward the frames based on the information that they have in their MAC address tables. Remote MAC addresses have the name of the switch to which they are attached instead of the upstream interface as in Classical Ethernet.

 © 2012 Cisco Systems, Inc.

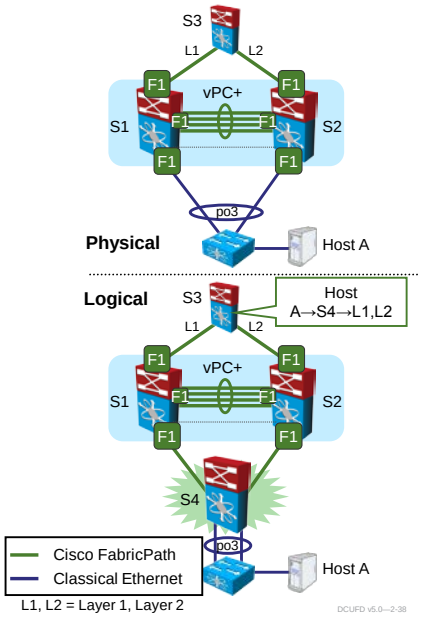
https://t.me/learningnets

certcollecion.net

Virtual PortChannel Operation—vPC+

Interaction between Cisco FabricPath and vPC:

- vPC+ allows dual-homed connections from edge ports into a Cisco FabricPath domain with active/active forwarding.
  - CE switches, Layer 3 routers, dual-homed servers, and so on
- vPC+ requires F1 or F2 modules with Cisco FabricPath enabled in the VDC.
  - Peer link and all vPC+ connections must be to F1 or F2 ports.
- vPC+ creates a "virtual" Cisco FabricPath switch S4 for each vPC+ attached device to allow load balancing within a Cisco FabricPath domain.
- The subswitch ID in the Cisco FabricPath header is used to identify the vPC toward the CE device.



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

Cisco FabricPath and vPCs

vPC support was added to Cisco FabricPath networks in order to support switches or hosts that dual-attach through Classical Ethernet.

A vPC+ domain allows a CE vPC domain and a Cisco FabricPath cloud to interoperate. A vPC+ domain also provides a First Hop Routing Protocol (FHRP) active/active capability at the Cisco FabricPath to Layer 3 boundary.

Note	vPC+ is an extension to vPCs that run CE only. You cannot configure a vPC+ domain and a vPC domain on the same Cisco Nexus 7000 Series device.

A vPC+ domain enables Cisco Nexus 7000 Series devices that are enabled with Cisco FabricPath devices to form a single vPC+, which is a unique virtual switch to the rest of the Cisco FabricPath network. You configure the same domain on each device to enable the peers to identify each other and to form the vPC+. Each vPC+ has its own virtual switch ID.

vPC+ must still provide active/active Layer 2 paths for dual-homed CE devices or clouds, even though the Cisco FabricPath network allows only one-to-one mapping between the MAC address and the switch ID. vPC+ provides the solution by creating a unique virtual switch to the Cisco FabricPath network.

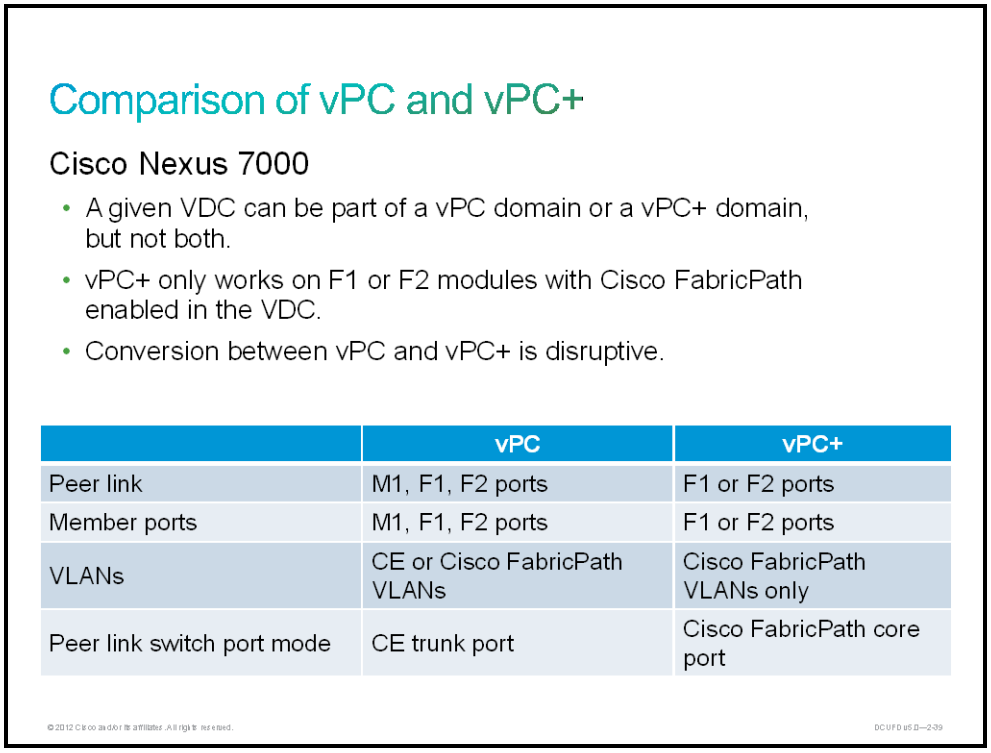
The F1 Series modules have only Layer 2 interfaces. To use routing with F1 Series modules and vPC+, you must have an M Series module inserted into the same Cisco Nexus 7000 Series chassis. The system then performs proxy routing using both the F1 Series and the M1 Series modules in the chassis.

The F2 Series modules cannot exist in the same VDC with F1, M1, or M1XL Series modules. Therefore, you cannot mix F1 and F2 interfaces in vPC+.

The subswitch ID field in the Cisco FabricPath header identifies the vPC port channel of the virtual switch where the frame should be forwarded.

https://t.me/learningnets

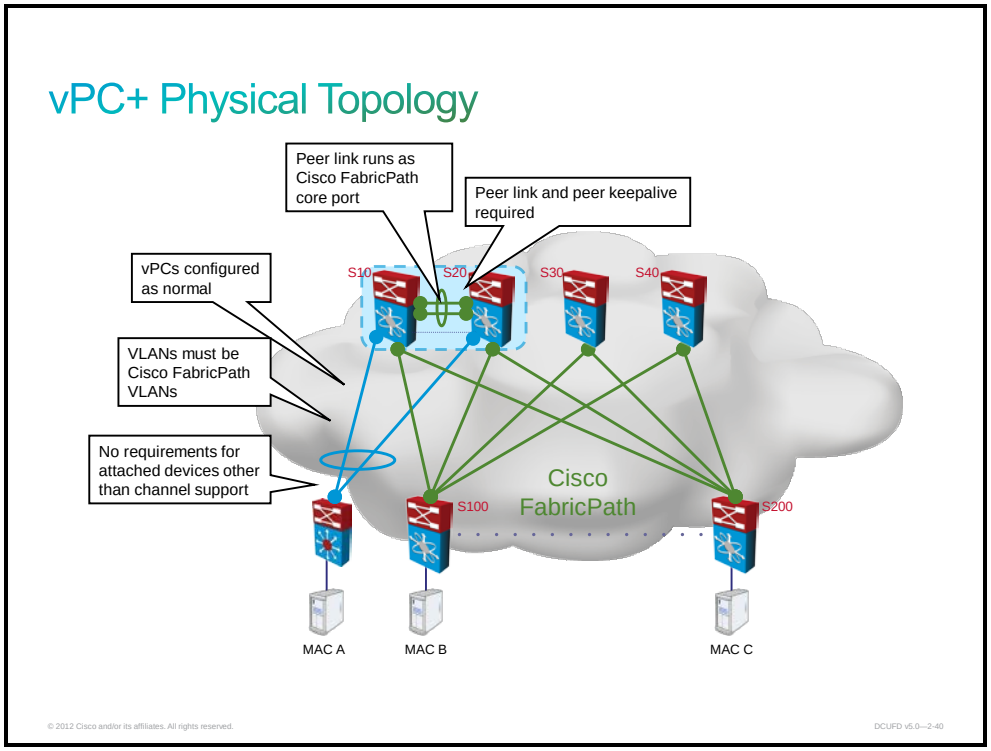
certcollecion.net



This figure shows the differences between vPC and vPC+. You must have all interfaces in the vPC+ peer link as well as all the downstream vPC+ links on an F-Series module with Cisco FabricPath enabled. The vPC+ downstream links are Cisco FabricPath edge interfaces, which connect to the CE hosts.

https://t.me/learningnets

certcollecion.net



This figure explains vPC+ physical topology:

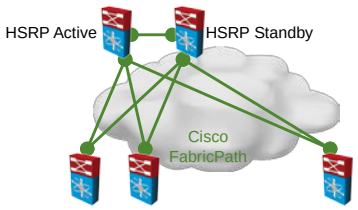
- Cisco FabricPath devices can be dual- or multi-attached to the core switches arbitrarily and form the Cisco FabricPath domain. Classical Ethernet switches must be dual-attached to one pair of Cisco FabricPath switches.
- On the Cisco FabricPath switches, a virtual switch is created in the topology to accommodate the vPC+. Through the vPC+, the CE switch can take full advantage of multipathing.

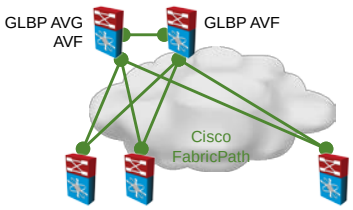
https://t.me/learningnets

certcollecion.net

Cisco FabricPath and Default Gateway

- HSRP
  - Active/standby deployment
  - HSRP on spine switches
  - Hosts learn one default gateway MAC address through conversational learning.
  - HSRP MAC is active only on one gateway.
- GLBP
  - Active/active deployment
  - GLBP on spine switches
  - Hosts learn multiple default gateway MAC addresses through conversational learning.
  - GLBP MAC addresses are active on multiple AVFs.





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

AVF = actual virtual forwarder

Default Gateway Routing and Cisco FabricPath

Cisco FabricPath is a Layer 2 technology that provides a fabric that is capable of efficient forwarding of traffic and load balancing, which is especially suitable for “East-West” traffic. Traffic that must be routed to other networks (“Northbound”) needs to be forwarded to the default gateway.

HSRP Active/Standby

The gateway protocols work in Cisco FabricPath the same way that they work in typical Ethernet networks. The default active/standby behavior of HSRP does not negate the value of Cisco FabricPath.

Even without any further configuration than the standard configuration of HSRP, the Layer 2 traffic (east to west) would benefit from multipath forwarding.

Routed traffic (south to north) would, instead, be forwarded only to the active HSRP device.

GLBP Active/Active

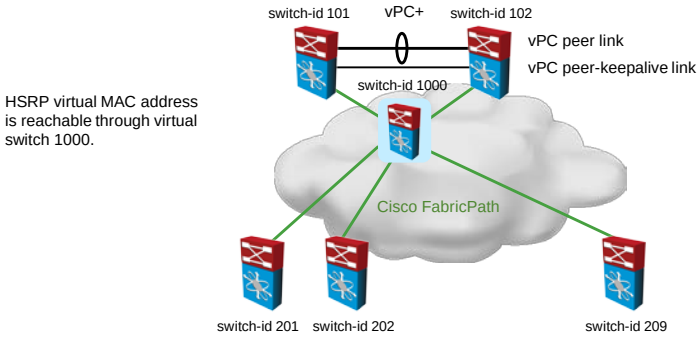
To perform traffic forwarding from multiple spine switches, you can use the Gateway Load Balancing Protocol (GLBP), which hands out a different gateway MAC (up to four different MACs) in a round-robin fashion.

https://t.me/learningnets

certcollecion.net

Cisco FabricPath and Default Gateway (Cont.)

- HSRP active/active is also possible by adding a vPC+.
  - HSRP virtual MAC address is bound to the virtual switch ID of vPC+.
  - Switches in the fabric perform ECMP to the virtual switch across both spine switches.



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

HSRP Active/Active

By taking full advantage of the concept of vPC+ at the spine, you can achieve HSRP forwarding in active/active fashion.

The spine devices are connected by an additional Cisco FabricPath link (which would be recommended in any case to optimize multidestination tree forwarding) and by defining it as a vPC+ peer link. The use of vPC+ at the spine is strongly recommended. vPC+ gives you the ability to forward routed traffic to multiple routing engines as well as to optimize failover times.

Note	The vPC+ peer link must be built using F1 ports and not M1 ports, because it must be configured as a Cisco FabricPath link.

As a result of configuring HSRP and vPC+, the edge switches learn the association of HSRP virtual MAC addresses with the emulated switch ID instead of the individual spine switch ID.

The use of vPC+ allows the routed traffic between any two given endpoints to benefit from both Layer 2 equal-cost multipathing and from the aggregated routing capacity of the spine switches.

Peer Link Failure

As a result of declaring the link that connects the spines as a vPC peer link, the default behavior of the vPC applies, whereby if the peer link goes down, the SVIs on the vPC secondary device are shut down.

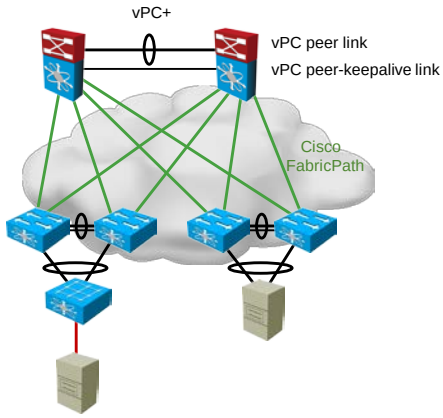
In the context of Cisco FabricPath designs, this behavior is not beneficial, because the Cisco FabricPath links are still available, and there is no good reason to shut down the SVIs on the secondary device. To continue forwarding over the Cisco FabricPath fabric to the HSRP default gateway, exclude the SVIs to be shut down by properly configuring the vPC domain.

https://t.me/learningnets

certcollecion.net

vPC+ at Cisco FabricPath Edge

- The link between edge switches allows for direct server-to-server traffic.
- vPC+ at the edge allows multihoming for a FEX or for a vPC to a host.



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

vPC+ at the Edge

Additional links at the Cisco FabricPath edge are beneficial because they allow direct server-to-server traffic between edge switches and do not need to traverse one of the spine switches. The additional path is considered by the Cisco FabricPath IS-IS control plane.

Another example is using the link as the vPC peer link to either multihome a FEX or to create a port channel to a host that intensively uses the network, such as a virtual machine host. Traffic from the network to the server (“Southbound”) is sent to the emulated vPC+ virtual switch, which then load-balances the traffic across both edge switches, forwarding it to the server.

In all cases, the link between edge switches must be configured as a Cisco FabricPath link.

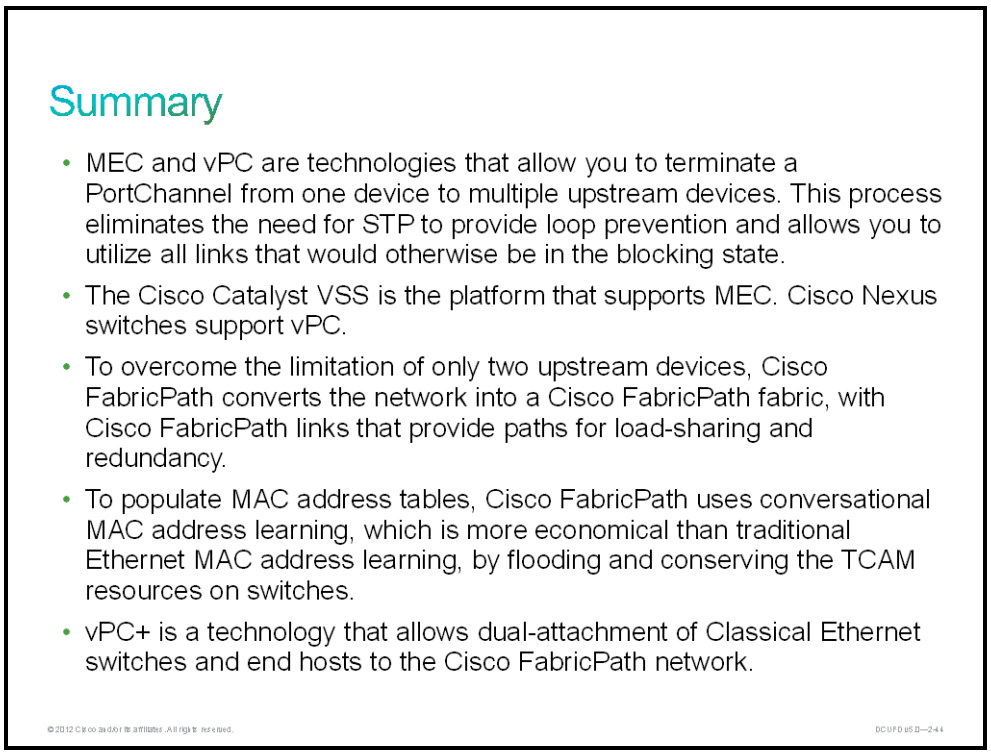
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

Summary

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



References

For additional reference, please refer to the following material:

- Cisco FabricPath Design Guide: Using FabricPath with an Aggregation and Access Topology at http://www.cisco.com/en/US/prod/collateral/switches/ps9441/ps9670/guide_c07-690079.html#wp9000285

https://t.me/learningnets

certcollecion.net

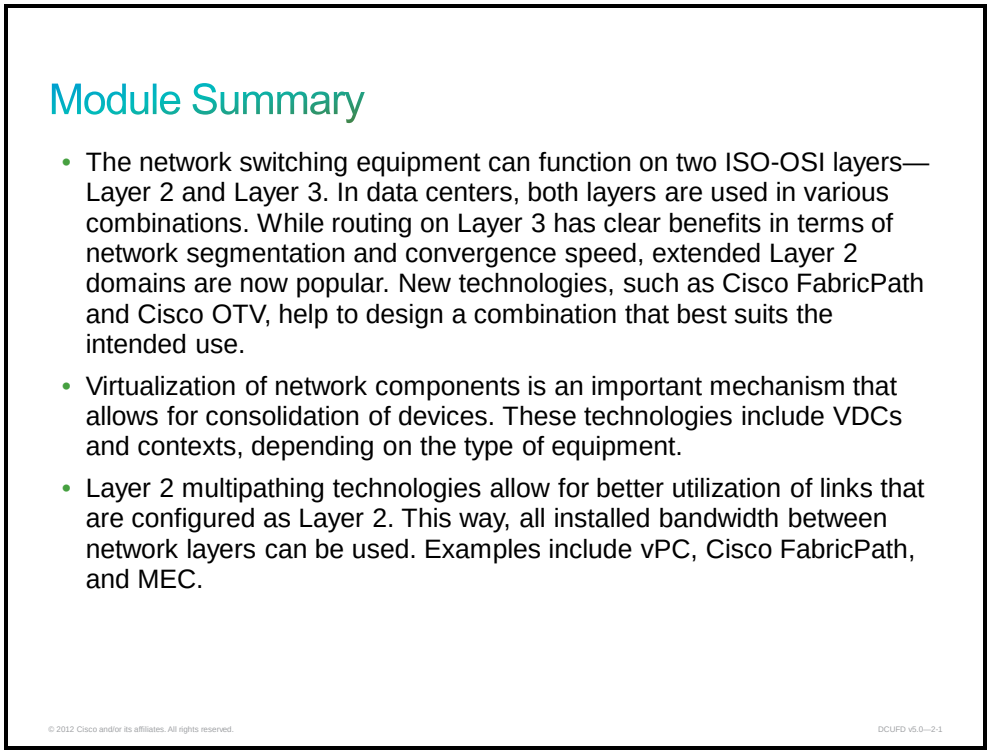
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

Module Summary

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



In this module, you learned about Cisco technologies that are used in data centers.

The first lesson covered traditional routing and switching technologies that are widely supported.

The second lesson covered various device virtualization technologies that are used to virtualize physical equipment into multiple logical devices. These device virtualization technologies allow for equipment consolidation in data centers, increasing the operational efficiency of devices.

The last lesson described Layer 2 multipathing technologies, allowing for designs that utilize all links between the switches, including those that would otherwise be blocked by Spanning Tree Protocol (STP).

https://t.me/learningnets

certcollecion.net

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

Module Self-Check

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

Q1) Which forwarding database is used on the I/O modules to decide where to forward the packet? (Source: Designing Layer 2 and Layer 3 Switching)

A) OSPF topology table
B) dFIB
C) RIB
D) MAC address table

Q2) Which type of forwarding is being used when a switch is forwarding traffic using its supervisor engine? (Source: Designing Layer 2 and Layer 3 Switching)

A) MAC address-based forwarding
B) distributed forwarding with central forwarding engine
C) centralized forwarding
D) routed transport

Q3) What must be considered when designing IP addressing for the data center management network? (Source: Designing Layer 2 and Layer 3 Switching)

A) number of managed devices
B) firewall inspection capabilities
C) use of contiguous subnets that can be easily summarized
D) use of IPv6 exclusively to prevent attacks to the management network

Q4) How is the Cisco Nexus 7000 Series Switch virtualized using VDCs? (Source: Virtualizing Data Center Components)

A) Multiple switches join the same VDC and establish vPCs.
B) A switch is divided into multiple VDCs.
C) The switch is divided using multiple VLANs and VRFs.
D) VDCs virtualize the switch on the hypervisor layer.

Q5) What are two considerations to make when you are assigning interfaces to VDCs? (Choose two.) (Source: Virtualizing Data Center Components)

A) The sum of the required resources for every VDC must not exceed the available resources on the I/O modules.
B) You can mix all I/O modules in the VDCs.
C) Power requirements increase when you enable VDCs.
D) The ports must be allocated considering the distribution of ports in port groups on the I/O module.

Q6) What is the purpose of a fabric extender? (Source: Designing Layer 2 Multipathing Technologies)

A) It provides cost-effective connectivity to gigabit-only endpoints.
B) It provides switching functions on the extended fabric.
C) It shortens cable lengths.
D) It provides network self-management capabilities.

https://t.me/learningnets

certcollecion.net

Q7) What are the three most commonly used Layer 2 multipathing technologies? (Choose three.) (Source: Designing Layer 2 Multipathing Technologies)

A) vPC
B) MEC
C) STP
D) HSRP
E) Cisco FabricPath

Q8) Which address learning mechanism does Cisco FabricPath use? (Source: Designing Layer 2 Multipathing Technologies)

A) MAC address learning by flooding
B) IS-IS-based MAC address learning
C) conversational MAC address learning
D) prolonged MAC address retention

Q9) What is the primary benefit of a Cisco FabricPath fabric against vPC? (Source: Designing Layer 2 Multipathing Technologies)

A) incorporated routing for HSRP
B) the ability to use more than two upstream switches for load balancing to other parts of the network
C) vPC support for Cisco Catalyst switches
D) faster data plane operation

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

Module Self-Check Answer Key

Q1) B

Q2) C

Q3) A

Q4) B

Q5) A, D

Q6) A

Q7) A, B, E

Q8) C

Q9) B

https://t.me/learningnets

certcollecion.net

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

Module 3

Data Center Topologies

Overview

In this module, you will learn about designing data centers from the topology point of view. Several devices need to be interconnected, and depending on the size of the data center, the layered approach is used. The access layer connects the physical devices, the aggregation layer aggregates the connections, and the core layer interconnects multiple data center aggregation blocks and the rest of the network. In addition to this classic example, there are other examples as well, such as the collapsed core layer, the virtual access layer, and so on.

In this module, you will also study data center topologies that will enable you to determine which technologies, such as virtual port channel (vPC), Cisco FabricPath, and so on, are the best fit for any particular design requirements.

Module Objectives

Upon completing this module, you will be able to design data center connections and topologies in the core, aggregation, and access layers. This ability includes being able to meet these objectives:

- Design data center connections and topologies in the core layer
- Design data center connections, topologies, and services in the aggregation layer
- Design the data center physical access layer
- Design the data center virtual access layer and related physical connectivity, and describe scalability limitations and application impact
- Design for data center high availability with various technologies, including IP routing, clusters, next-hop redundancy protocols, and LISP
- Design data center interconnects for both data traffic and storage traffic, over various underlying technologies

https://t.me/learningnets

certcollecion.net

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

Lesson 1

Designing the Data Center Core Layer Network

Overview

In this lesson, you will learn about the role of the data center core layer, and design considerations for it. The role of the data center core is to provide an interconnection between data center aggregation blocks and the campus core network. High-speed switching and high-bandwidth links are provisioned in the core network.

Objectives

Upon completing this lesson, you will be able to design data center connections and topologies in the core layer. This ability includes being able to meet these objectives:

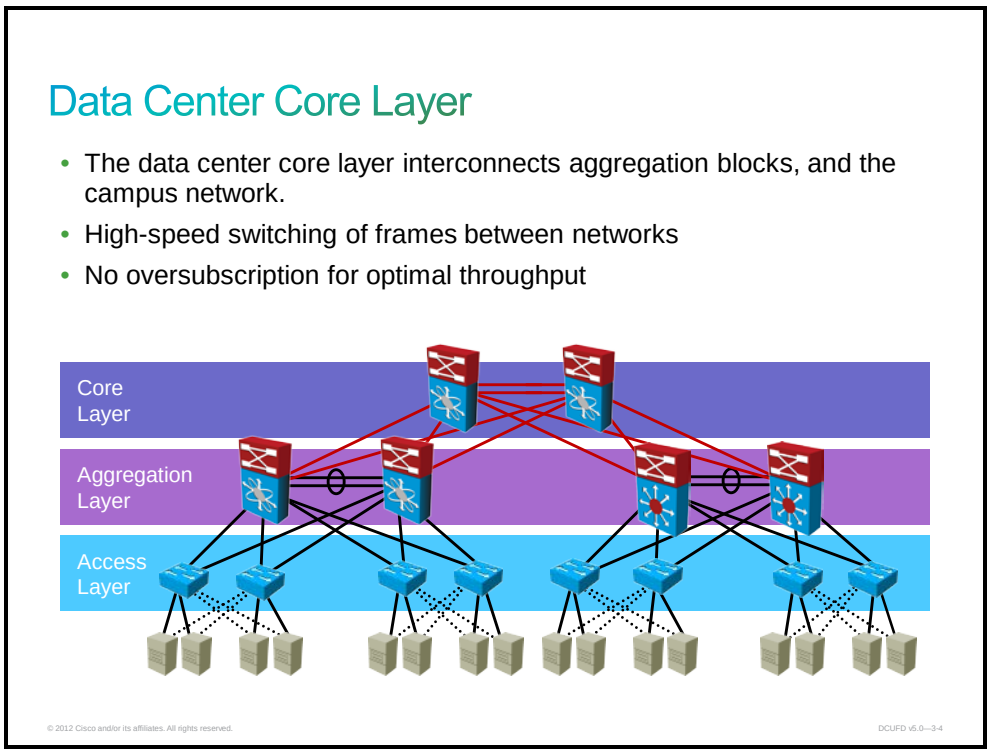
- Identify the need for the data center core layer
- Design a Layer 3 data center core layer
- Design a Layer 2 data center core layer
- Evaluate designs using data center collapsed core

https://t.me/learningnets

certcollecion.net

Data Center Core Layer

This topic describes how to identify the need for the data center core layer.



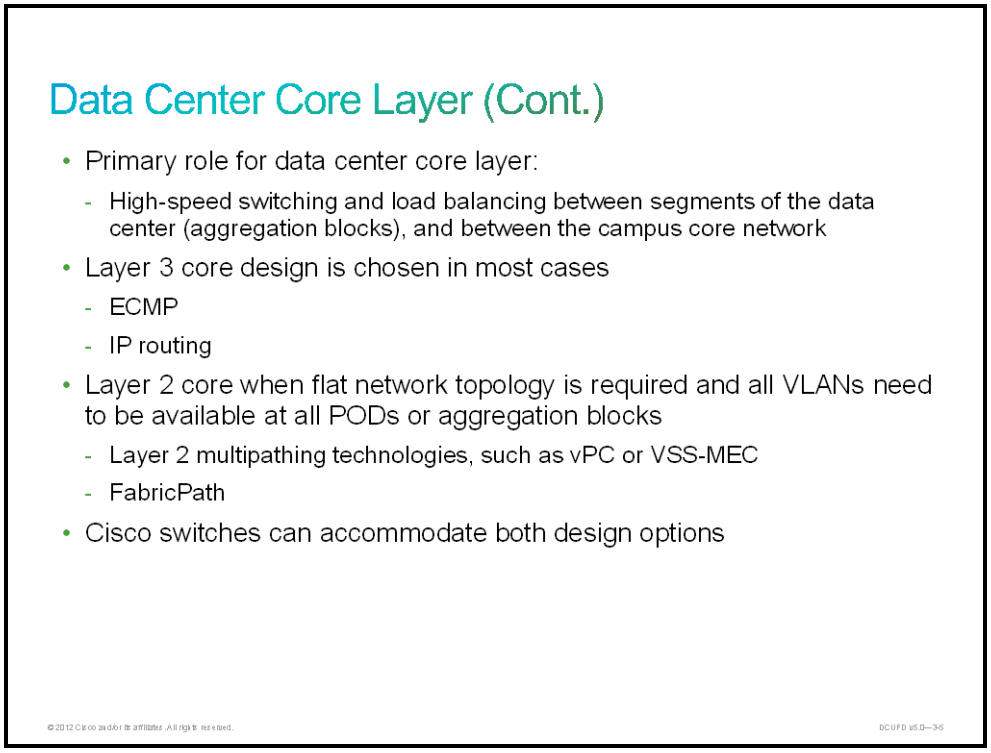
The primary function of the data center core layer is to provide a high-speed interconnection between different parts of the data center, which are grouped in several aggregation blocks.

When provisioning links and equipment for the core, you should allow as little oversubscription as possible. The data center core should not drop frames because of congestion. An efficient core improves the performance of applications.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net



Most data center core deployments are based on Layer 3 forwarding technologies. These include IP routing protocols and forwarding technologies, such as Equal-Cost Multipath (ECMP).

ECMP supports up to 16 paths with Cisco Nexus 7000 and Catalyst 6500 Series equipment.

Using IP routing implies that you segment the data center into multiple networks, and route traffic between them. When network segmentation is not desired, traffic can be forwarded on a Layer 2 basis. Using traditional Layer 2 technologies introduces severe drawbacks in link scalability (that is, the Spanning Tree Protocol (STP), broadcast flooding, and so on), so such deployments had limited scalability.

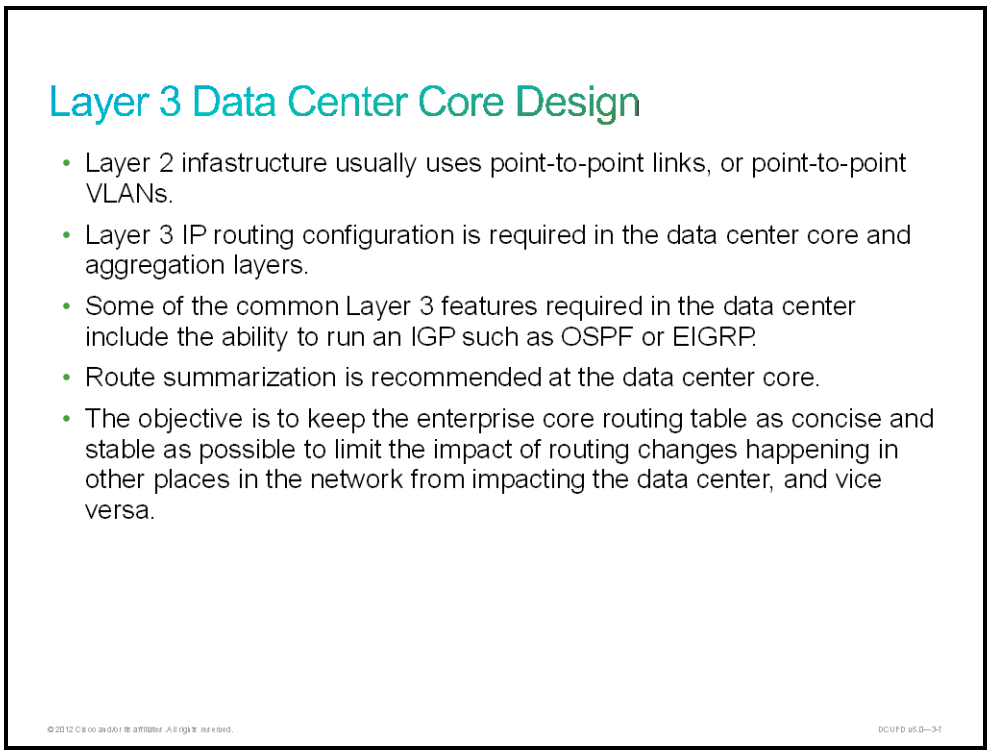
To overcome limitations of traditional single-domain Layer 2 deployments, you can use Cisco FabricPath as the core forwarding technology, which can control the whole Layer 2 domain and select optimal paths between devices. Alternatively, you can achieve good scalability and throughput by using multilink aggregation technologies, such as the virtual port channel (vPC), or the Multichassis EtherChannel (MEC) (when using Cisco Catalyst Virtual Switching System [VSS]).

https://t.me/learningnets

certcollecion.net

Layer 3 Data Center Core Design

This topic describes how to design a Layer 3 data center core layer.



The Layer 2 infrastructure usually uses point-to-point links, or point-to-point VLANs between switches. Over those links, you can establish a routing protocol adjacency and forward traffic between devices.

Layer 3 IP routing configuration is required in the data center core and aggregation layers. This includes routing adjacency configuration, possible route filtering, summarization, and default route origination.

Some of the common Layer 3 features required in the data core include the ability to run an interior gateway protocol (IGP) such as Open Shortest Path First (OSPF) or Enhanced Interior Gateway Routing Protocol (EIGRP). OSPF in particular allows route summarization on network area borders.

Route summarization is recommended at the data center core, so that only summarized routes are advertised out of the data center, and only the default route is advertised into the data center. The objective is to keep the enterprise core routing table as concise and stable as possible to limit the impact of routing changes happening in other places in the network from impacting the data center, and vice versa.

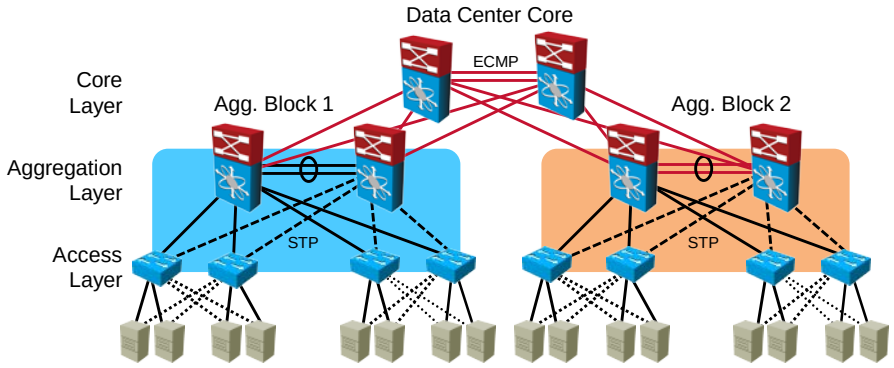
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

Layer 3 Data Center Core Topology

- Design using segmented networks provides good failure isolation
- Can implement centralized control over traffic, since all traffic between aggregation blocks passes through the core
- Drawback: network segmentation does not allow extension of the same VLAN between two aggregation blocks



A design that uses segmented networks provides good failure isolation. A failure in the STP or in a VLAN (Layer 2 broadcast domain) in one aggregation block does not reach the core switch. This provides good failure isolation and stability for the rest of the data center network.

When using a Layer 3 core layer, you have control over traffic because all traffic between aggregation blocks needs to be forwarded through the core switches.

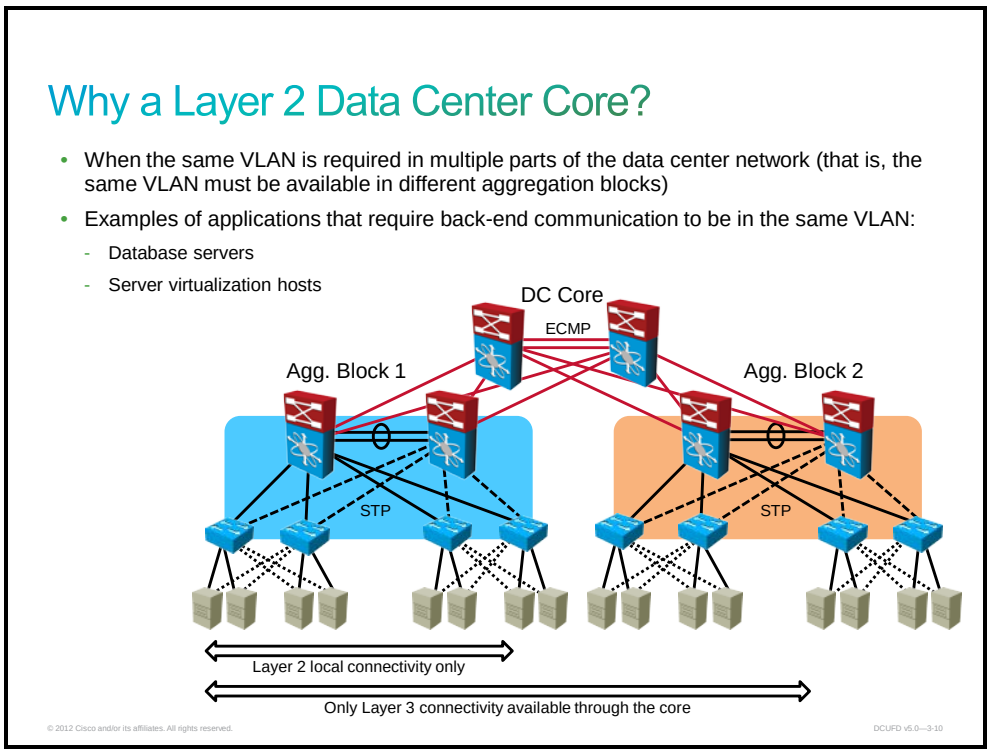
On the other hand, the Layer 3 segmented network has some drawbacks: network segmentation does not allow you to extend the same VLAN between two aggregation blocks. This has become the requirement for enabling workload mobility (live virtual machine migration), and application clusters, such as database servers.

https://t.me/learningnets

certcollecion.net

Layer 2 Data Center Core Design

This topic describes how to design a Layer 2 data center core layer.



The requirement for data center-wide VLANs has become apparent with the arrival of technologies that enable workload mobility (live virtual machine migration), and application clusters, such as database servers. Database server clusters use Layer 2 connectivity for operation, cluster load balancing, and back-end database synchronization. Clusters of server virtualization hosts may also require the same VLAN to be extended across various parts of the data center. Such VLANs are used for control traffic, or virtual machine mobility solutions.

Note	Some databases may be able to synchronize over IP-based networks. In this case, you do not need to have all servers that are part of clusters be in the same subnet.

Because there is a Layer 3 core between the aggregation blocks of the data center, the Layer 2 connectivity cannot be contiguous.

To enable host-to-host reachability through the data center core, the hosts need to be in different subnets and in different Layer 2 domains. Under these conditions, direct Layer 2 connectivity is not possible.

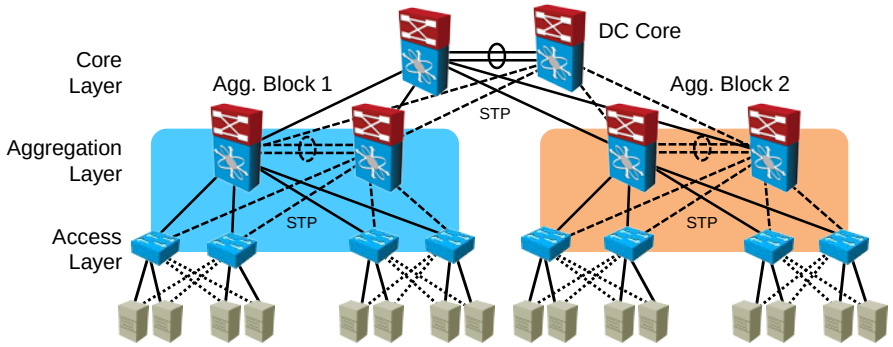
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

Layer 2 Data Center Core Challenges

- Very large spanning tree domain
- Only 50 percent of installed bandwidth available throughout the network
- Large MAC tables on all switches
- High amount of traffic flooding
- One malfunctioning device is a threat to the whole network
- Deployment of IP Services at the core: FHRP, IP routing, firewalling, and so on.



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

The Layer 2 data center core provides a foundation for data center-wide VLANs. However, there are a couple of issues that introduce potential risks, or are challenging to implement:

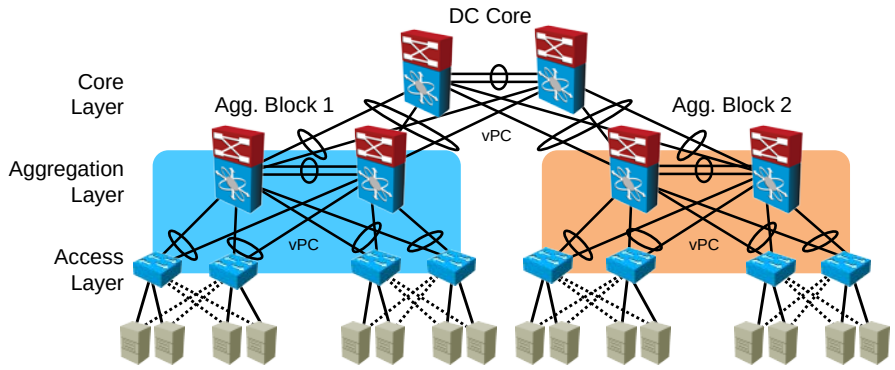
- **Very large spanning tree domains:** Very large spanning tree domains greatly affect the convergence time of STP when changes occur in the network. The more network devices you have in such a network, the greater is the chance that some devices change state from online to offline, and generate a topology change. If a crucial switch goes offline (that is, any switch that has STP root ports), the topology will need to reconverge and trigger an event.
- **50 percent of bandwidth available:** The STP will block half of the links to prevent Layer 2 loops.
- **Large MAC tables on all switches:** There are many hosts in such networks, and switches need to constantly age-out and relearn MAC table entries from traffic flows.
- **High amount of traffic flooding:** Broadcast frames for protocols such as Address Resolution Protocol (ARP), various discovery mechanisms, application broadcast traffic, and unknown unicast traffic generates much flooding traffic that burdens the entire network and all attached network devices.
- **Malfunctioning devices:** One malfunctioning device is a threat to the whole network, because it can generate broadcast storms. Broadcast storm suppression can be used on the links, but it also stops legitimate broadcast traffic.
- **Deployment of IP Services at the core:** While not a real challenge, services such as First Hop Redundancy Protocol (FHRP), IP routing, firewalling, and so on, need to be implemented at the core and require devices that manage more bandwidth.

https://t.me/learningnets

certcollecion.net

Layer 2 Data Center Core Challenges (Cont.)

- Layer 2 multipathing using vPC solves only the available bandwidth issue.
- Data center-wide, end-to-end VLANs are available, but with all of the potential issues of flat Layer 2 networks/



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

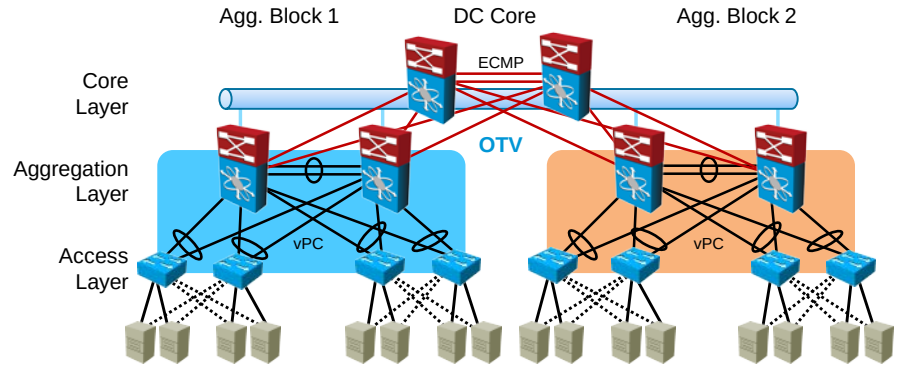
Layer 2 multipathing mechanisms will solve the issue of only 50 percent of bandwidth being available because they manage multiple uplinks, so that they can be used at the same time. This increases available bandwidth and reduces convergence time. However, this does not protect you from the downsides of Layer 2 flat networks, such as broadcast storms, large MAC table usage, excessive flooding, and so on.

https://t.me/learningnets

certcollecion.net

Layer 3 Data Center Core with Layer 2 Extension

- "Best of both worlds" design
- Layer 3 core with network segmentation and ECMP
- Extended subnets across aggregation blocks provided by Cisco OTV
- Layer 3 core acts as Cisco OTV IP backbone



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

An elegant solution is to use a common and well-known design for the Layer 3 core, and extend the VLANs that you need at multiple aggregation blocks with a Layer 3 encapsulation technology, such as Cisco Overlay Transport Virtualization (OTV). Cisco OTV is supported on Cisco Nexus 7000 Series Switches.

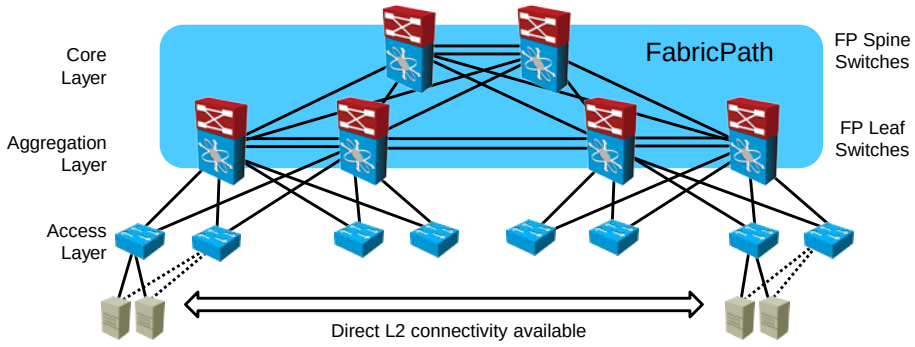
In this case, the Layer 3 core acts as an IP backbone for Cisco OTV, the VLAN extension mechanism. In this case, you can have all VLANs that you require present at all aggregation blocks, and allow for live mobility of virtual machines between the segments of the network.

https://t.me/learningnets

certcollecion.net

Layer 2 Data Center Core Design Option

- Cisco FabricPath can be used to enable data center-wide VLANs.
- FabricPath uses an underlying routed topology for loop prevention.
- Can distribute the load over several links between aggregation blocks
- Implementation of other IP services (firewalling, and so on) in the core



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

Another option to bring the same VLAN to multiple aggregation blocks is to use Cisco FabricPath. FabricPath uses an underlying routed topology and the Time to Live (TTL) field in the frame to prevent Layer 2 loops and broadcast storms. To conserve core switch resources, it uses conversational MAC address learning.

One of the most important benefits of Cisco FabricPath is that it can use several links to carry data for the same VLAN between several aggregation switches, which allows up to 16-way load balancing.

Optimal path selection and load balancing is achieved by Cisco FabricPath. The switches that connect the access switches are called leaf switches, which connect to other leaf switches across spine switches. In the example in the figure, traffic can be load balanced between two leaf switches across both spine switches.

You still need to provide IP Services such as FHRP, IP routing, server load balancing, and connection firewalling at the data center core layer.

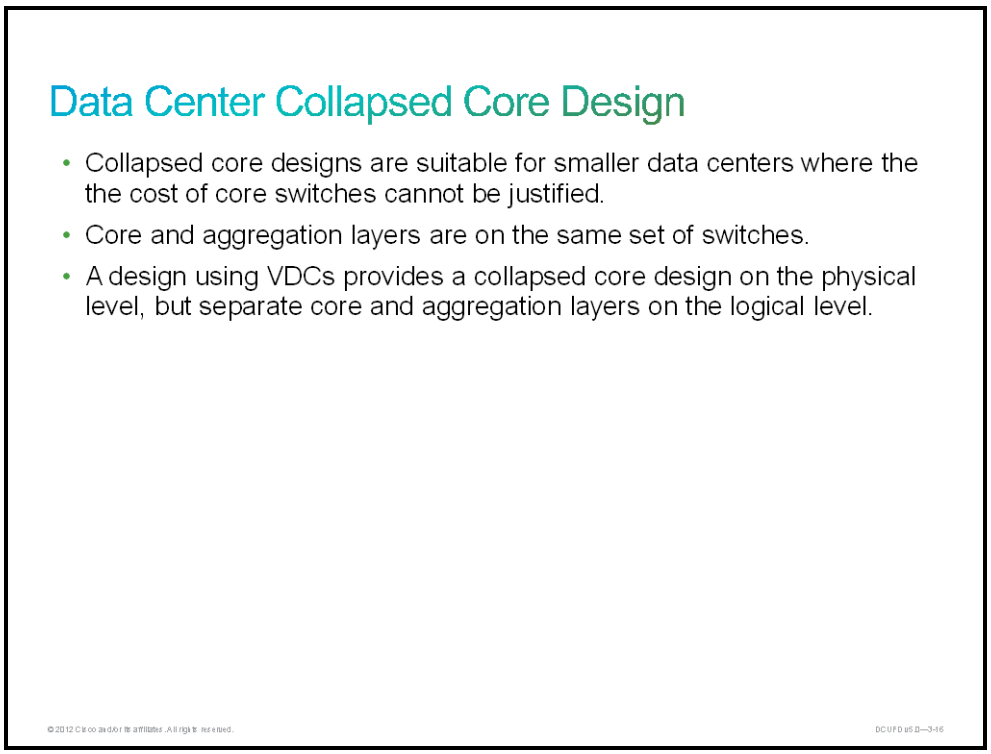
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

Data Center Collapsed Core Design

This topic describes how to evaluate designs using data center collapsed core.



Collapsed core designs are suitable for smaller data centers where the cost of core switches cannot be justified. In this case, core and aggregation layers are on the same set of switches.

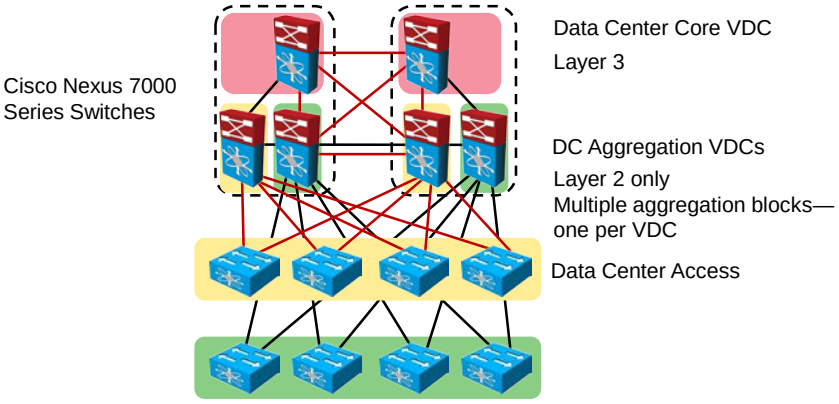
When using the Cisco Nexus 7000 Series Switches, you can use virtual device contexts (VDCs) for device segmentation. VDCs provide a collapsed core design on the physical level, but separate core and aggregation layers on the logical level.

https://t.me/learningnets

certcollecion.net

Data Center Core and Aggregation Consolidation Example

- When using VDCs on Cisco Nexus 7000 Series Switches, there is a possibility to consolidate data center core and aggregation switches:



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

This example describes the collapsed core layer using Cisco Nexus 7000 Series Switches with VDCs configured.

On every Nexus 7000 Series Switch, there is a core VDC configured. This VDC has IP routing configuration and route summarization. The links between the core VDCs, and between the core VDC and the campus core, are all Layer 3 (routed). This VDC forms routing adjacencies with the campus core switches, and with other aggregation VDCs.

The aggregation VDCs are the boundary between switched and routed traffic. They run IP routing with the core, and a FHRP between them.

The access layer is Layer 2 only. Each aggregation VDC has connections to its set of access switches or fabric extenders (FEXs).

The example shows the left and right aggregation VDCs on each switch. Both aggregation VDCs have dedicated links to the access layer. The topmost VDCs form the data center core.

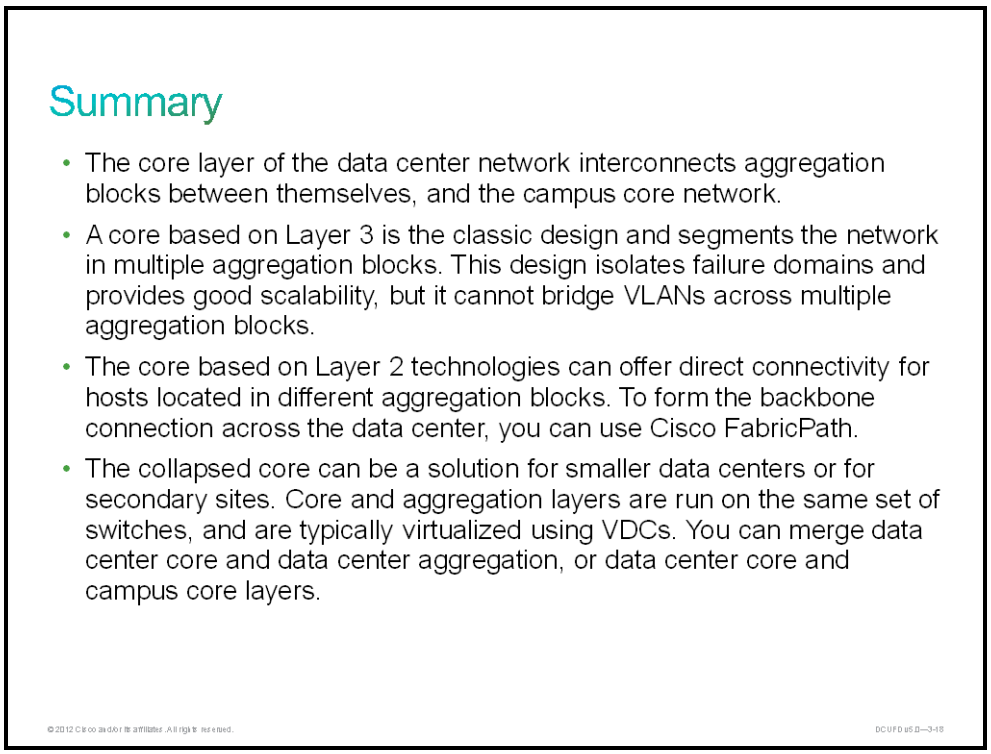
Note There is a drawback of this design. To interconnect the VDCs, you must use physical connections (cables). For 10 Gigabit Ethernet connections, this approach is expensive, and consumes 10 Gigabit Ethernet ports that would otherwise be used for interconnecting devices.
In the example, you need 10 links between the pair of switches to accommodate all needed connections. These consume twenty 10 Gigabit Ethernet ports.

https://t.me/learningnets

certcollecion.net

Summary

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



https://t.me/learningnets

certcollecion.net

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

Lesson 2

Designing the Data Center Aggregation Layer

Overview

This lesson describes the approaches and possibilities for designing the data center aggregation layer. Based on the equipment available, there are many technology options, from existing, Spanning Tree Protocol (STP)-based deployments, to modern, Cisco FabricPath enabled, or Cisco Unified Fabric-enabled aggregation layer designs.

Objectives

Upon completing this lesson, you will be able to design data center connections, topologies, and services in the aggregation layer. This ability includes being able to meet these objectives:

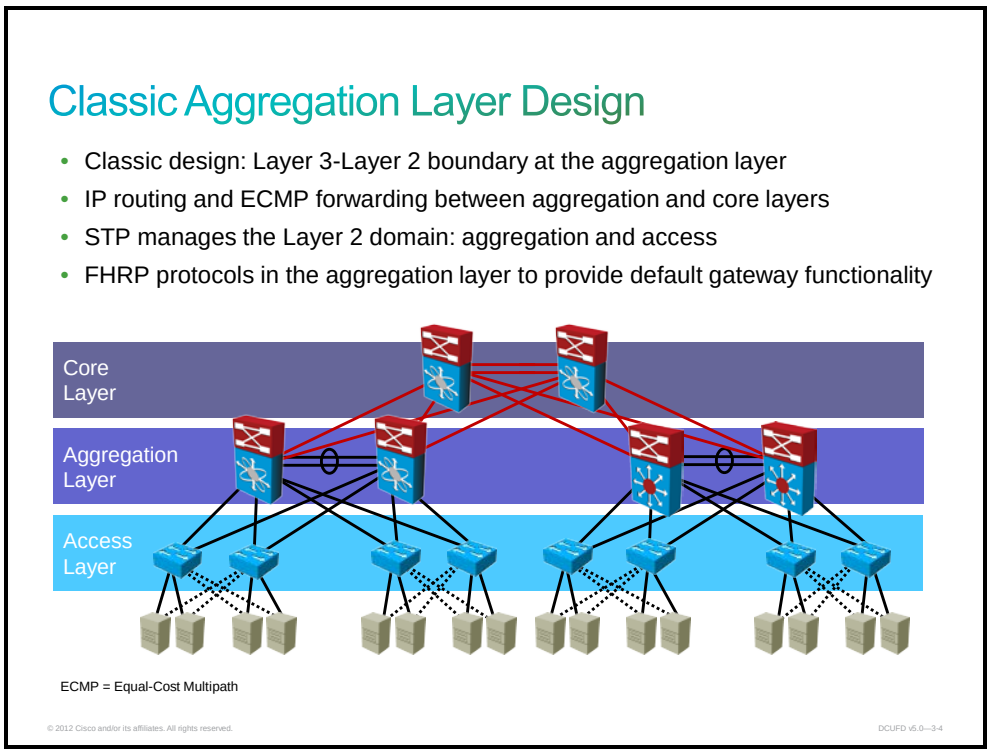
- Describe classic aggregation layer designs
- Design the aggregation layer with VDCs
- Design the aggregation layer using Cisco Unified Fabric
- Design the aggregation layer with IP storage-related specifics in mind

https://t.me/learningnets

certcollecion.net

Classic Aggregation Layer Design

This topic describes classic aggregation layer designs.



The classic aggregation layer design using spanning tree is still used in many existing data centers. These designs are based on STP (some of its variants, like per-VLAN Rapid STP [RSTP]) and are generally robust and with relatively fast reconvergence upon topology changes.

However, the biggest flaw of STP-based designs is the poor utilization of links, and with this, a high oversubscription ratio. Several new technologies emerged that reduce oversubscription by enabling all links. Cisco offers the Multichassis EtherChannel (MEC) for Cisco Catalyst devices, and virtual port channel (vPC) for Cisco Nexus devices.

Classic aggregation layer designs also do not provide any lower-level separation of traffic between parts of data centers. Device capabilities typically include VLANs and virtual routing and forwarding (VRF) instances, but these still use the same control and data planes on the switch. A possible solution to this concern is the virtual device contexts (VDCs) that are implemented on the Cisco Nexus 7000 Series Switches.

Another classification of the type of data center aggregation layer is whether the aggregation layer uses Layer 2 only, or introduces the demarcation point between Layer 2 and Layer 3 forwarding. When using the aggregation layer as a Layer 2-Layer 3 boundary, you need to configure switch virtual interfaces (SVIs), routing, high availability, and possibly other IP-based services, such as firewalling.

When designing a classic aggregation layer, follow these recommendations:

- Interswitch link belongs to the Layer 2 domain.
- Match the STP primary root bridge and the Hot Standby Router Protocol HSRP and Virtual Router Redundancy Protocol (VRRP) active router to be on the same device. This action will route traffic on Layer 2 and Layer 3 on the same path.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

First Hop Redundancy Protocols

There are three options that First Hop Redundancy Protocols (FHRP) you can use:

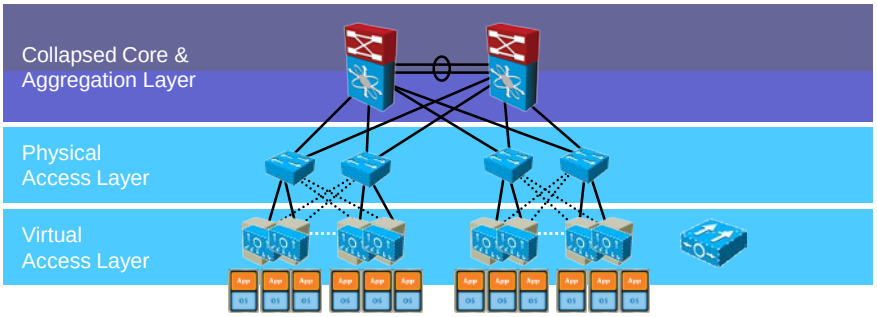
- **HSRP:** The Cisco proprietary default gateway redundancy protocol has the widest selection of features. It provides several options for object tracking and interaction with other router processes, as well as customizations for vPC and virtual port channel plus (vPC+).
- **VRRP:** The standards-based VRRP offers functionality that is similar to HSRP. The primary difference is that the virtual IP address can be same as one of the interface addresses.
- **Gateway Load Balancing Protocol (GLBP):** GLBP is another FHRP that allows you to use several default gateways to forward traffic upstream for server subnets. Returning traffic (downstream) usually enters the network at a single point. Load distribution between default gateways is managed by GLBP and is configurable.

https://t.me/learningnets

certcollecion.net

Collapsed Core Layer Design

- Optimized classic design for smaller networks
- IP routing and ECMP forwarding between collapsed aggregation and core and campus core layers
- STP manages the Layer 2 domain: aggregation and access



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

The collapsed core design is suitable for smaller data center networks, or when you wish to obtain deliberate concentration on core switches. The core and aggregation functions are combined on the same switch, with the same protocols configured on the collapsed switches: STP for the Layer 2 domain, and IP routing protocol for the upstream, Layer 3 domain.

This design still uses the STP protocol as the loop-prevention mechanism for the access network; today, there are better options available, such as vPC and Cisco FabricPath.

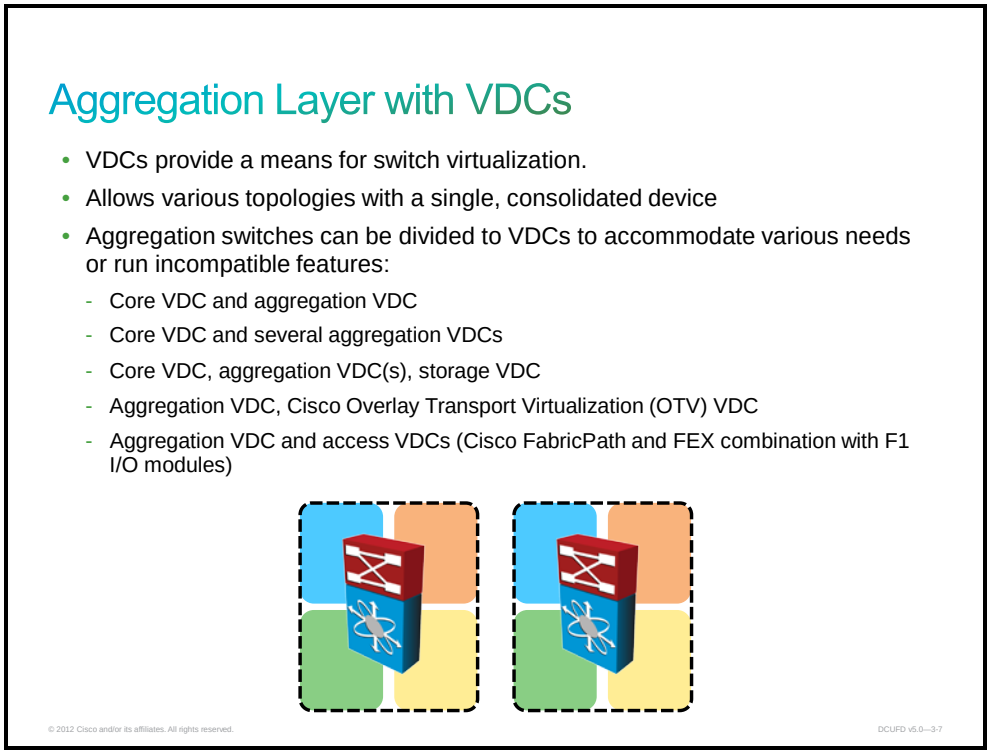
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

Aggregation Layer with VDCs

This topic describes how to design the aggregation layer with VDCs.



VDCs are a Cisco Nexus Operating System (NX-OS) mechanism that allows virtualization of a single physical switch into multiple logical switches. As opposed to virtualization that uses only VLANs, VDCs also virtualize the switch on the Cisco NX-OS software level.

When using VDC separation in the aggregation layer, VDCs can be used to separate various aggregation layers. This allows you to connect various networks and to construct multiple topologies using a single (or a pair) Cisco Nexus 7000 Series Switch.

VDCs provide isolation on the control plane and on data plane. From the control plane perspective, a copy of the entire protocol stack is started for every VDC, with all required higher-level protocols that are required in that VDC: IP, routing protocol, and so on.

On the data plane, traffic is separated using a VDC tag. The VDC tag has local meaning that identifies the VDC to which the packet belongs.

Note You must take care when assigning interfaces to the VDCs and to resource consumption. I/O modules have a fixed amount of ternary content addressable memory (TCAM) available for packet forwarding, access lists, and quality of service (QoS). Having a VDC that utilizes many these resources on a particular I/O module might prevent you from having any other VDC on that I/O module.

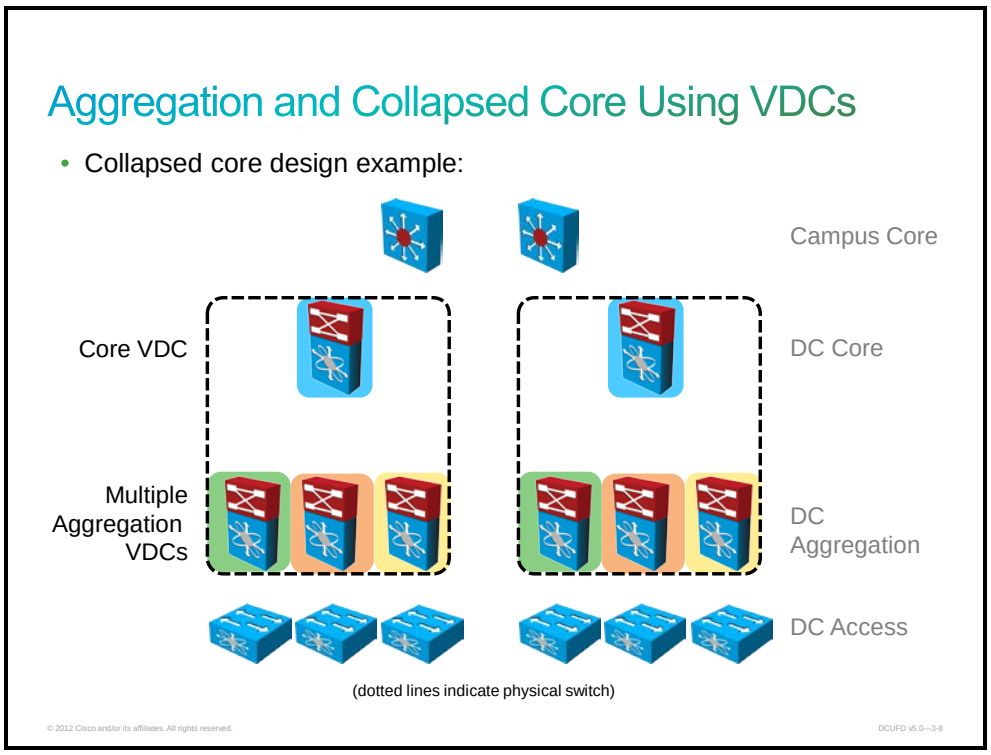
There are several designs that are possible using VDCs, depending on your use case:

- Core VDC and aggregation VDC (collapsed core design with VDCs)
- Core VDC and several aggregation VDCs
- Core VDC, aggregation VDC(s), storage VDC

https://t.me/learningnets

certcollecion.net

- Aggregation VDC and access VDCs (Cisco FabricPath and Cisco Fabric Extender [FEX] combination with F1 I/O modules)
- Aggregation VDC “above” network services and subaggregation VDC “below” network services



In this example, you provide core and aggregation services on the same physical switch.

A similar solution to the aggregation and subaggregation layer VDC “sandwich” is the collapsed core data center topology.

One VDC hosts the core services, while another VDC offers aggregation services. Both VDCs are then consolidated in the same physical switch. In this case, the switch can accommodate multiple aggregation layers for multiple segments, as long as the total VDC count does not exceed the maximum that is supported by the Cisco Nexus 7000 platform.

This example isolates multiple segments and makes all the traffic flow through the core VDC, where traffic can be controlled.

You will need physical cables to interconnect VDCs with each other. You should provision in such a way that you have enough physical ports available, and that they are in different port groups to maximize available bandwidth.

Cisco FabricPath and Cisco FEXs on M1 and F1 Modules

VDCs can be used to bridge the Cisco FabricPath domain to Cisco FEXs. In the case of the Cisco Nexus 7000 F1 modules, you cannot connect Cisco FEXs to these modules directly. Connections need to be physically cabled from F1 I/O modules to M1 I/O modules, where Cisco FEXs are connected.

Note	This limitation does not exist anymore with F2 I/O modules.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

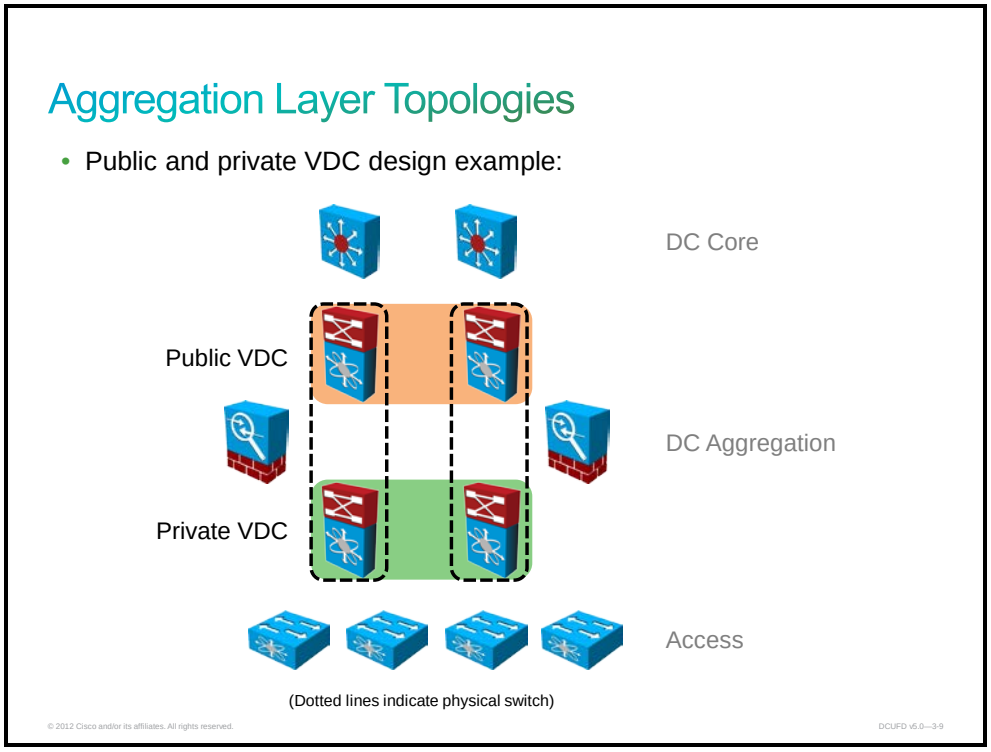
certcollecion.net

Cisco OTV in a Separate VDC

Another example of such usage is Cisco Overlay Transport Virtualization (OTV). In Cisco OTV, there is a limitation that a VLAN cannot be extended over a Cisco OTV overlay to another site if there is an SVI that is configured for that VLAN. Since most VLANs do have SVIs configured, along with FHRP, the recommendation is to create a dedicated VDC for Cisco OTV, and patch that VLAN over a trunk link from the production VDC into the Cisco OTV VDC. In the Cisco OTV VDC, that VLAN is Layer 2 only, without a configured SVI.

VDC Sandwich Configuration

One of the examples is the “VDC sandwich” configuration, where the aggregation layer is divided into public and private VDCs, with security services interlinking these VDCs.



One of the use cases for VDCs is to divide the aggregation layer to public and private parts.

In this case, the VDCs are used for traffic separation when implementing IP Services in the aggregation layer. One VDC is configured as an outside or public VDC, while another VDC can be configured as an internal or private VDC.

After this separation, you can also segment the networks using VLANs and VRFs. This segmentation involves isolating applications unto themselves, maintaining the inside-outside demarcation, and inserting services in between, forming a “sandwiched” network architecture.

Note	Multiple VDCs within the same physical switch need to be linked using a physical cable. Keep in mind the requirements of each I/O module, and how ports can be assigned in multiple VDCs.

https://t.me/learningnets

certcollecion.net

The benefit of this solution is that you consolidate the public and the private part of the network in the same physical switch. To ensure service redundancy, you still need a pair of aggregation switches, and high-availability technologies (vPC, HSRP, and so on).

You can decide which VDC can provide Layer 3 functionality, and which one can provide only Layer 2 functionality, if applicable. Service appliances (firewalls, in this example) may dictate your selection on where would you implement the boundary between Layer 2 and Layer 3 forwarding. They may or may not support dynamic routing, and so on.

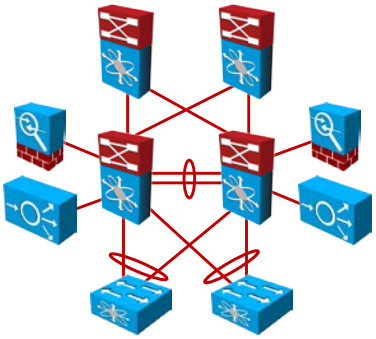
 © 2012 Cisco Systems, Inc.

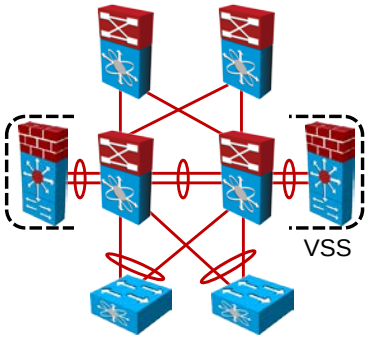
https://t.me/learningnets

certcollecion.net

Aggregation Layer with Network Services

- Standalone services:
  - Implemented with Cisco ASA and Cisco ACE
  - Physical connections between aggregation switches and service appliances
- Services in a service chassis:
  - Implemented with Cisco Catalyst 6500 service chassis with ASA-SM/ FWSM and ACE
  - Physical connections to service chassis, internal connections to service modules





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

There are two ways to deploy network services: as standalone services, using appliances for firewalling and possibly server load balancing, or by using Integrated Services to exploit the capability of the Cisco Catalyst Series switches that have Integrated Services. Both solutions are valid choices and can be used depending on the hardware that you have available.

Performance-wise, there are differences depending on the capacity of each component: the standalone firewall or the load balancer. Integrated Services communicate with the switch through its internal backbone and can connect using higher speed to the aggregation switches. You are limited only by the capacity of the port channel, and not really with the number of available ports on the appliance.

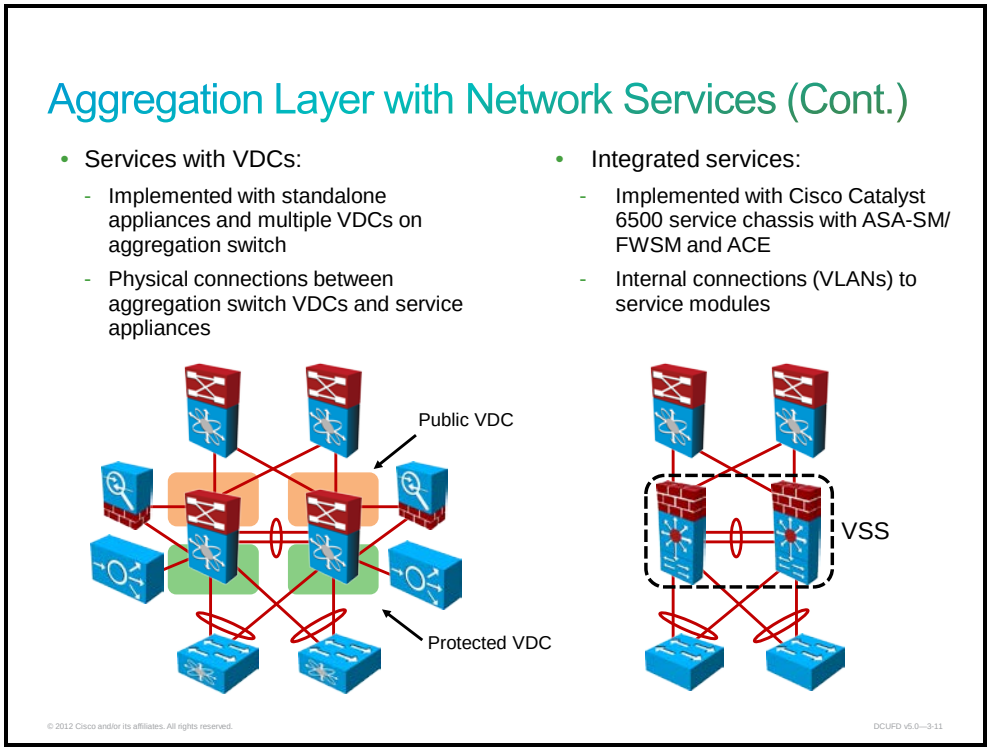
Note Integrated Services can provide another functionality—the route health injection (RHI)—which cannot be done using external appliances. A routing protocol must be running between the service chassis and the aggregation switch.

Standalone services are implemented using Cisco ASA standalone adaptive security appliances, and using the Cisco Application Control Engine (ACE) module to support server load balancing. Other service options are also available, such as data center traffic optimization using Cisco Wide Area Application Services (WAAS).

Integrated Services are limited to the Cisco Catalyst 6500 platform, which has a wide selection of service modules, including the Cisco Adaptive Security Appliance Security Module (ASA-SM) and Cisco Firewall Services Module (FWSM) for firewalling, and the Cisco ACE30 module to provide server load-balancing functionality, at a higher capacity than the standalone appliance. The service chassis can form a Virtual Switching System (VSS) for increased redundancy, shorter convergence time, easier management, and easier configuration of EtherChannels to the pair of Nexus 7000 Series Switches.

https://t.me/learningnets

certcollecion.net



Implementing Services with VDCs

You can deploy network services in the aggregation layer using VDCs as well. On the left side of the figure, there is a combination of standalone services with a VDC sandwich design. The aggregation switches have the public and internal VDCs deployed, and the traffic from the security appliance to the load balancer flows through the default VDC. The default or internal VDC can host all VLANs that are required for redundant operation of the service appliances.

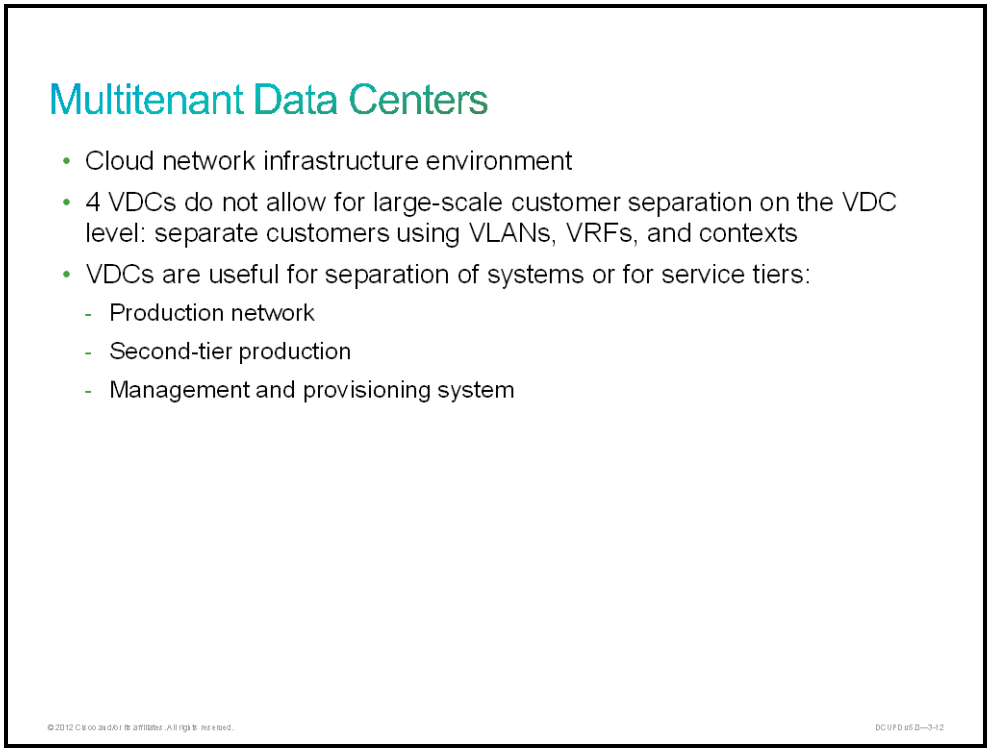
Implementing Fully Integrated Services

Another design is a fully integrated option that takes advantage of the Catalyst 6500 VSS, and Integrated Service modules that plug into the Catalyst 6500. These include the Cisco ASA-SM and the Cisco ACE30 server load-balancing device. The VSS operates as a single switch, while the service modules operate in active-standby operation mode. Active-active load distribution is achieved by using multiple contexts on the service modules.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net



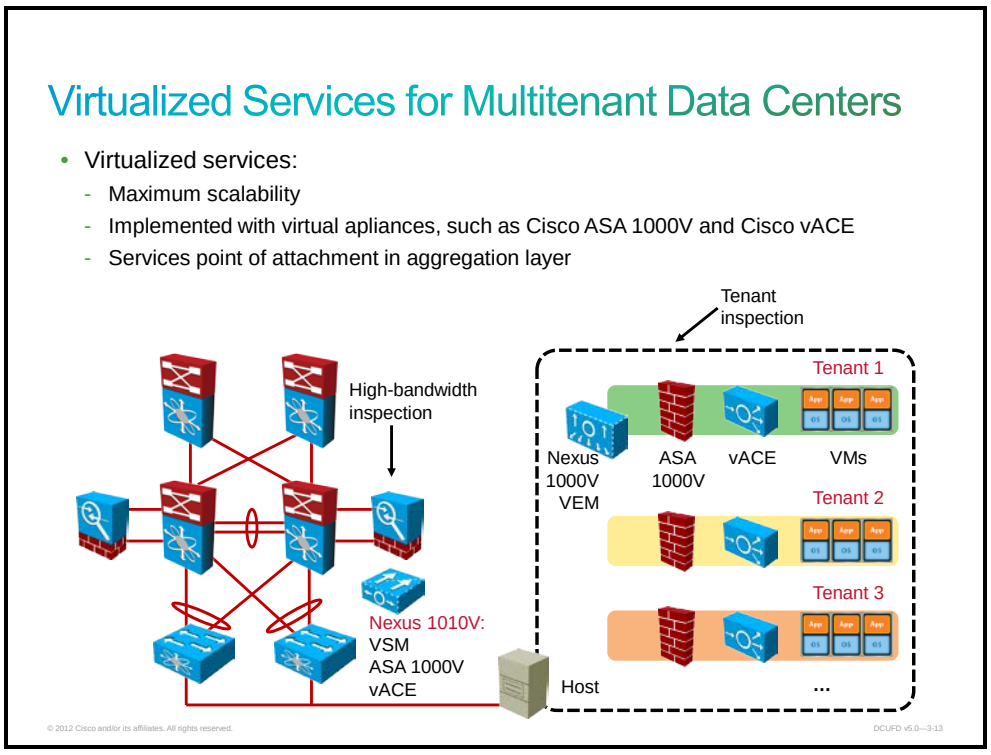
VDCs can be used in multitenant data centers that support various types of cloud solutions. VDCs do provide separation, but do not scale to hundreds of customers, so it makes sense to use VDCs to isolate systems within the data center.

VDCs can be used for production network tiering, when you need to provide different levels of service (or different services) to customers. One VDC can be configured with external services in appliances, where another VDC can provide basic connectivity for “lightweight” users.

It is always a good idea to completely separate the management and provisioning network from the production network. This way, even if you have a breach into the system of one of your customers, the danger cannot spread in the management VDC.

https://t.me/learningnets

certcollecion.net



Implementing Virtualized Services

The aggregation layer hosts various IP Services, most typically using hardware-based, standalone devices or service modules.

With server virtualization taking place and virtual appliances being available, you can combine approaches for an even more robust and flexible deployment.

Physical devices feature high-bandwidth capacity and can be used for raw traffic inspection.

Virtualized devices can be deployed with maximum flexibility, allowing you to customize (or automatically generate) configuration for every tenant, customer, and department for which you are hosting services.

If the tenant has multiple networks or applications, you can also deploy another instance of the virtual services appliance, or another context within an existing instance.

Note	Keep in mind that virtualized services cannot have the same capacity as hardware-based appliances. They are dependent on the host CPU and on the networking hardware.

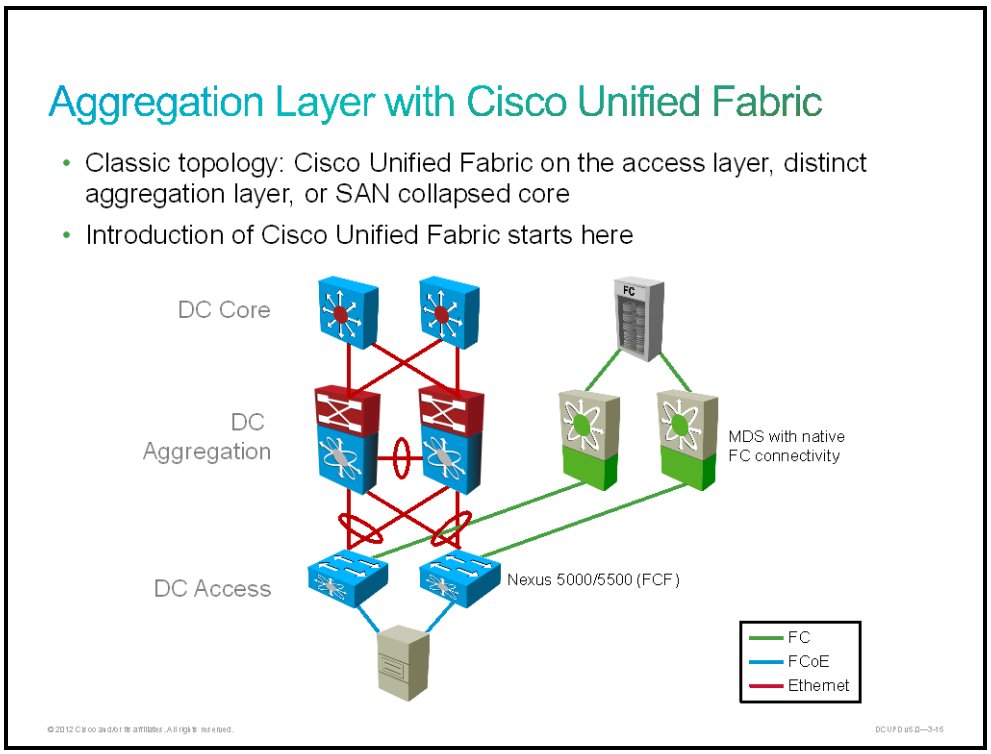
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

Aggregation Layer with Unified Fabric

This topic describes how to design the aggregation layer when using unified fabric.



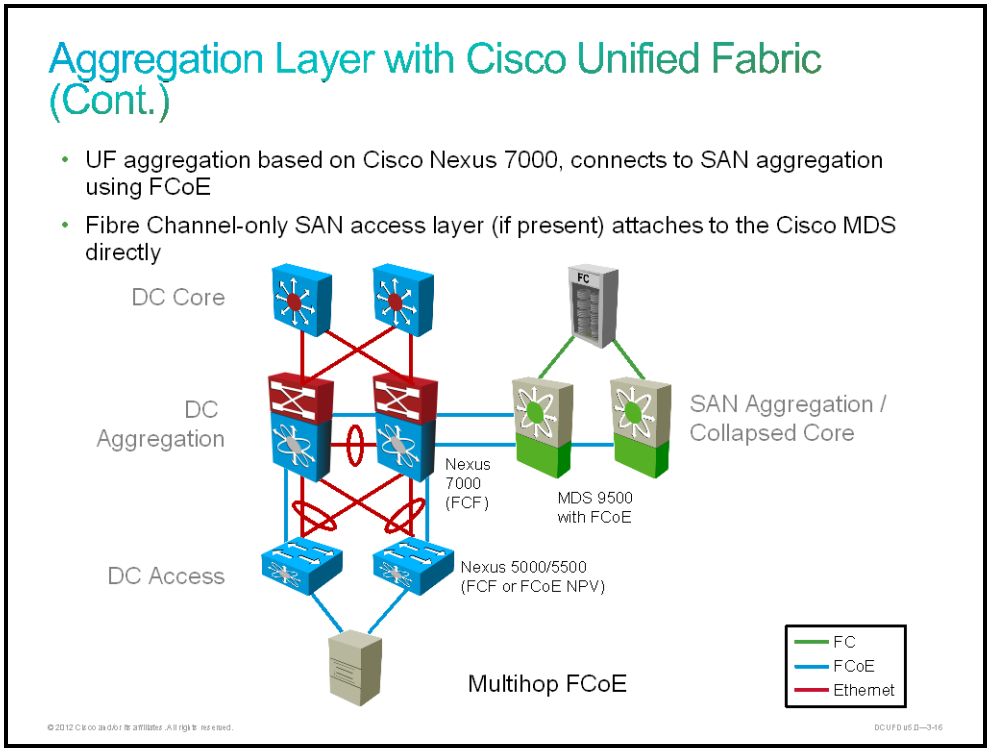
With the development of Cisco Unified Fabric and all supporting technologies and hardware, it makes sense to consolidate access and aggregation layers, mainly to simplify the network and to reduce costs.

Note	Migrations are usually not done before some equipment reaches the end of its investment cycle. Therefore, the equipment must support various ways of interconnecting new equipment and existing systems.

In the example in the figure, Cisco Unified Fabric is used only in the access layer, where Fibre Channel connections are “detached” and connected to the existing SAN infrastructure. This is one of the options and we move on from here.

https://t.me/learningnets

certcollecion.net



The Cisco Nexus 7000 or 5500 Series Switches can be used as Fibre Channel over Ethernet (FCoE) aggregation switches. These switches consolidate connections from access switches, and send storage traffic over a consolidated FCoE connection to an Cisco MDS 9500 Series Multilayer Director (which must have the FCoE I/O module), or to a FCoE-capable storage device directly.

Note	To connect FCoE hosts to the Cisco Nexus 7000, an exception is made to the rule that one physical port can belong only to one VDC. These ports are shared between the data connectivity VDC and the storage VDC. This functionality is available on F1 and F2 I/O modules.

When using the Cisco Nexus 7000 to aggregate storage connections, maintain the separation between fabric A and fabric B, to provide two independent paths to the SAN.

Note the connections between the access and aggregation switches. These switches are kept separate for data traffic and for storage traffic. Outer connections are storage-only, but use FCoE as their underlying transport. They are kept separate to provide fabric A and fabric B isolation. These connections use their dedicated links so that they do not suffer from bursts of data from the data network.

Each Cisco Nexus 7000 aggregation switch is then connected using FCoE to the Cisco MDS 9500 SAN aggregation or collapsed core layer. An FCoE connection must be used because the Nexus 7000 does not support Fibre Channel connections.

Such a situation is very common, so that you can connect a new data center aggregation network to the existing SAN infrastructure. The only upgrade you need for the Cisco MDS directors is FCoE connectivity.

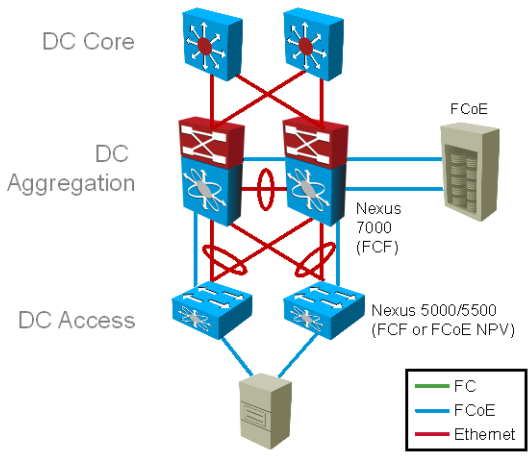
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

Aggregation Layer with Cisco Unified Fabric (Cont.)

- Attachment of FCoE storage to the aggregation layer:



This example is similar to the previous one, with the difference that you do not have an existing SAN network. FCoE storage devices are available, so you can connect them to the access (for small deployments) or to the aggregation layer directly (for larger deployments).

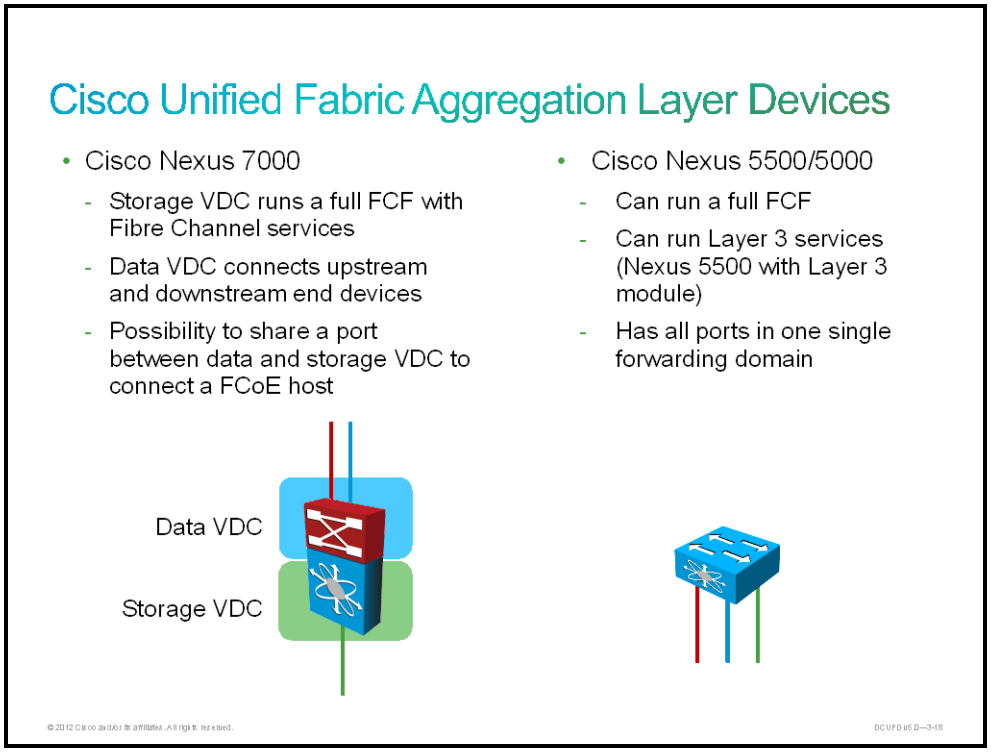
The design of the connection between the access and aggregation layers is maintained. You then connect the storage array directly to Cisco Nexus 7000 aggregation switches using FCoE.

To provide Fibre Channel services and security, the Cisco Nexus 7000 must be a full Fibre Channel Forwarder (FCF), which requires a storage VDC to be configured and appropriate licensing. Using this approach, the access switches can be in N-Port proxy mode, using the FCoE N-Port Virtualizer (NPV).

Using this design, a distinct separation between fabric A and fabric B is still maintained. The traffic of the data network travels across the vPC.

https://t.me/learningnets

certcollecion.net



The approaches of a collapsed core and unified fabric aggregation layer can be combined. You will need three to four VDCs on each Cisco Nexus 7000 Series Switch to implement this design, depending the complexity of your aggregation layer and the kind of services will you be implementing.

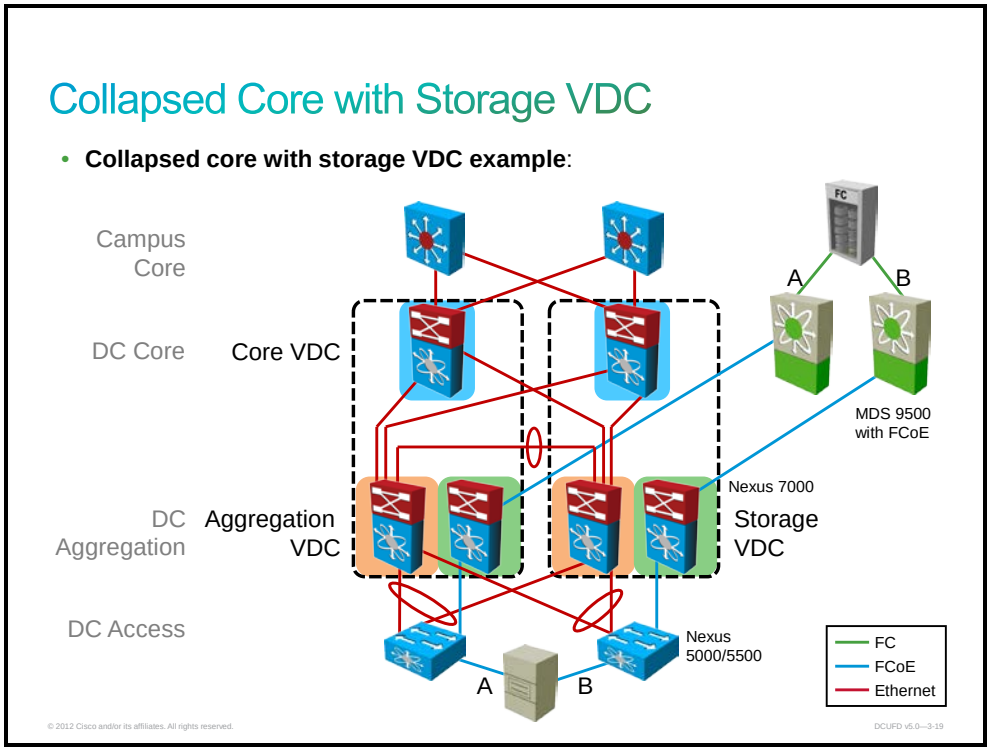
On Cisco Nexus 7000 Series Switches, you need to have storage and data in separate VDCs. Generally, you cannot share interfaces between VDCs, so in this sense the virtual contexts are totally separate. The exception is a shared port, which you need to use if you use the Nexus 7000 as an access switch. The data traffic will be forwarded by the data VDC, and the storage traffic will be forwarded and managed by the storage VDC.

When using the Cisco Nexus 5000 or 5500 Series Switches in the aggregation layer, all types of traffic are combined in a single switching domain. This design is common when extensively using FEXs.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net



This example on the Cisco Nexus 7000 includes a compact core-aggregation block that is provided by the Nexus 7000 Series Switch.

Core VDC

The core VDC provides its usual services, performing IP routing and Equal-Cost Multipath (ECMP) to the campus core, and advertising routes to and from the data center. In this example, the core layer features physical links to the campus core.

Aggregation VDC

The aggregation layer consists of one VDC that is connected to the core VDC with physical connections and performs ECMP Layer 3 forwarding.

There is a link between the aggregation VDCs between two physical Nexus 7000 switches, and it serves as a vPC peer link to extend aggregation VLANs, HSRP, provide routing adjacency, and so on.

Note This VDC can consist of one single or two VDCs, depending on your requirements for the services. It also depends on how you connect the equipment and if you are using Cisco Nexus 7000 M1 or F1/F2 type of I/O modules. You cannot combine FEXs and Cisco Unified Fabric on Cisco Nexus 7000 F1 modules.

Storage VDC

In this VDC, storage protocols are configured, such as the FCF services and the uplinks to the rest of the SAN. This topology includes multihop FCoE. You need to design the SAN according to storage design best practices, to maintain the separation between two distinct fabrics. The storage VDC is not connected to other VDCs and serves the purpose of aggregating storage connections from the aggregation layer, and forwarding packets to the Cisco MDS Fibre Channel switches.

https://t.me/learningnets

certcollecion.net

Storage links from the access layer need to be separate from the data links. This is to maintain fabric A and fabric B separation, and to prevent competition between data traffic and storage traffic. These links are inexpensive (twinax 10 Gigabit Ethernet copper), so cost should not be a concern.

Note **Shared interfaces:** If you are connecting the hosts directly to the Cisco Nexus 7000, you will need to share the access interfaces between the data and storage VDC. Shared interfaces are not applicable in our case, as we are not using the Cisco Nexus 7000 as an access switch for hosts.

Access Layer

The access layer consists of a pair of Cisco Nexus 5000/5500 Series Switches, with hosts connected to them using Cisco Unified Fabric. The switch then separates the data traffic and sends it over vPCs to the aggregation data VDCs, and sends storage traffic through dedicated unified fabric links to the storage VDC.

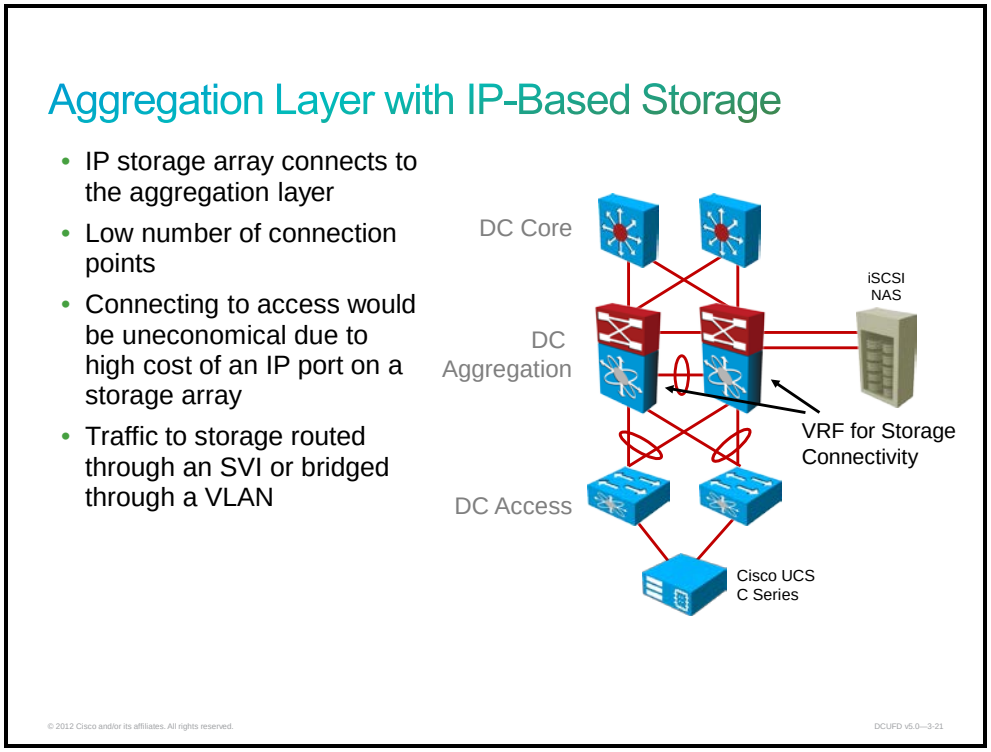
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

Aggregation Layer with IP-Based Storage

This topic describes how to design the aggregation layer with IP storage-related specifics in mind.



One of the options to implement storage is to do it over the IP protocol. IP-based storage uses protocols like Internet Small Computer Systems Interface (iSCSI), Network File System (NFS), and Common Internet File System (CIFS).

iSCSI is a block protocol, where a basic unit of saved data is an SCSI block. NFS and CIFS are file-level protocols, where the minimum unit of saved data is a file. This type of storage manages segmenting the files into blocks internally.

Block storage and file storage attach to the host operating system at different levels and at different points. File-level storage attaches on the file system level, where block storage attaches below the file system implementation in the operating system.

Note	These protocols typically utilize jumbo frames to carry the entire block of data. All data center switches in the path of this data must support jumbo frames.

From the server perspective, IP-based storage can be reached through the main production data network interface card (NIC) of the server, or by using a separate NIC. Some of these NICs also have additional TCP offload capabilities (relevant for iSCSI) to optimize storage packet processing.

The aggregation layer is where storage data is separated from other production data. For this reason, it is a practical point to attach IP-based storage.

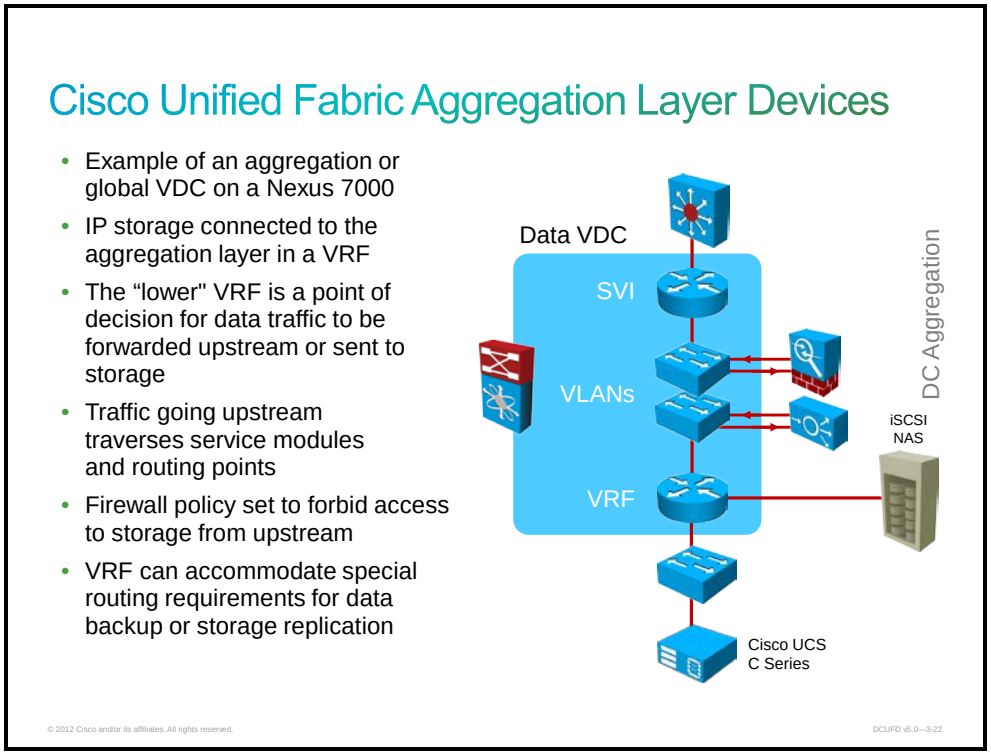
Note	Attaching storage to the access layer is considered less scalable, depending on the number of interfaces that you have on the storage array and the number of access switches.

https://t.me/learningnets

certcollecion.net

Depending on the solution, storage traffic can be isolated in a VLAN (the case if you have multiple NICs, or if you have multiple VLANs to the host), or in a VRF if you do not separate storage from data traffic on the host. The VRF is then used to route storage traffic away, toward the storage array instead of sending it further upstream to the data center core.

Using this approach, you also relieve the core from a high number of packets that use much bandwidth. Storage access is usually bandwidth-hungry and bursty, which makes it difficult to measure and use the correct average values for bandwidth.



When using network-based application services, it does not make sense to route IP storage traffic through these service modules or appliances. This would unnecessarily consume their resources. To divert the storage traffic from the path that production traffic takes, a VRF is created “below” the services, with a static route pointing to the storage subnet to a local interface, connected to the storage array.

On the firewall, you can make sure that the storage array can be accessed only from internal subnets. You can also open up dedicated paths for backup of the storage, and for storage replication, if necessary.

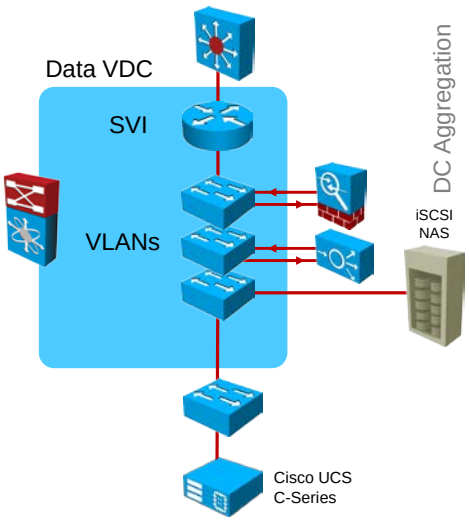
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

Cisco Unified Fabric Aggregation Layer Devices (Cont.)

- Simpler, but less flexible scenario:
  - Traffic from server to storage can be bridged, but requires a VLAN trunk connection to the server
  - If traffic to the IP storage is routed, it traverses service modules before getting to the SVI and then it is bounced back toward the storage
  - Unnecessary load on service appliances or service modules
  - Does not require VRF, no static or dynamic routing required between VRF and SVI



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

A simpler design does not use the VRF. Instead, it uses a flat VLAN where the storage array is connected, and all servers that require access to it. Storage traffic is kept isolated from production traffic with VLAN isolation. You must ensure that you do not extend the storage VLAN toward the (possibly present) service modules, as it may consume their resources unnecessarily.

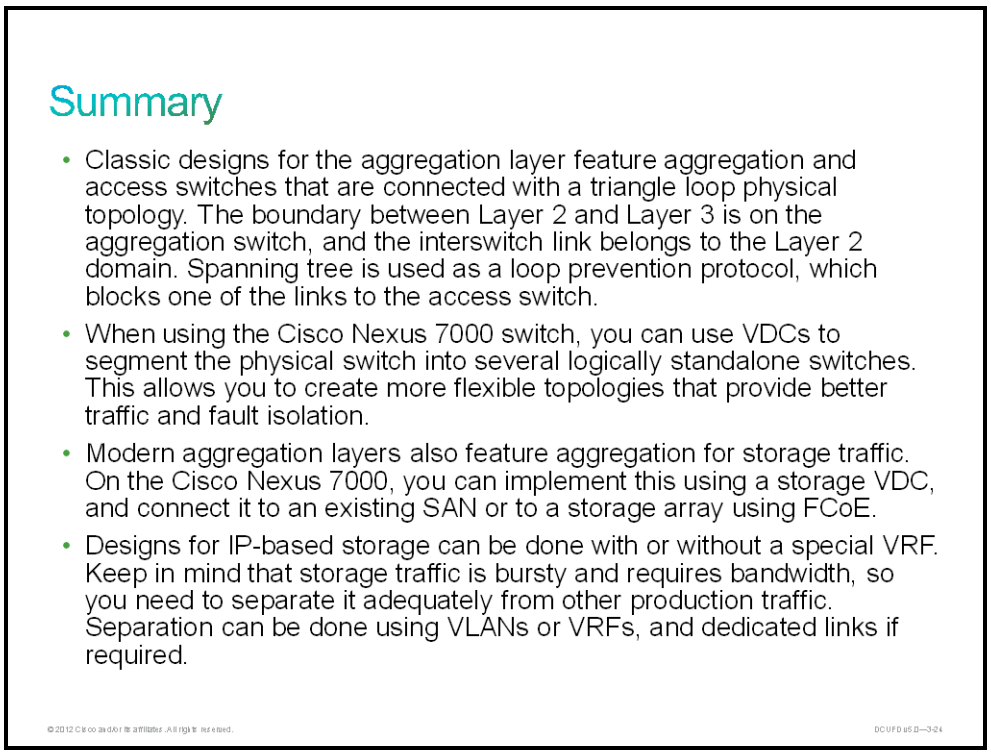
This design is simpler, but it requires either two NICs on the server (one for data, one for storage) or a VLAN trunk toward the server. The server then places storage traffic in the correct VLAN.

https://t.me/learningnets

certcollecion.net

Summary

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



 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

Lesson 3

Designing the Data Center Access Layer

Overview

In this lesson, you will learn about designing the access layer of the data center network. With the introduction of Layer 2 multipathing technologies, the access layer has evolved in recent years. The general goal is to make the access layer as efficient and as cost effective as possible by offering high port density, low oversubscription, and the lowest cost based on the needed features.

Objectives

Upon completing this lesson, you will be able to design the data center physical access layer. This ability includes being able to meet these objectives:

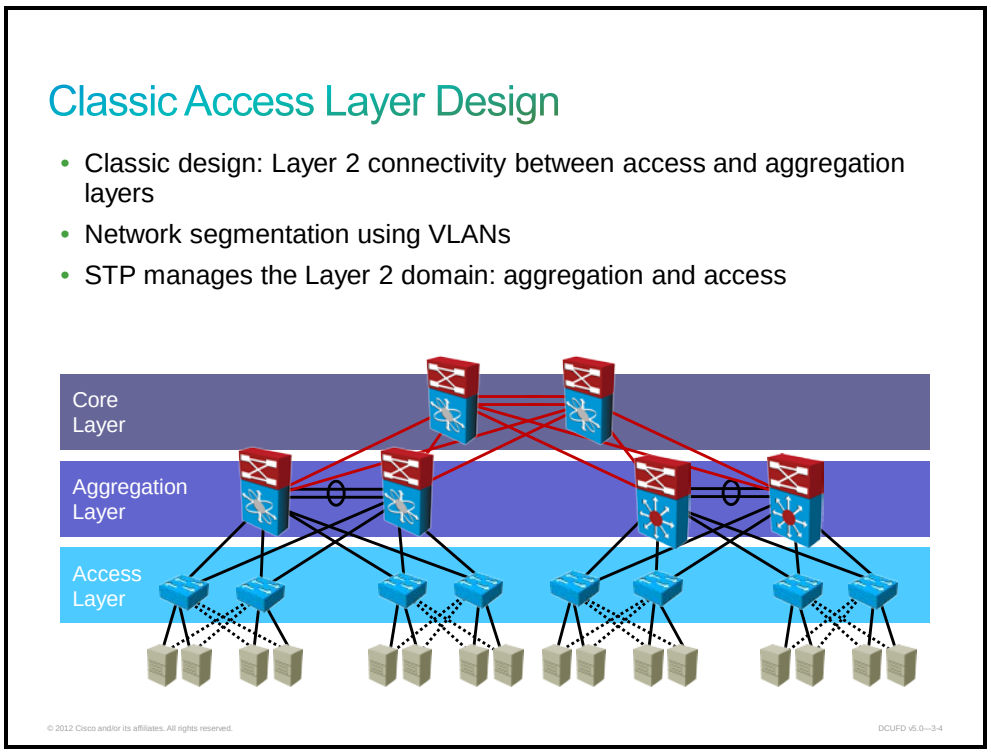
- Describe the classic access layer designs and design issues
- Design the access layer with vPC and MEC
- Design the access layer with FEXs
- Design the access layer with Cisco Unified Fabric

https://t.me/learningnets

certcollecion.net

Classic Access Layer Design

This topic describes the classic access layer designs and design issues.



For a long time, the classic access network design has used access switches that are connected to a pair of aggregation switches using one link to each aggregation switch. The aggregation switches are interconnected between themselves using an interswitch link. Typically, the aggregation switches terminate the Layer 2 domain and forward data traffic toward the core using Layer 3, or toward other aggregation blocks.

Networks are segmented using VLANs, and all needed VLANs are brought to access switches that connect physical servers.

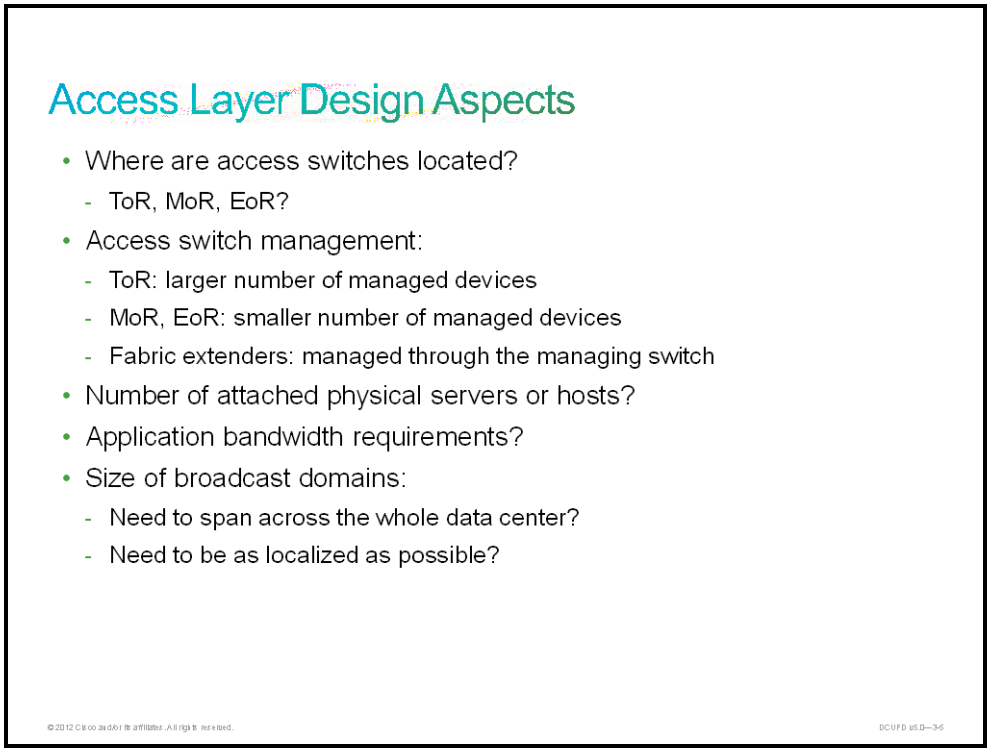
Physically, access switches are typically located as top-of-rack (ToR) switches. Their role is to provide network connectivity to servers that are installed in that rack.

Alternative designs include middle of row (MoR), where a switch is installed in the middle rack and with optimal connections to every server, and end of row (EoR), where an access switch is installed in a rack at the edge of server racks. MoR and EoR simplify management, but require large, modular switches. The ToR approach minimizes cabling because every rack has its local switch.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net



When designing the access layer, there are a few design considerations that the designer should take into account.

The access layer topology determines the placement of the switch and the form of the access layer. The physical topology can be ToR, MoR, and EoR. This defines the equipment type and the cabling.

The access layer utilizing ToR switches has optimized cabling, but in exchange, it has many more devices to manage. EoR and MoR designs have fewer managed devices, but more cumbersome cabling.

Note	A modern design of an access layer using fabric extenders (FEXs) combines all benefits from ToR and EoR designs: optimized cabling and a reduced number of managed devices.

The number of attached devices defines the number of ports that is needed. The number of ports is completely dependent on the application, and on the scale of offered services. These are your design inputs.

The bandwidth requirement is similar. It needs to be defined by the application vendor as a design input parameter. This requirement can be satisfied by using link scaling technologies, such as port channels and virtual port channels (vPCs).

The size of the Layer 2 domains also defines the data center design. If the application requires Layer 2 domains to span across several aggregation blocks, you need to consider Layer 2 multipathing technologies, data center fabrics, or Layer 2 encapsulation technologies.

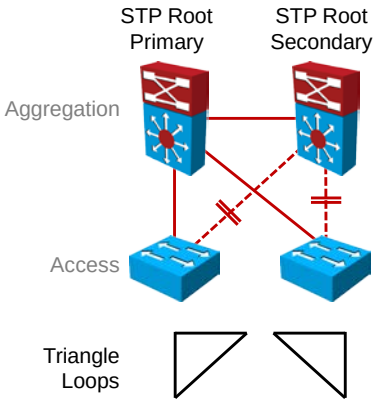
Localized broadcast domains call for Layer 3 access layers, and VLANs that span across multiple switches need Layer 2 data center designs.

https://t.me/learningnets

certcollecion.net

Classic Access Layer Design

- Looped-triangle physical wiring
- STP as loop prevention mechanism:
  - STP on access switches blocks path to secondary root bridge
  - If the link to primary root bridge fails, the link to secondary root bridge begins forwarding
  - If the primary root bridge fails, the link to secondary root bridge begins forwarding
- Main downside: one link always blocked, only 50 percent of installed bandwidth available



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

The classic access layer requires triangle looped physical wiring. This means that an access switch connects to both aggregation switches, which are also interconnected themselves. This forms a triangle, with a potential to create a Layer 2 loop. This is why the Spanning Tree Protocol (STP) is used as a loop prevention mechanism. If STP is set correctly, the link from the access switch to the STP secondary root bridge will be blocked.

STP then manages network convergence in case of topology changes:

- If the link to the primary root bridge fails, STP on the access switch will enable the link to the secondary root bridge.
- If the primary root bridge fails, STP on the access switch will enable the link to the secondary root bridge.

The main downside of the classic access layer design is that one link is always in the blocking state, allowing utilization of only half of the installed bandwidth.

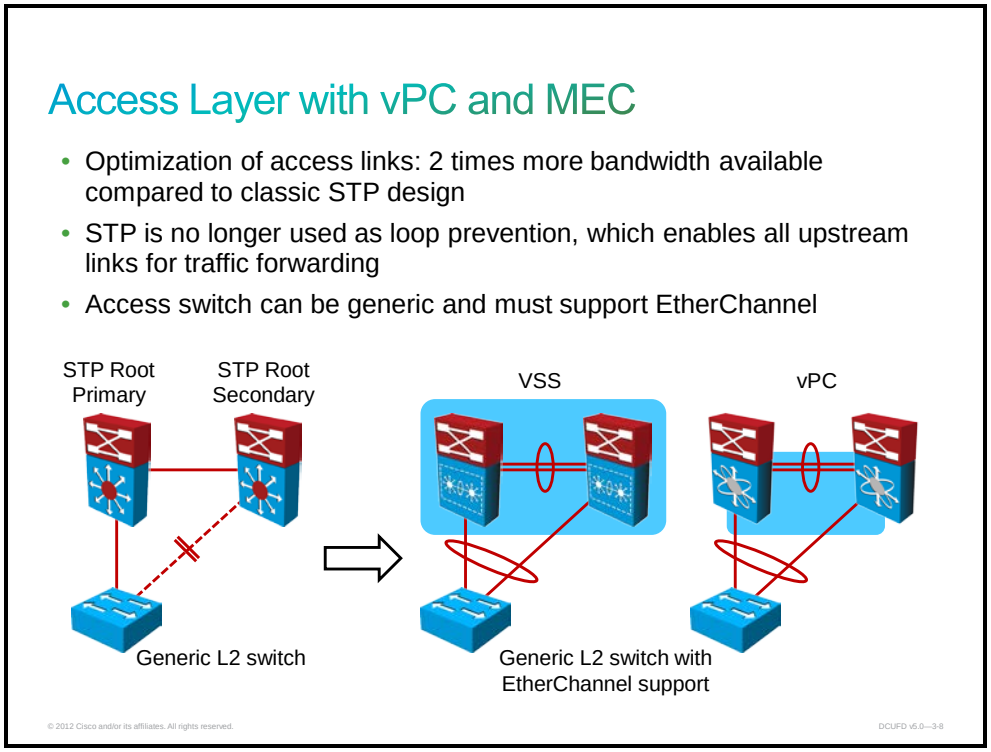
© 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

Access Layer with vPC and MEC

This topic describes how to design the access layer with the vPC and Multichassis Ethernet (MEC).



The classic access layer has a disadvantage that is inherited from STP: the high oversubscription ratio that is caused by one link in the blocking state.

Two technologies are available that can overcome this limitation: vPC and MEC. You use one or the other based on the hardware in the aggregation layer. MEC is used with the Cisco Catalyst 6500 Virtual Switching System (VSS), and vPC is employed when Cisco Nexus series switches are used. The access switch can be generic. It only needs to support EtherChannel.

The physical wiring must be the same—the looped triangle.

With MEC and vPC, all uplinks from access switches can be used, doubling the bandwidth available to the servers, and reducing by half the oversubscription ratio.

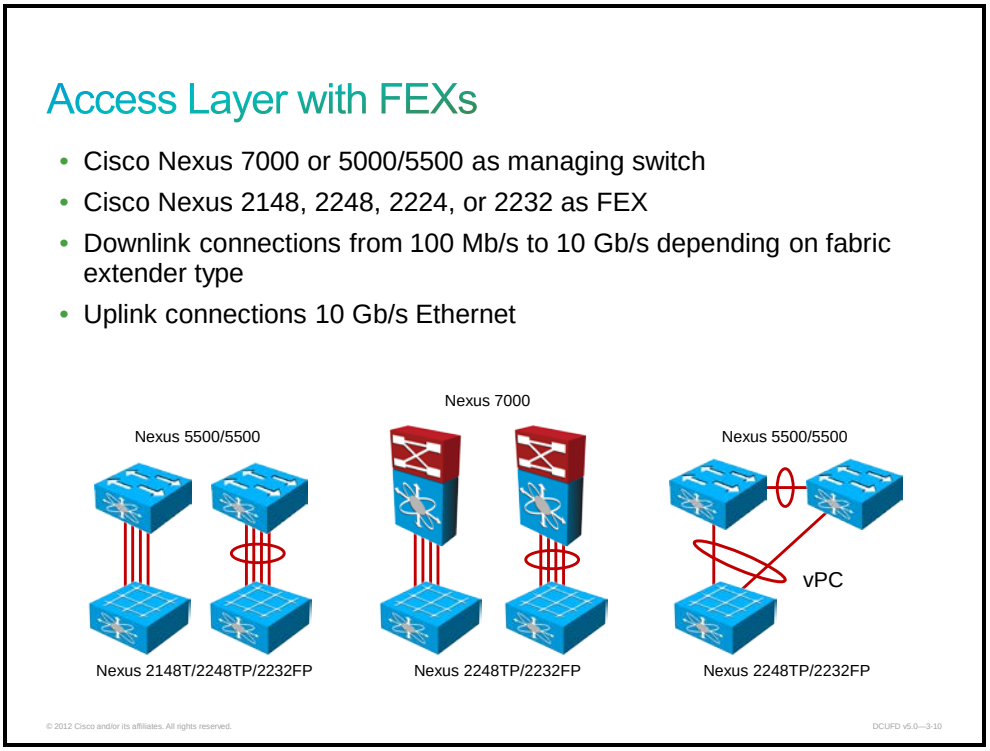
There are differences in how MEC and vPC function on the control plane and data plane levels; this is discussed in detail in a separate lesson.

https://t.me/learningnets

certcollecion.net

Access Layer with FEXs

This topic describes how to design the access layer with FEXs.



FEXs are a cost-efficient way to design the access layer.

FEXs are unmanaged devices. They are managed by their upstream, managing device. This substantially simplifies the network design because the FEX is not an additional logical network element. It does not require a visible management IP address on its own because it is managed through the managing switch.

The FEX is often used to provide low-cost, 1-Gb native Ethernet connectivity to devices that would otherwise consume an “expensive” 10 Gigabit Ethernet port on the Cisco Nexus 5000/5500 or 7000 Series Switches. FEX is an inexpensive solution to connect Fast Ethernet ports in the data center that are typically used for management without requiring a dedicated management switch.

Note	The FEX enables increased access layer scalability without increasing management complexity.

FEX Models

These are the FEX models:

- **Cisco Nexus 2148T FEX:** This is the first FEX and offers Gigabit Ethernet connectivity to servers, and four 10 Gigabit Ethernet upstream connections.
- **Cisco Nexus 2224TP and 2248TP FEXs:** An enhanced version of the original FEX, these extenders offer Fast Ethernet and Gigabit Ethernet downstream connectivity. There are two sizes: the Cisco Nexus 2224 has two uplink ports, while the Cisco Nexus 2248 has four. Additional Fast Ethernet connectivity support was added to connect to server management interfaces.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

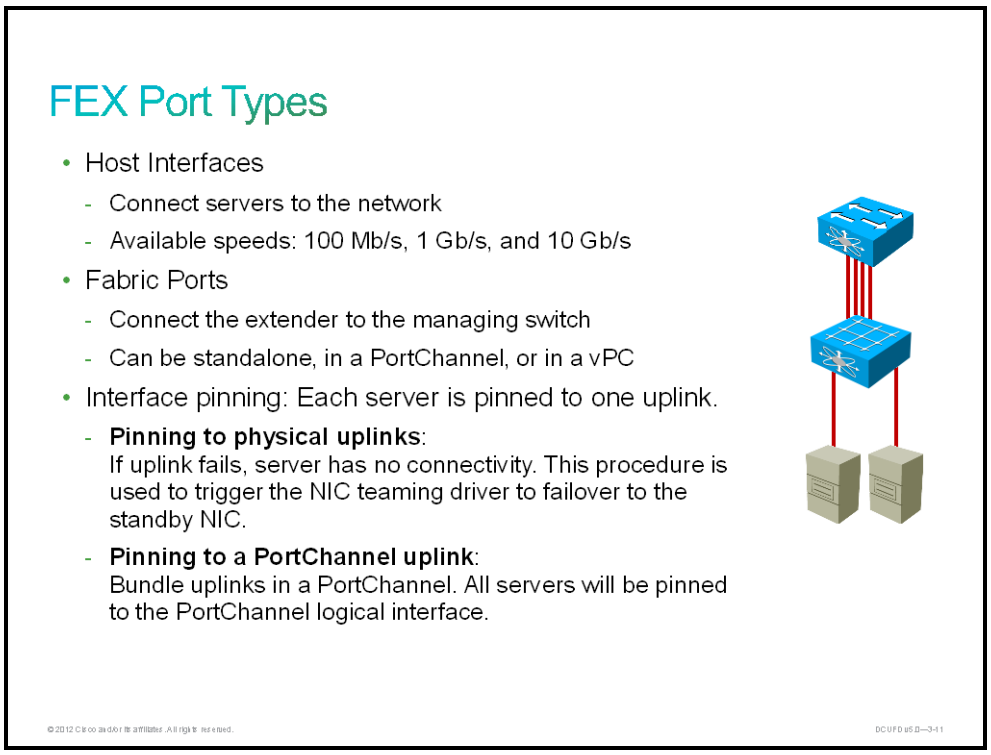
certcollecion.net

- **Nexus 2232PP FEX:** This FEX has 32 10 Gigabit Ethernet Fibre Channel over Ethernet (FCoE) and Data Center Bridging (DCB) ports for downlinks, and eight FCoE and DCB 10 Gigabit Ethernet ports for uplinks. This FEX also supports Cisco Unified Fabric for data and storage connectivity using the extender.

Connecting the FEX

FEXs can be connected using these methods:

- Single or multiple direct FCoE and DCB or native Ethernet connections
- Single port channel connection consisting of multiple physical connections to the same managing switch
- vPC connection, connecting the FEX to multiple upstream switches for additional redundancy and load balancing



Ports on the FEX are of two types: server ports and uplink ports.

Host Interfaces

These ports connect the servers to the network. In some cases, an additional switch can be connected to a host or server port to additionally or temporarily further extend the network.

Fabric Ports

Uplink ports connect the FEX to the network managing switch or switches. The managing switch performs all switching functions for devices that are attached to the FEX. Even if devices are connected to the FEX, traffic between them needs to be switched by the managing switch.

https://t.me/learningnets

certcollecion.net

Note This makes the FEX less desirable for servers that have much traffic between them, such as clustered servers, application serves, and so on. When using FEXs, the best solution is to connect servers that have much upstream traffic (“North-South”), and little traffic between themselves (“East-West”). Such servers are still best connected using a proper switch.

Interface Pinning

When connecting a server to an FEX that has standalone upstream connections to the managing switch, the extender performs the server interface pinning. Interface pinning can be static or dynamic. Each server is assigned its own upstream interface that it uses to send and receive traffic. In case of failure of that link, the server loses connectivity. For the server to regain connectivity, host interfaces need to be repinned.

Note The FEX does not automatically repin lost connections to a working link. This behavior is designed to trigger the NIC teaming driver on the server to switch over to a standby NIC.

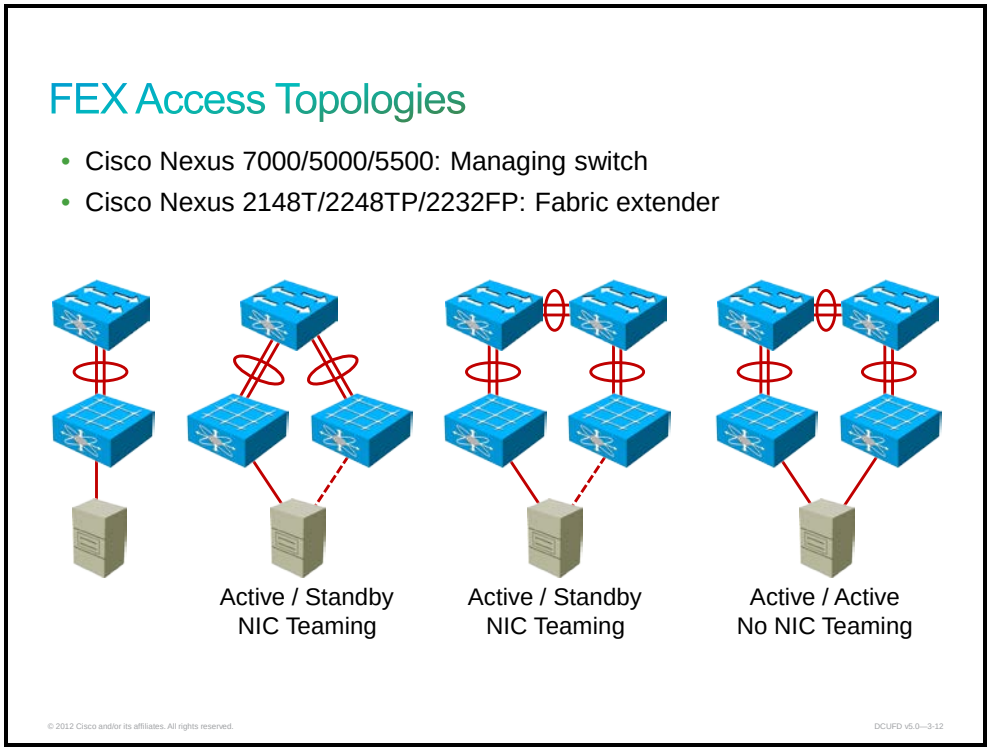
The load-balancing paradigm for this case is per-server load balancing. Load distribution on the links depends greatly on the amount of traffic that each server produces.

Another option is to bundle the uplinks into a port channel. Port channel achieves better load balancing and has better resiliency. All servers are pinned to the same port channel interface, which does not go offline until at least one physical interface (a member of port channel) is still functioning.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net



There are several topologies that are available when using FEXs that depend on the requirements for servers and the physical topology of the network. Later, if the server has multiple connections to the network, it depends on the mode in which these connections operate. Having several network interface cards is called network interface card (NIC) teaming. Teaming can work in active/standby or active/active operating regime.

The topologies with FEXs and the Cisco Nexus 7000 and 5000/5500 Series Switches that are supported depend on the NIC operating regime.

- The first connection in the figure is the basic setup. The FEX is connected by two physical connections (forming a port channel) to the network. All servers pin to this logical uplink.
- In the second example, a server is connected to two FEXs, which are managed by a single managing switch. In this case, the NICs are allowed to operate in active/standby regime.
- The third example shows a connection where a server is connected to every FEX using a link, but these connect to a pair of switches. In this mode, teamed NICs can operate in active/standby mode.
- The last example shows an active/active connection. In this case, the server uses MAC address pinning and forwards traffic for various internal and virtual MAC addresses through one NIC only.

Note	The Nexus 7000 Series Switches support attachment of Nexus 2232 FEXs that support Cisco Unified Fabric. The line cards on the Nexus 7000 must be with the F2 forwarding engine, and the Cisco Nexus Operating System (NX-OS) needs to be version 6.0. All links can be native Ethernet or DCB.

https://t.me/learningnets

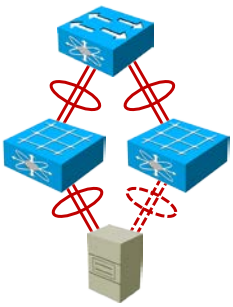
certcollecion.net

FEX Access Topologies (Cont.)

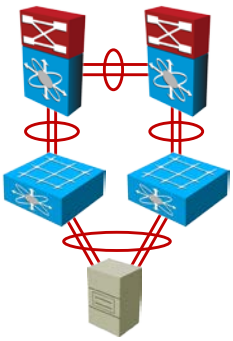
- Nexus 7000/5000/5500: Managing switch
- Nexus 2248TP/2232FP: Fabric extender



Server PortChannel
Active / Active
NIC Teaming



Server PortChannel
Active / Standby
NIC Teaming



Server PortChannel
Active / Active
NIC Teaming

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

The following designs are supported when using a port channel connection from the FEX to the server, both on the uplink side and on the downlink side of the FEX.

- The first example extends the port channel from the FEX to the server.
- The second example supports an active/standby NIC teaming scenario with the server, with a port channel extended to the server.
- The third example supports an active/active NIC teaming scenario, with a vPC extended to the server.

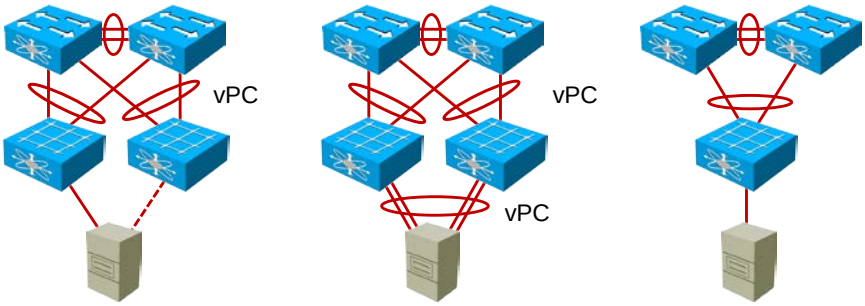
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

FEX Access Topologies (Cont.)

- Nexus 5000/5500: Managing switch
- Nexus 2248TP/2232FP: Fabric extender



Dual-homed FEX
Active / Standby
NIC Teaming

Enhanced vPC – **EvPC**
Active / Active
(N5500, N2200 only)

Dual-homed FEX
with single-NIC
server

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

The following designs are supported when using vPC between the managing switch and the FEX. These designs can be used when the managing switch is a Cisco Nexus 5000 or 5500 Series Switch.

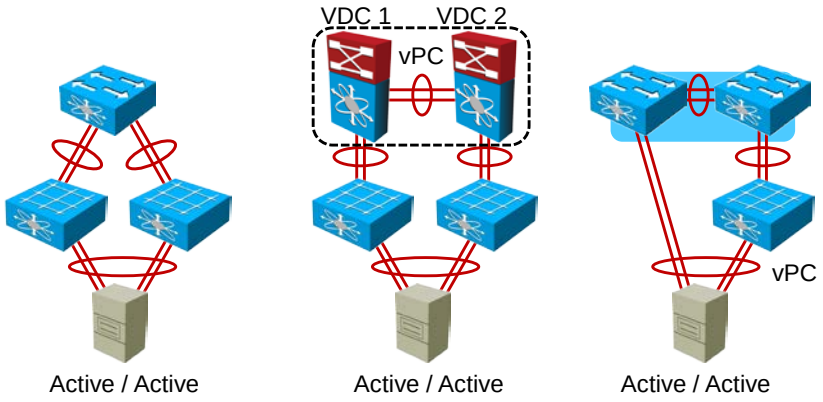
- The first example dual-homes the FEX to two Cisco Nexus 5000 managing switches. The server then connects in an active/standby manner using simple NIC teaming.
- The second example connects the FEX using a vPC to the managing switches, and additionally forms a vPC to the server. In this case, the server can have either two active connections to the network, or have an active/standby connection.
- The last example multihomes the FEX, and connects the server with a single NIC.

https://t.me/learningnets

certcollecion.net

FEX Access Topologies (Cont.)

Unsupported topologies:



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

This slide lists the topologies that are *not* possible:

- The first example would, in reality, form a port channel from the managing switch to the server, over two FEXs.
- The second example is not possible because two virtual device contexts (VDCs) of the same Nexus 7000 chassis cannot form a vPC.
- The third example is not possible due to an asymmetrical connection. The server would connect directly to the managing switch using one NIC, but over an extender using the second NIC, and try to form a vPC.

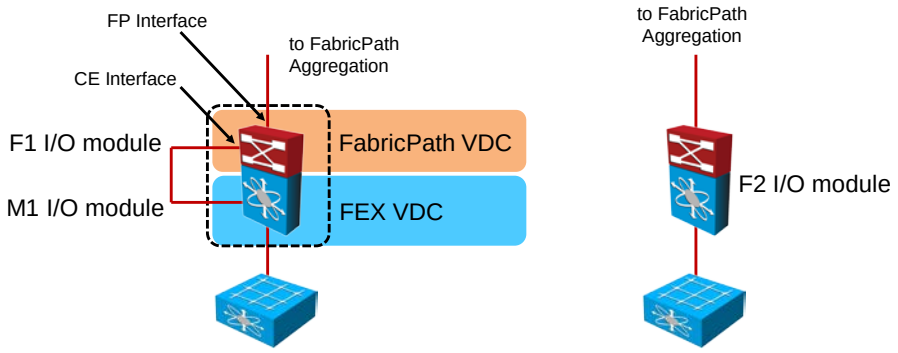
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

Using Fabric Extenders and Cisco FabricPath

- Combination of FEX and FabricPath:
  - F1 and M1 modules need to work in different VDCs, which need to be interconnected with a physical link.
  - F2 modules support FabricPath and FEX in the same VDC.



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

Modern data center designs may use Cisco FabricPath in the aggregation layer due to its advantages regarding network resiliency, load balancing, and its ability to use multiple paths between any two switches in the fabric.

Due to hardware design reasons, there are two ways to implement FEXs in such a network when using a Cisco Nexus 7000 Series Switch as the access switch.

Note	Such setup is not very common, but takes advantage of high port density and low price of 10 Gigabit Ethernet ports offered by the Cisco Nexus 7000.

- The first scenario uses a Cisco Nexus 7000 switch with two types of I/O modules installed. The F1I/O modules support Cisco FabricPath, while the M1 I/O modules support connecting FEXs.

Note	Only the M1 32-port 10 Gigabit Ethernet I/O modules can connect FEXs. The 8-port, 10 Gigabit Ethernet module and the 48-port, 1 Gigabit Ethernet I/O module do not have this ability.

  Another limitation is that you need a dedicated VDC to face the Cisco FabricPath network, so the only way of using both FabricPath toward the network and FEXs to the servers is to create two VDCs and link them with a physical cable.

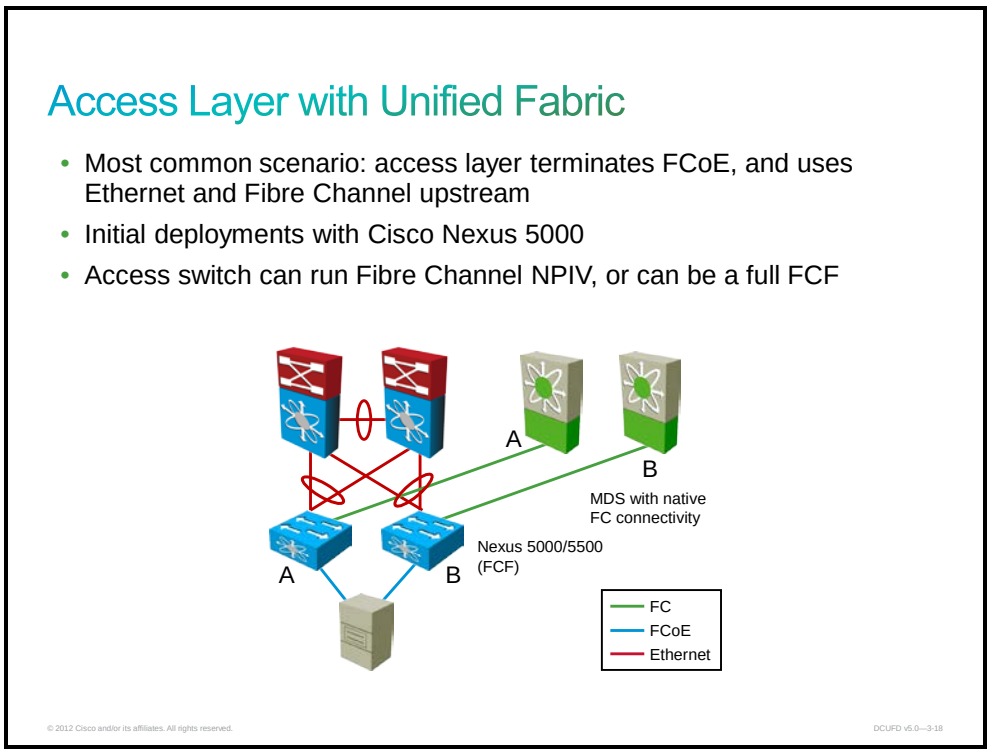
- The second scenario uses the F2 I/O module, so connecting both types of ports to the same switch is not an issue and is supported.

https://t.me/learningnets

certcollecion.net

Access Layer with Unified Fabric

This topic describes how to design the access layer with unified fabric.



The most traditional way to implement unified fabric in the access layer is to connect the servers with Converged Network Adapters (CNAs) to two access switches with a UF connection. On the access switch, you may or may not run Fibre Channel services.

- You can use the switch as a full Fibre Channel Forwarder (FCF)
- You can use the switch in N-Port Virtualizer (NPV) mode, and rely on the upstream Fibre Channel switch to provide Fibre Channel services.

Note	Such were the initial deployments with Cisco Nexus 5000 as an FCoE access switch. The benefit of this design is simplicity, wide support, and easy integration into an existing Fibre Channel SAN.

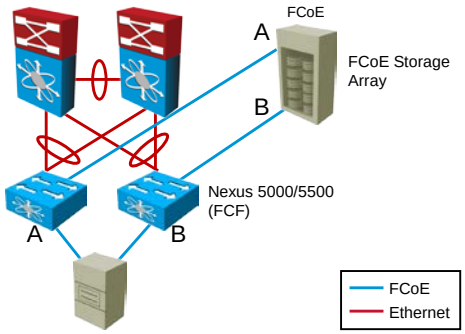
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

Access Layer with Unified Fabric (Cont.)

- Option with Unified Fabric only: directly attached FCoE storage array
- Access switch is a full FCF to support advanced Fibre Channel configuration
- Some FCoE storage arrays may work without an FCF
- Keeps distinct separation of Fabrics A and B



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

Another option is not to terminate FCoE on the access switch, but instead carry it forward to either an FCoE-capable storage array, or upstream to the aggregation switch.

Note For this design example, see the “Designing the Data Center Aggregation Layer” lesson.

From the access switch perspective, there is not much difference. The access switch may be the full FCF if Fibre Channel services are needed (that is, when attaching directly to FCoE storage array), or may run in FCoE NPV mode if the upstream aggregation switch is the FCF.

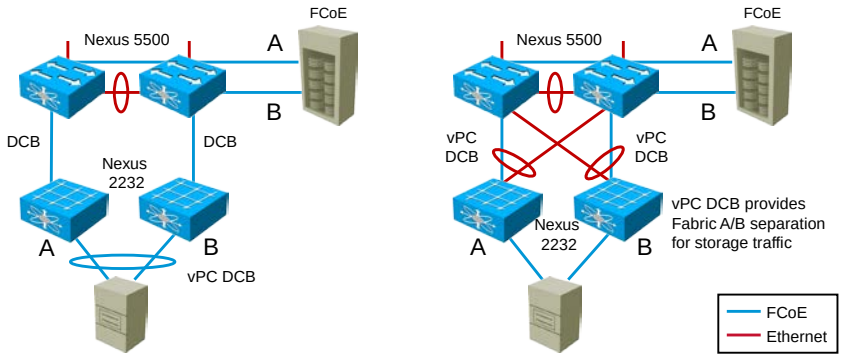
Keep in mind that you need distinct separation between Fabric A and Fabric B to provide two independent paths from the server to the storage array.

https://t.me/learningnets

certcollecion.net

Access Layer with Unified Fabric and FEX

- Option with Unified Fabric only and a fabric extender
- Access switch is a full FCF to support advanced Fibre Channel configuration if using a directly attached FCoE storage array
- Dual-homing the server to 2 FEXs keeps distinct separation of Fabrics A and B



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

You can use the Cisco Nexus 2232PP FEX to connect the server to the network using the CNA. The Cisco Nexus 2232 supports 10 Gigabit Ethernet FCoE/DCB host interface (HIF) downlinks.

If you need Fibre Channel services, you can use the access (FEX managing) switch as a full FCF. If connecting directly to an FCoE-capable storage array, you typically need full Fibre Channel services on the access switch. If connecting to another FCoE/DCB FCF switch upstream, the access switch can be in FCoE NPV mode.

- In the first example, the server is dual-homed to two FEXs to provide two independent paths to the storage array. Both links carry data and storage traffic.
- In the second example, the server is dual-homed to two FEXs, which then separate data traffic and transport it over a vPC (to provide better load balancing). Storage traffic is transported over the direct uplinks within the vPC. The vPC uses the cross links for native Ethernet data only, while the direct links are used for data and storage traffic.

  The access switch then separates data traffic from storage traffic and sends each type of traffic to its appropriate upstream connection.

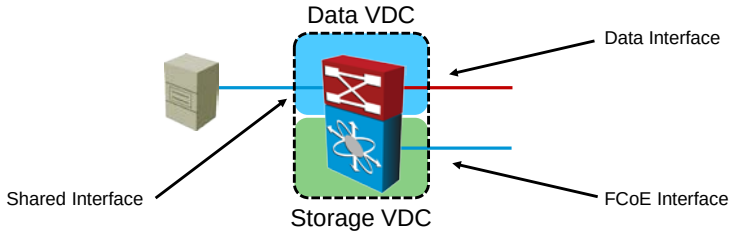
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

Access Layer with Unified Fabric on Nexus 7000

- Nexus 7000 as MoR Layer 2 access switch
- Nexus 7000 has distinct separation between data and storage traffic
  - Data VDC carries data traffic
  - Storage VDC runs the FCF
- Server connects to an interface that is shared between two VDCs. The FCoE VLAN is added to data VLANs.
- Data VDC connects upstream to data network, storage VDC connects upstream to FCoE network (Nexus 5000/5500/7000, MDS 9500, FCoE storage array)



F1/F2 I/O module

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

The Cisco Nexus 7000 has a distinct configuration when used as a unified fabric switch. The FCF needs to be run in its own VDC, while data traffic is forwarded through regular data VDCs.

The server connects to the switch using a CNA, and the switch port that terminates this connection is shared between the data VDC and the storage VDC. The FCoE VLAN is relayed to the storage VDC, the data VLANs remain in the data VDC.

The storage VDC then connects upstream to either a FCoE-capable storage array, or to a Nexus 5000/5500/7000 FCF, or to an Cisco MDS 9500 Series Multilayer Director with an FCoE I/O module.

Only FCoE connectivity is possible on the Nexus 7000 because the Nexus 7000 does not have native Fibre Channel ports.

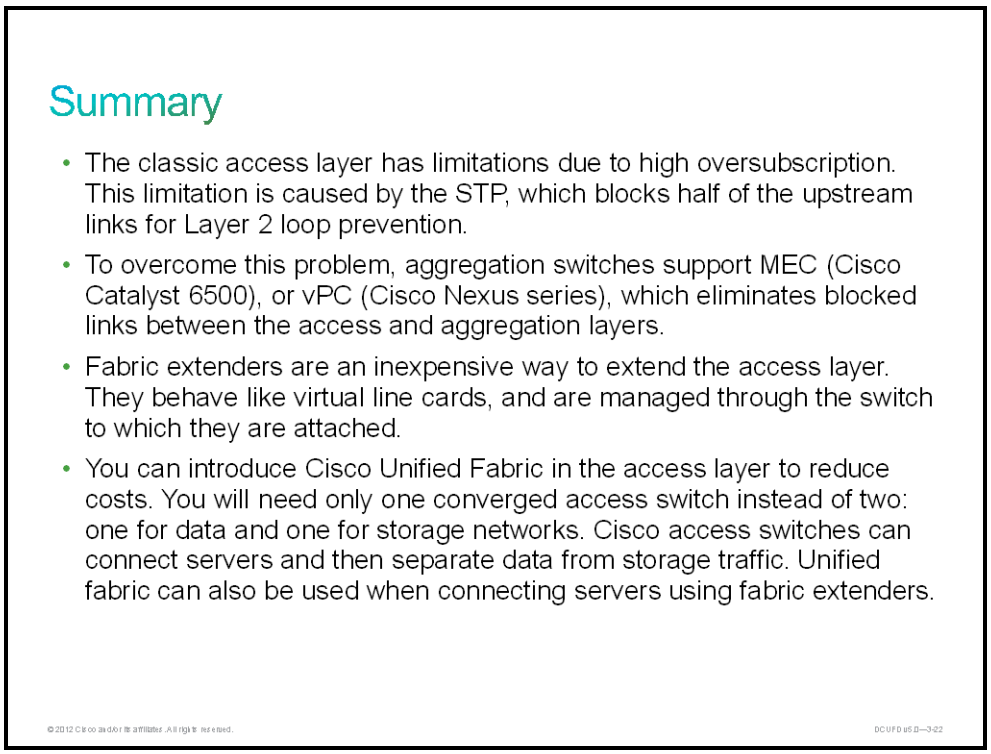
Note This setup is very suitable for high-performance standalone servers, which may use the Cisco Nexus 7000 as a MoR access switch. The most suitable configuration of the switch is the Nexus 7009 because it is in a very compact factor suitable for MoR deployment, and has high port density. In an MoR deployment, most servers can be reached using 10 Gigabit Ethernet copper-based twinax cables, which can be up to 10 meters long and are inexpensive.

https://t.me/learningnets

certcollecion.net

Summary

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



 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

Lesson 4

Designing the Data Center Virtualized Access Layer

Overview

With the usage of server virtualization growing, more focus is put on switching traffic between virtual machines. Initially, this situation was entirely managed by the hypervisor software, which provided poor visibility into network traffic by network administrators. Cisco is extending the network infrastructure into the virtualized hosts to provide more control of traffic flows that would otherwise be contained within the host.

Objectives

Upon completing this lesson, you will be able to design the data center virtual access layer and related physical connectivity, and describe scalability limitations and application impact. This ability includes being able to meet these objectives:

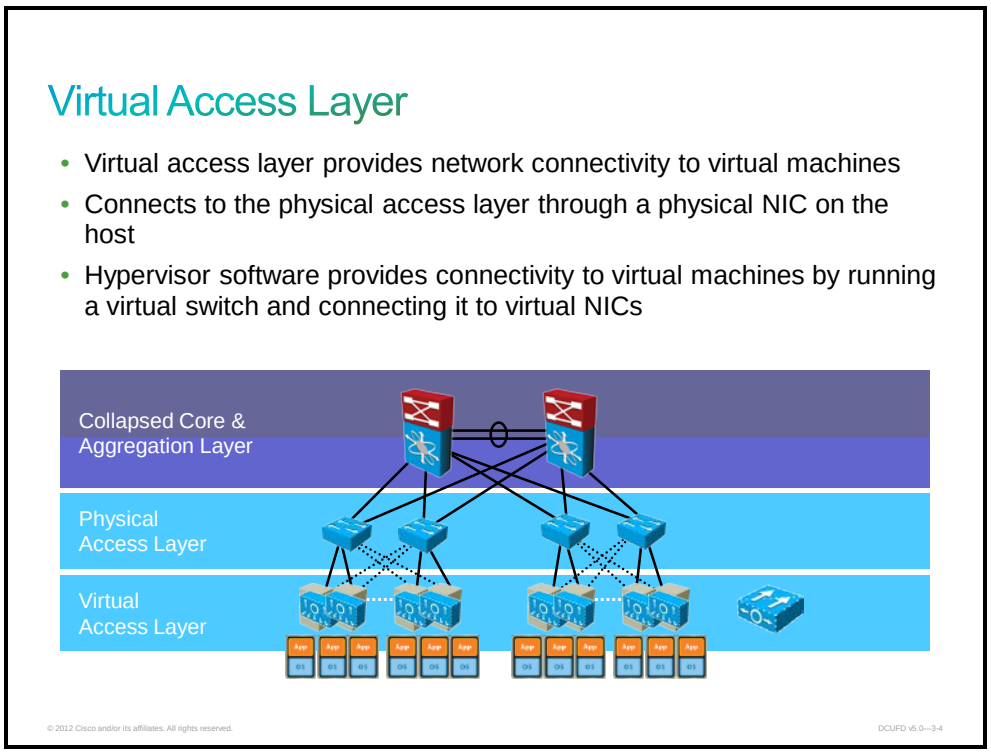
- Define the virtual access layer
- Describe the virtual access layer solutions within virtual machine hosts
- Design solutions with Cisco Adapter FEX
- Design solutions with Cisco VM-FEX
- Design solutions with the Cisco Nexus 1000V switch

https://t.me/learningnets

certcollecion.net

Virtual Access Layer

This topic describes the virtual access layer.



The virtual access layer is below the physical access layer and its main role is to provide network connectivity to virtual machines inside virtualized hosts.

Virtual machines run inside virtualized hosts. This environment is controlled by the hypervisor, which is a thin, special-purpose operating system. To manage network connectivity for and between virtual machines, hypervisors have virtual switches, which is software that presents network interface cards (NICs) to virtual machines and manages the data between the virtual machines and the physical network.

This virtual access layer is connected to the physical access layer through physical NICs installed on the host. To the physical network infrastructure, such a host appears as a device with multiple MAC addresses that appear at the same port.

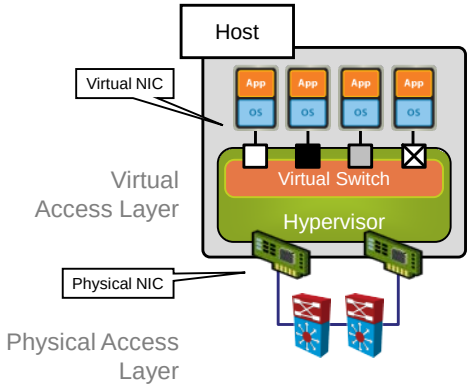
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

Virtual Access Layer (Cont.)

- Virtual access switch runs inside the host
- Embedded in the hypervisor
- Connects virtual NICs on virtual machines
- Processes packets and sends them to the physical NIC if the destination MAC address is outside the host
- Switches packets between virtual machines if the destination MAC is inside the host
- Different options for VLAN tagging:
  - Virtual switch tagging
  - No tagging
  - Virtual guest tagging



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

The example shows a generic virtualized host, with the hypervisor and virtual machines. The hypervisor runs the virtual switch as well, which is a network element of the virtual access layer.

When the virtual switch receives a packet, it determines what to do with it. The virtual switch with forward the packet through the physical NIC to the physical infrastructure or forward the packet to a virtual machine.

VLAN Tagging

The virtual switch can be set to tag the Ethernet frames that it receives with a VLAN tag, or not. The virtual switch can do the following with frames:

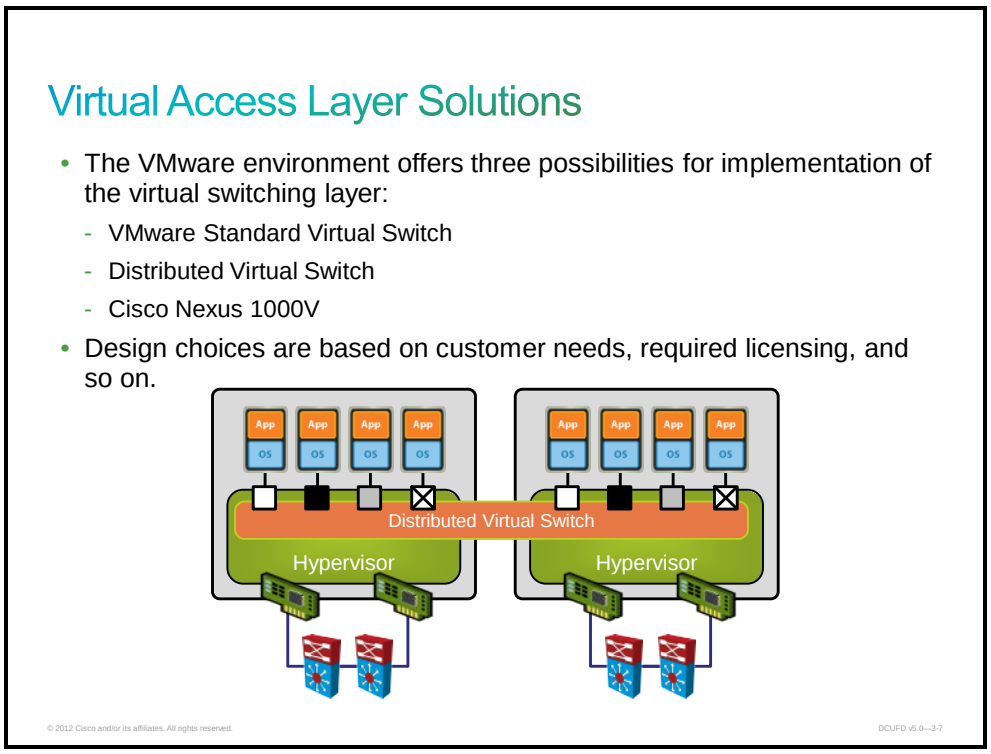
- **Virtual switch tagging:** The virtual machine gets an access port, which means that the virtual switch, upon receiving the frame, imposes the VLAN tag. This approach is the most common.
- **No tagging:** The virtual switch does not perform any tagging. The frame is forwarded as it is to the physical switch. The physical switch then adds the VLAN tag, after receiving the frame on its access port.
- **Virtual guest tagging:** The virtual machine has a trunk port and imposes the VLAN tag.

https://t.me/learningnets

certcollecion.net

Virtual Access Layer Solutions

This topic describes virtual access layer solutions within virtual machine hosts.



The general design is the same for most virtualization environments that are based on the hypervisor model.

In the case of the VMware vSphere solution, the hypervisor component is the VMware ESXi software. Within ESXi, you have two virtual switching solutions:

- VMware standard virtual switch
- VMware distributed virtual switch

The difference between them is that the standard switch is a standalone network element within one host, where the distributed virtual switch spans across multiple hosts and acts as a single switch.

Note	The distributed virtual switch (DVS) needs the vCenter Server component to function. The vCenter Server is part of the VMware vSphere solution.

The third option for VMware vSphere is an extension to the Cisco Nexus 1000V Distributed Virtual Switch. The Nexus 1000V provides additional functionality in addition to the VMware DVS, including the following:

- Cisco Nexus Operating System (NX-OS) CLI
- Centralized manageability by providing network administrators with a way to manage virtual networks
- Improved redundancy, and so on

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

Customers choose their virtual switching layer based on their requirements:

- Does the customer require virtual machine mobility?

  All three solutions support it. Distributed switches retain counter values upon virtual machine motion, and do not require separate configuration of every host.

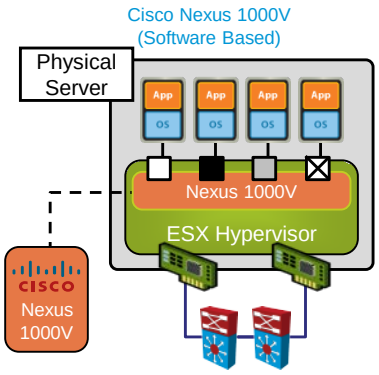
- Does the customer require network administrator access and the Cisco NX-OS CLI for improved management?

  The Cisco Nexus 1000V solution supports this requirement.

Virtual Access Layer Solutions (Cont.)

Cisco Nexus 1000V

- Distributed virtual switching solution
- Virtual Supervisor Module and Virtual Ethernet Modules
- Leverages functionality of VMware VDS and adds Cisco functionality
- Policy-based VM connectivity
- Mobility of network and security properties
- Nondisruptive operational model



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

Cisco server virtualization uses technology that was 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 to enable automated network centralized management.

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.

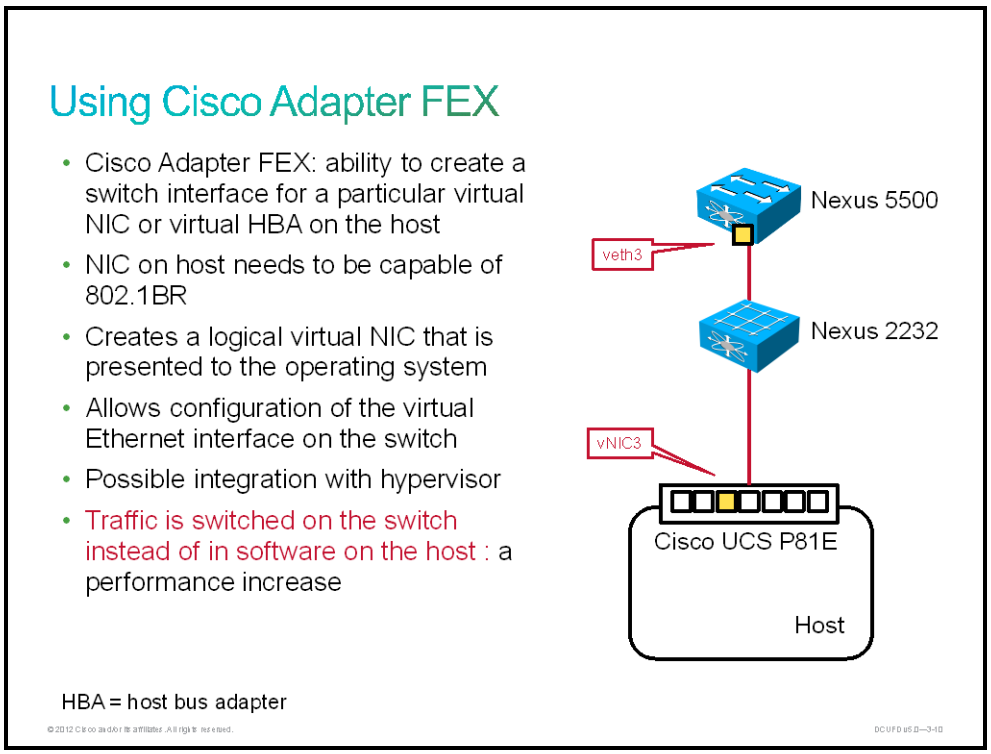
The Cisco Nexus 1000V is a software-based solution that provides VM-level network configurability and management. The Cisco Nexus 1000V works with any upstream switching system to provide standard networking controls to the virtual environment.

https://t.me/learningnets

certcollecion.net

Using Cisco Adapter FEX

This topic describes how to design solutions with Cisco Adapter Fabric Extender (FEX).



Cisco Adapter FEX is a technology that is available for a combination of Cisco equipment:

- Cisco Nexus 5500
- Cisco Nexus 2232 FEX
- Cisco Unified Computing System (UCS) P81E virtualized CNA

The solution allows the NIC to create a virtual interface on the upstream switch for that particular dynamically created vNIC on the host. The virtual NIC (vNIC) is presented to the host operating system as a physical NIC, and presented to the switch as a virtual Ethernet interface. This situation allows you to directly configure parameters for that interface or vNIC on the Cisco Nexus 5500 Series Switch, including assigning VLANs, quality of service (QoS), access lists, and port profiles.

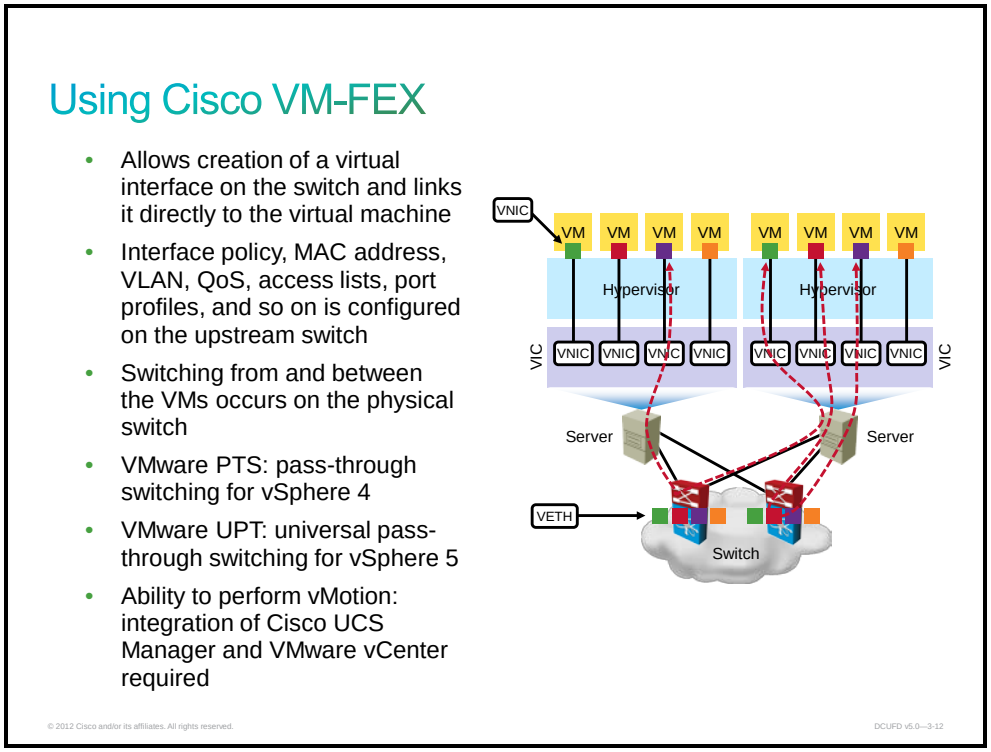
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

Using Cisco VM-FEX

This topic describes how to design solutions with Cisco Virtual Machine Fabric Extender (VM-FEX).



Port-extension-like functions with Cisco VM-FEX. The voice interface card (VIC) is the first implementation of VM-FEX technology from Cisco.

The VM-FEX technology eliminates the vSwitch within the hypervisor by providing individual virtual machine virtual ports on the physical network switch. Virtual machine I/O is sent directly to the upstream physical network switch, which takes full responsibility for virtual machine switching and policy enforcement.

This approach leads to consistent treatment for all network traffic, virtual or physical. Cisco VM-FEX consolidates virtual and physical switching layers into a single layer and reduces the number of network management points.

The Cisco VM-FEX solution offers higher performance than the DVS or Cisco Nexus 1000V because the host CPU is not involved in switching network traffic from and between the VMs. The traffic between the VMs is switched on the switch, between the virtual Ethernet (vEthernet) interfaces.

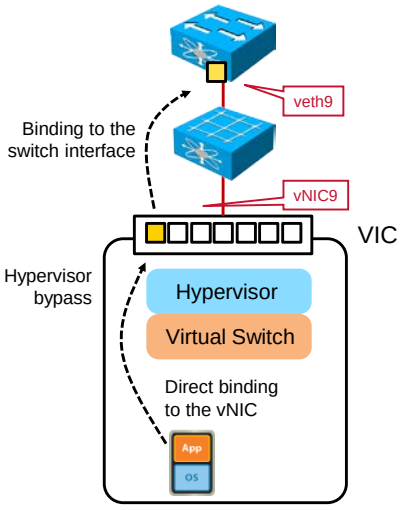
Workload mobility (VMware VMotion) is possible because the VMware vSphere environment and the Cisco UCS Manager environment are integrated and process a move of a vNIC from the VIC on one host to the VIC on another host. The change is also registered on the Cisco UCS Fabric Interconnect, which then moves the binding of the vEthernet interface from one physical downlink (leading to the first VIC) to the new physical downlink (leading to the second, new VIC).

https://t.me/learningnets

certcollecion.net

Hypervisor Bypass

- In some cases, allows for hardware forwarding of traffic between the VM and the switch without involvement of the hypervisor
- Requires customized hardware drivers to be installed into the VM guest operating system
- VM communicates directly with hardware, bypassing the supervisor
- vNIC is registered as a logical interface on the physical switch—the UCS Fabric Interconnect
- VM does not have ability to use vMotion
- Higher performance than PTS/UPT and DVS (switching in software)



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

The hypervisor bypass allows for hardware forwarding of traffic between the VM and the switch without involvement of the hypervisor. The virtual machine uses a customized driver that is installed in the VM guest operating system to communicate directly with virtual hardware, which is the virtual NIC created for that virtual machine on the VIC.

The VIC is then registered to the physical switch in the same way as when using Cisco VM-FEX.

This solution offers even higher performance compared to Cisco VM-FEX, where the hypervisor is still involved to some extent to manage VM network traffic. The hypervisor bypass networking completely bypasses the hypervisor and the VM communicates directly with the hardware.

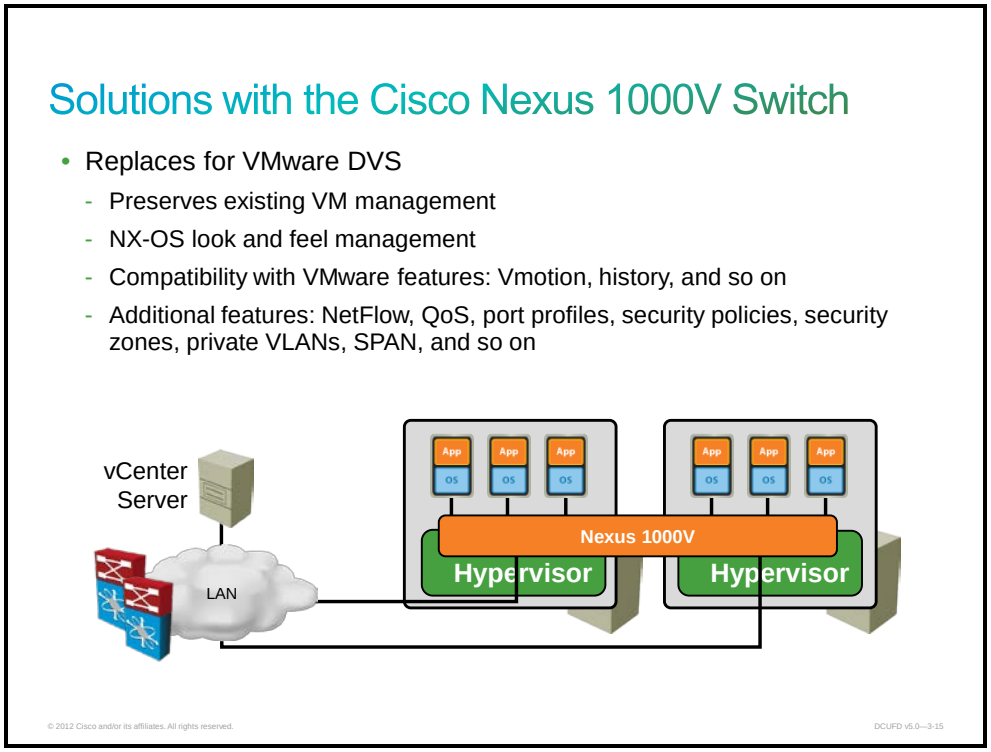
Because of direct hardware attachment of the VM, the capability of moving the VM to another host while the VM is in operation is not available. The VM is tied to the host.

https://t.me/learningnets

certcollecion.net

Solutions with the Cisco Nexus 1000V Switch

This topic describes how to design solutions with the Cisco Nexus 1000V Distributed Virtual Switch.



Cisco Nexus 1000V Series Switches are virtual distributed software-based access switches for VMware vSphere environments that run the Cisco NX-OS operating system. Operating inside the VMware ESX hypervisor, the Cisco Nexus 1000V Series supports Cisco VN-Link 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 a networking environment of multiple ESX hosts.

When server virtualization is deployed in the data center, virtual servers typically are not managed in 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.

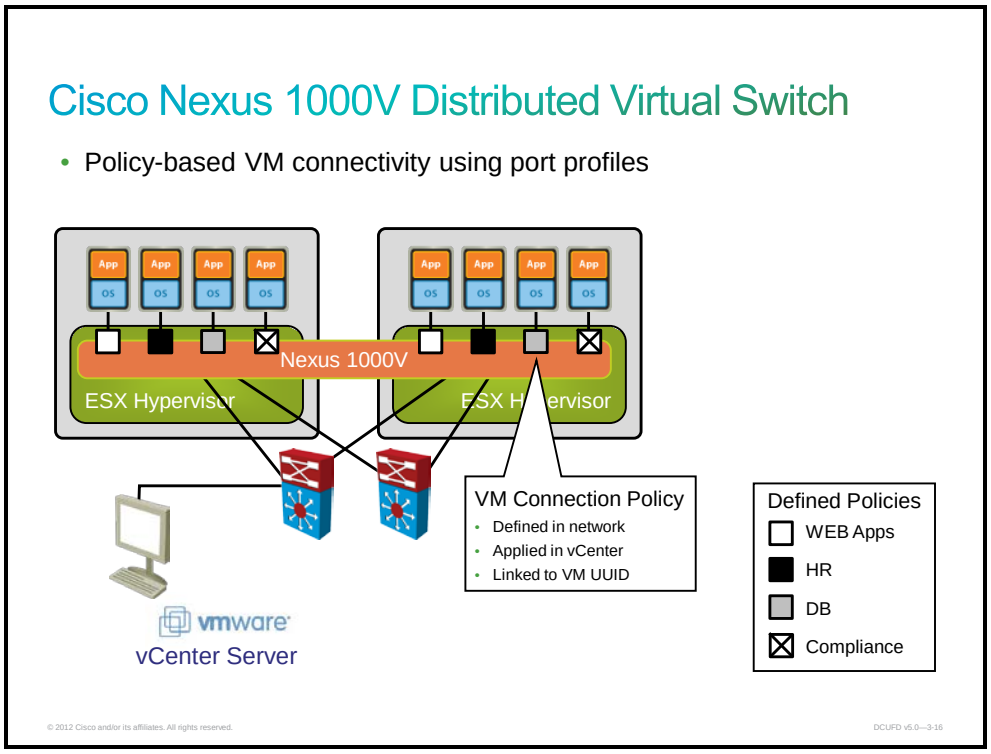
With the Cisco Nexus 1000V Series, you can have a consistent networking feature set and provisioning process all the way from the VM access layer to the core of the data center network infrastructure. Virtual servers can now leverage the same network configuration, security policy, diagnostic tools, and operational models as 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 saving valuable time to focus on virtual machine administration.

https://t.me/learningnets

certcollecion.net

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



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), then apply the profiles to individual VM virtual NICs through the VMware vCenter GUI for transparent provisioning of network resources. Port profiles are a scalable mechanism to configure networks with large numbers of VMs.

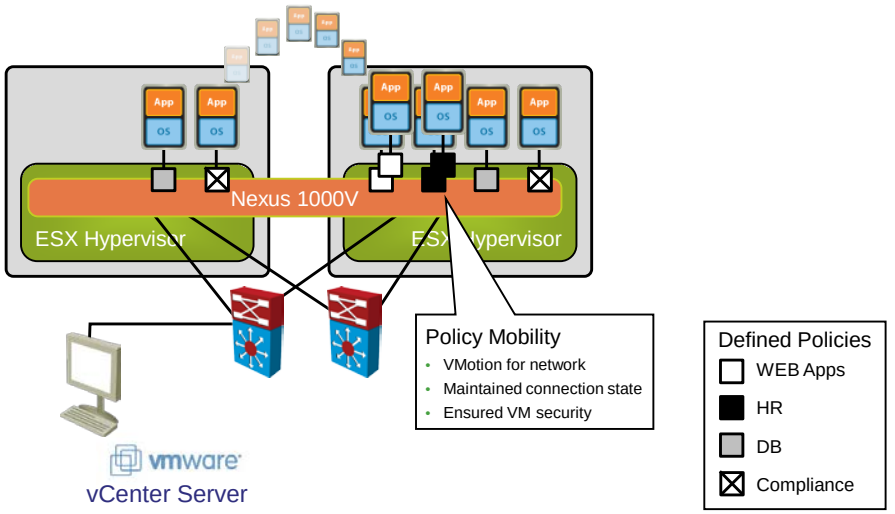
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

Cisco Nexus 1000V Distributed Virtual Switch (Cont.)

- Mobility of network and security properties



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

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

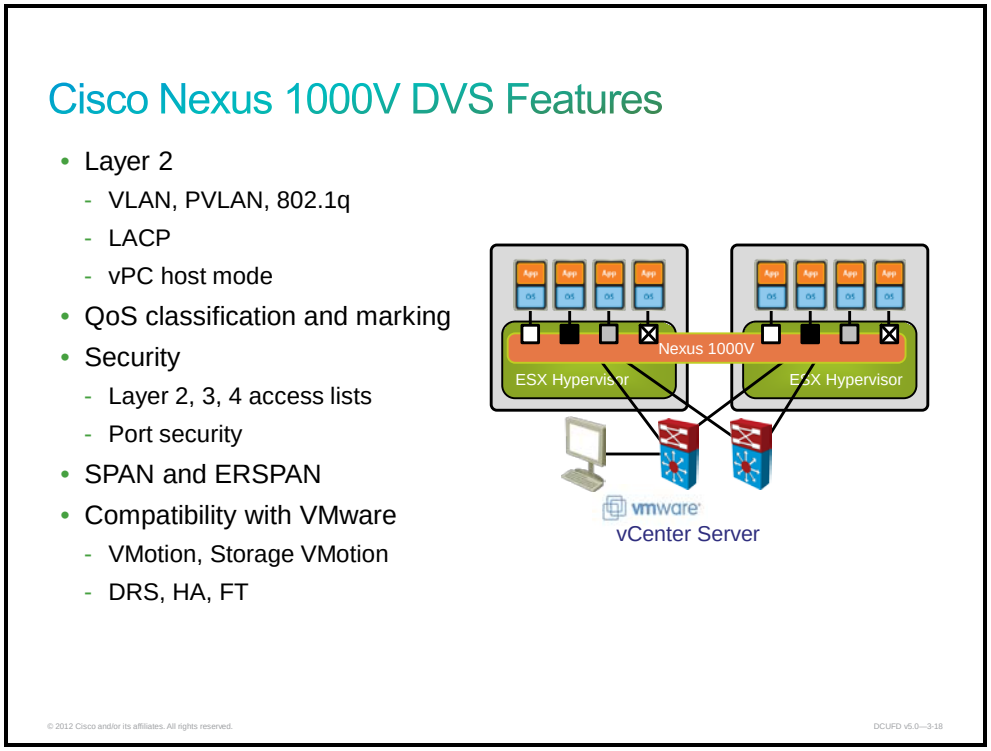
Network and security policies defined in the port profile follow the VM throughout its life cycle, whether it is being migrated from one server to another, suspended, hibernated, or restarted.

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

certcollecion.net



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

vPC Host Mode

Virtual port channel 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 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
- vPC-HM
- 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 (BPDU) filter without running Spanning Tree Protocol (STP)

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

QoS Features

The following QoS features are supported by Cisco Nexus 1000V:

- Classification per access group (by 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), 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 (MQC) compliance.

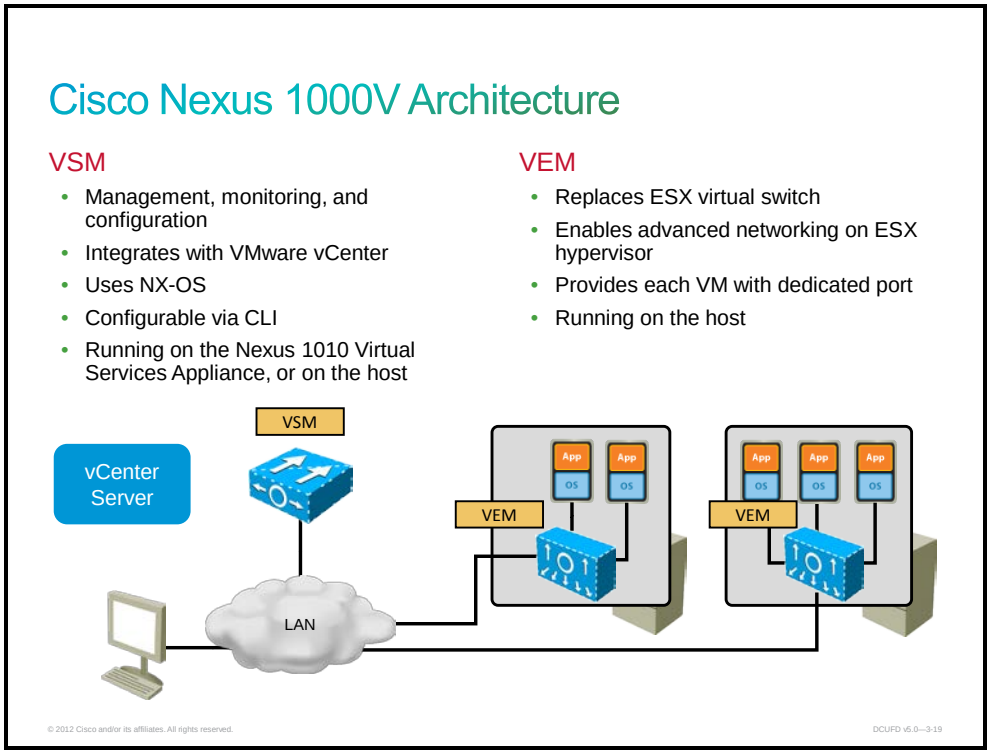
Security Features

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

- Ingress and egress ACLs on Ethernet and vEthernet ports
- Standard and extended Layer 2 ACLs
- Standard and extended Layer 3 and Layer 4 ACLs
- Port ACLs (PACLs)
- Named ACLs
- ACL statistics
- Cisco Integrated Security features
- Virtual service domain for Layer 4 through 7 virtual machine services

https://t.me/learningnets

certcollecion.net



Virtual Supervisor Module

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

- **Cisco Nexus 1000V Virtual Supervisor Module (VSM):** This module 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.
- **Cisco Nexus 1000V Virtual Ethernet Module (VEM):** This module **e**nables advanced networking capability on the VMware ESX hypervisor and provides each VM with a virtual dedicated switch port.

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

A VSM is the control plane on a supervisor module much like in regular physical modular switches, whereas VEMs are remote Ethernet line cards to the VSM.

In Cisco Nexus 1000V deployments, VMware provides the virtual network interface card (vNIC) and drivers while the Cisco Nexus 1000V provides the switching and management of switching.

Virtual Ethernet Module

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

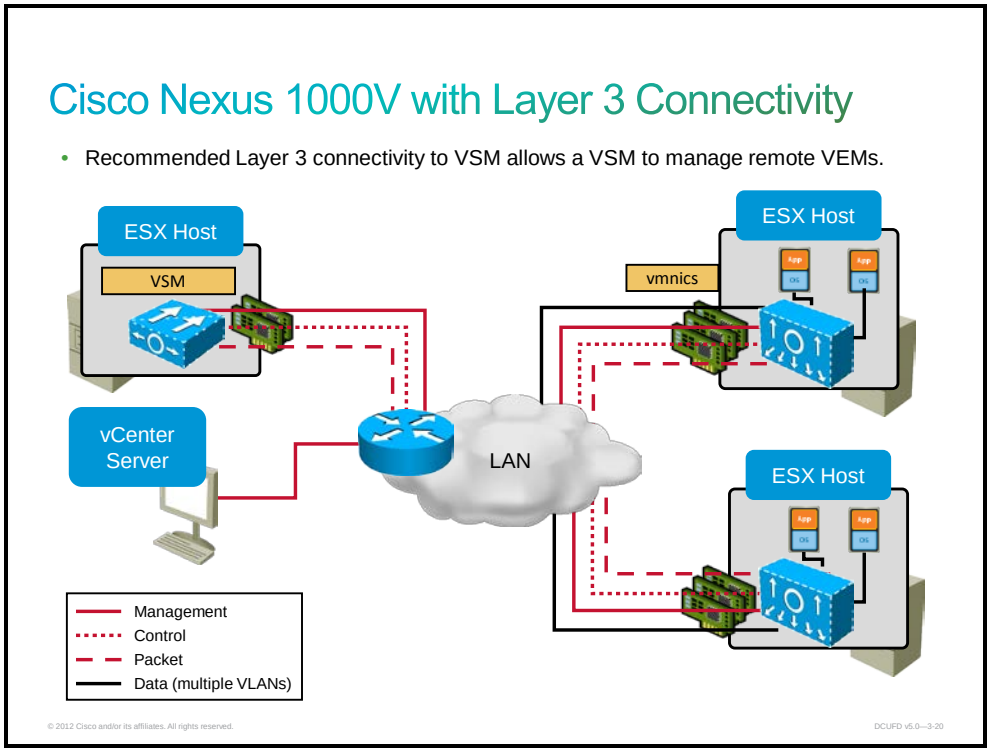
The VEM leverages the VMware vNetwork Distributed Switch (vDS) API, which was developed jointly by Cisco and VMware, to provide advanced networking management for virtual machines. This level of integration ensures that the Cisco Nexus 1000V is fully aware of all server virtualization events, such as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

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

- Port channels
- Port profiles
- Quality of service (QoS)
- Security: Private VLAN, access control lists, port security
- Monitoring: NetFlow, Switch Port Analyzer (SPAN), ERSPAN



Cisco Nexus 1000V VSM-VEM Connectivity Options

Layer 3 Connectivity

The Cisco Nexus 1000V VSM and VEM need to communicate in order to maintain the control plane of the switch (VSM) and to propagate the changes to the data plane (VEMs).

The VSM and the hosts need to be reachable over the IP protocol.

Note The minimum release of Cisco NX-OS for Cisco Nexus 1000V Release 4.0(4)SV1(2) is required for Layer 3 operation.

Deployment details for a Layer 3 VSM-VEM connection are the following:

- Layer 3 connectivity between VSM and VEMs:
  - VSM: Software virtual switch (SVS) Layer 3 mode with control or management interface
  - VEM: VMkernel interface and Generic Routing Encapsulation (GRE) to tunnel control traffic to VSM
  - Requires per-VEM Layer 3 control port profile

https://t.me/learningnets

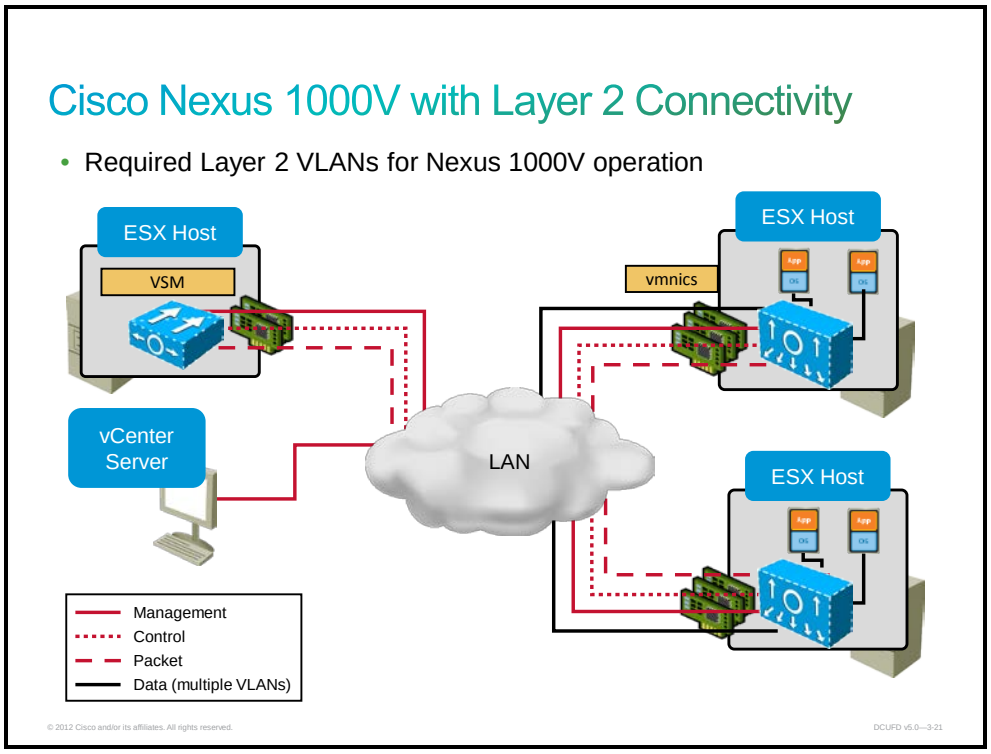
certcollecion.net

- Option 1: Management interface:
  - Out-of-band (OOB) management for VSM: mgmt0 port
  - Should be the same as VMware vCenter and ESX management VLAN
  - VSM-to-VEM traffic mixed with vCenter management traffic
- Option 2: Special control interface with own IP address:
  - Dedicated control0 interface for VSM-to-VEM communication

Note	VSM-VEM Layer 3 connectivity allows a VSM in a data center to manage VEMs in a remote data center. In such cases, the VSM in the primary data center is primary for local VEMs and secondary for remote VEMs.

Layer 2 Connectivity

The original option for Cisco Nexus 1000V VSM-VEM connectivity is using Layer 2.



Layer 2 connectivity (VLANs) is required between the VSM and VEMs:

- **Management VLAN, OOB for VSM (mgmt0 port):** Should be the same as vCenter and ESX management VLAN
- **Domain ID:** Single Cisco Nexus 1000V instance with dual VSM and VEMs
- **Control VLAN:** Exchanges control messages between the VSM and VEM
- **Packet VLAN:** Used for protocols like Cisco Discovery Protocol, LACP, and Internet Group Management Protocol (IGMP)
- **Data VLANs:** One or more VLANs are required for VM connectivity. It is recommended that separate VLANs are maintained.

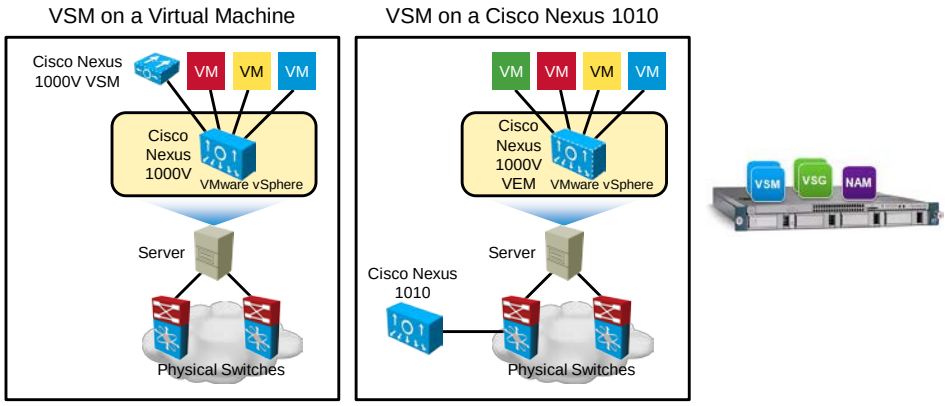
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

Cisco Nexus 1010 Virtual Services Appliance

- Hardware platform for Nexus 1000V VSM and other service appliances
- Provides VSM independence of existing production hosts
- Cisco Nexus 1010V comes bundled with licenses
- Platform for additional services: Cisco virtual NAM, VSG



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

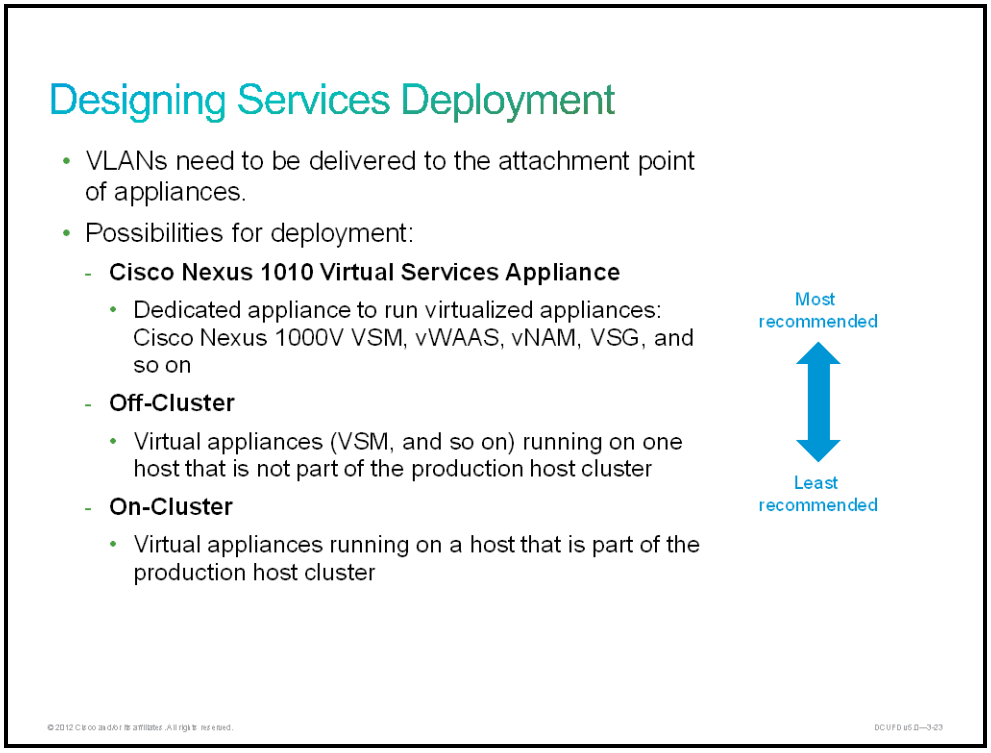
The Cisco Nexus 1010 Virtual Services Appliance is the hardware platform for the VSM. It is designed for customers who want to provide independence for the VSM so that it does not share the production infrastructure.

As an additional benefit, the Nexus 1010V comes bundled with VEM licenses.

The Nexus 1010V also serves as a hardware platform for various additional services, including the Cisco virtual Network Analysis Module (NAM), Cisco Virtual Security Gateway (VSG), and so on.

https://t.me/learningnets

certcollecion.net



These are the three possibilities for deploying virtualized services, listed from most to least recommended.

Deployment on Cisco Nexus 1010 Virtual Services Appliance

The Cisco Nexus 1010 Virtual Services Appliance is a dedicated appliance that runs virtualized appliances, such as the Cisco Nexus 1000V VSM, virtual Wide Area Application Services (vWAAS), vNAM, Virtual Security Gateway (VSG), and so on.

This option is the best because all appliances are running in a controlled and dedicated environment.

Off-Cluster Deployment

Virtual appliances run on one host that is not part of the production host cluster. Typically, a host system that is not used for production VMs is used, if it has enough CPU and memory resources.

This host is not part of the production cluster and is isolated from failures on the production environment, such as control software or cluster management software failure.

On-Cluster Deployment

In this case, the virtual appliances run on a host that is part of the production host cluster, together with the production VMs. This is the least desired option because the virtual appliances are not isolated from failures that may occur in the production network.

If a host in the production network goes offline, so do virtual appliances. Such a solution is meant to be temporary (for example, during the migration phase of VMs to new hardware, or when repurposing existing hosts).

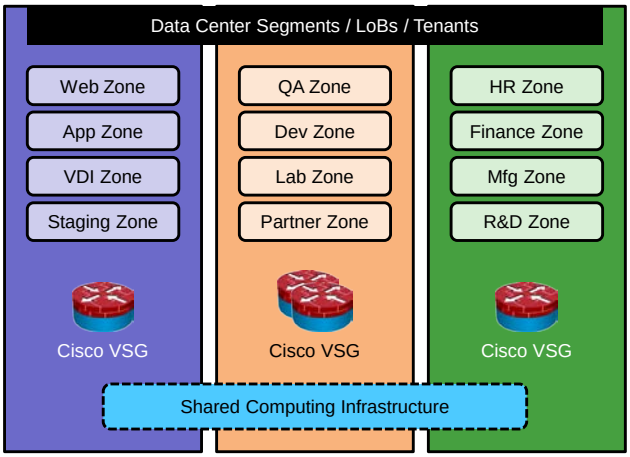
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

Cisco VSG

- Virtual network is divided to multiple security zones
- Security zones belong to network segments, departments, or tenants



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

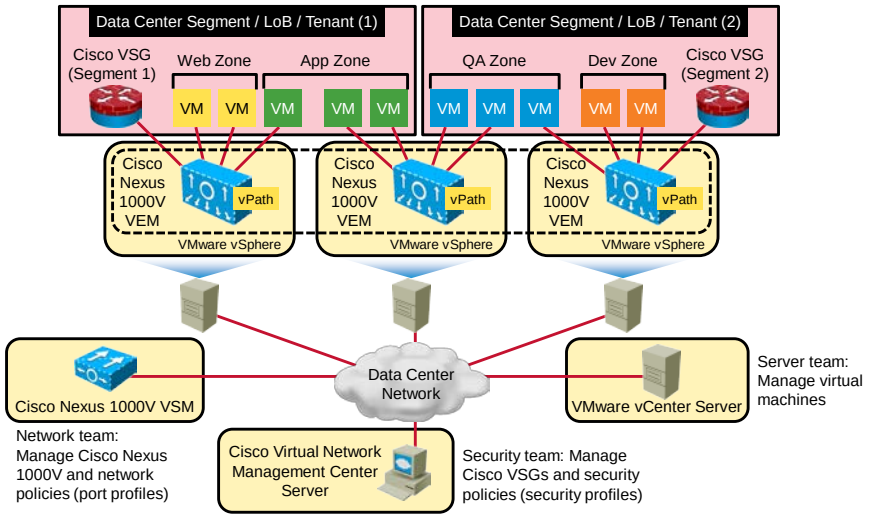
The Cisco VSG for Cisco Nexus 1000V Series Switches is a virtual appliance that provides trusted access to secure virtualized data centers in enterprise and cloud provider environments while meeting the requirements of dynamic policy-driven operations, mobility-transparent enforcement, and scale-out deployment for dense multitenancy.

https://t.me/learningnets

certcollecion.net

Cisco VSG (Cont.)

- Cisco VSG inspects an incoming flow, and if it is administratively permitted, allows the flow through the Nexus 1000V switch. A vPath is created.



Cisco vPath technology steers traffic, whether inbound or traveling from virtual machine to virtual machine, to the designated Cisco VSGs. A split-processing model is applied in which initial packet processing occurs in the Cisco VSG for policy evaluation and enforcement. Subsequent policy enforcement for packets is offloaded directly to vPath. Cisco vPath provides these advantages:

- **Intelligent traffic steering:** Flow classification and redirection to associated Cisco VSGs
- **Fast path offload:** Policy enforcement of flows offloaded by Cisco VSG to vPath

Cisco vPath is designed for multitenancy and provides traffic steering and fast path offload on a per-tenant basis.

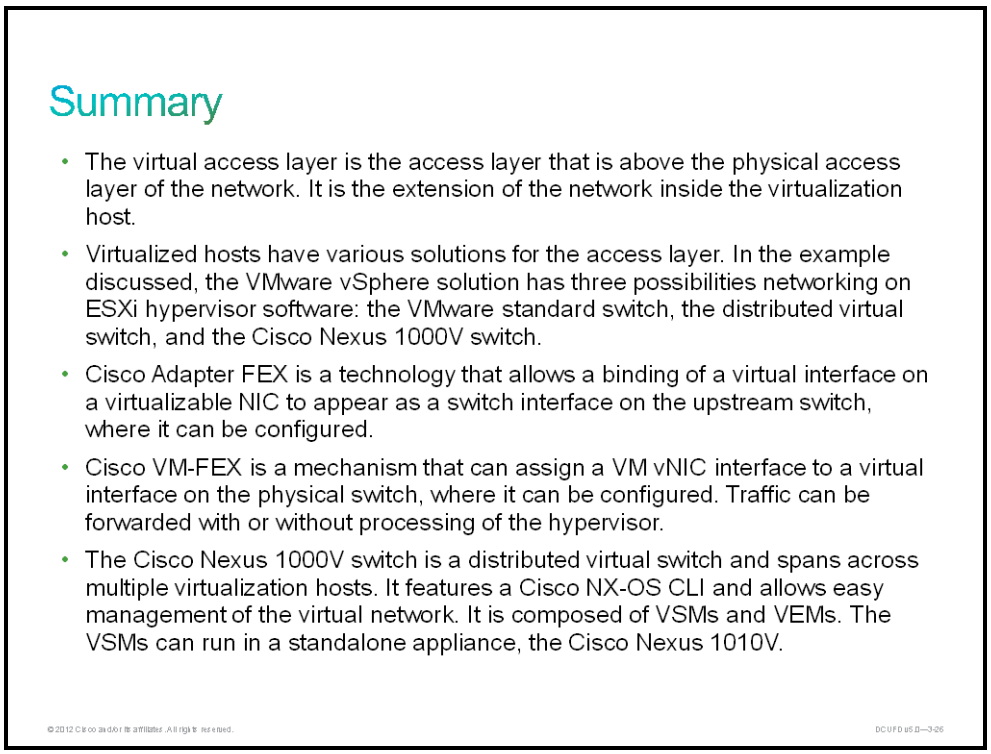
© 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

Summary

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



https://t.me/learningnets

certcollecion.net

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

Lesson 5

Designing High Availability

Overview

In this lesson, you will analyze various technologies that can provide high availability on Layer 3. These include IP routing protocols, first hop redundancy protocols, Locator Identity Separation Protocol (LISP), and clustered applications to some extent.

In addition to IP protocol high availability, there are other high-availability approaches. One variant is to provide high availability on the data link layer, where both virtual port channel (vPC) and Cisco FabricPath can be used. On the equipment level, there are options to provide high availability by employing redundant supervisor engines and similar technologies.

This lesson focuses on high availability provided by Layer 3 protocols.

Objectives

Upon completing this lesson, you will be able to design for data center high availability with various technologies, including IP routing, clusters, next-hop redundancy protocols, and LISP. This ability includes being able to meet these objectives:

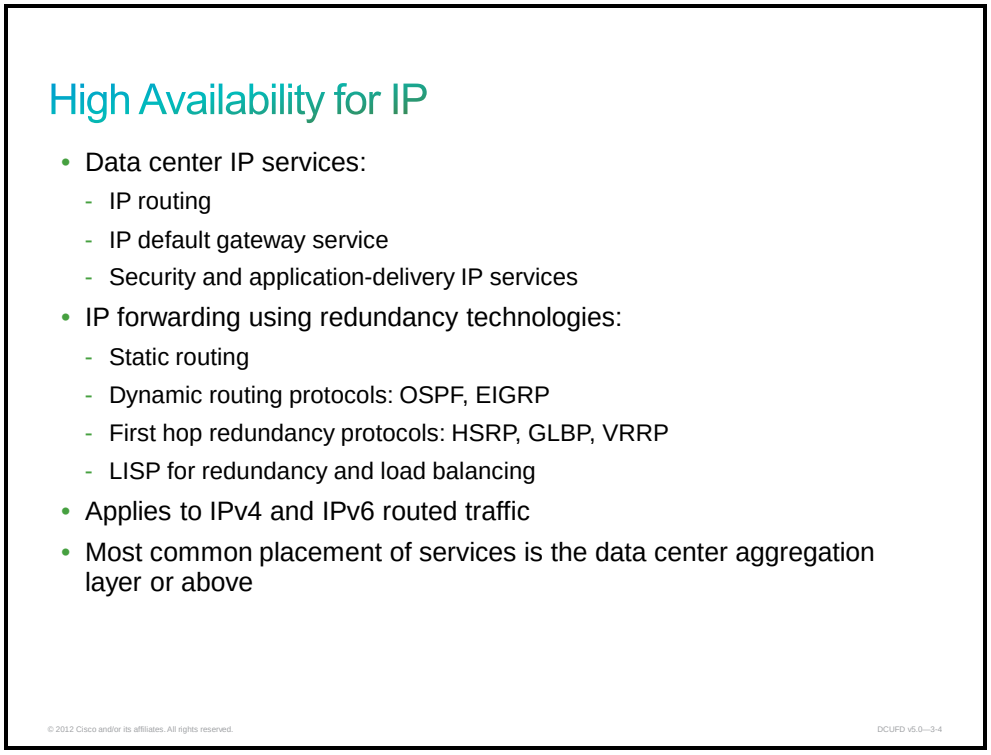
- Design high availability for IP-based services
- Design high availability by implementing link aggregation
- Design high availability of services using IP routing and FHRPs
- Provide high availability with RHI
- Design high availability of services using LISP

https://t.me/learningnets

certcollecion.net

High Availability for IP

This topic describes how to design high availability for IP-based services.



When designing highly available data centers, you need to provision the IP layer as well. The IP protocol is not highly available as such, so various enhancements and protocols are available to guarantee continuous operation.

The first set of protocols that makes IP highly available is the First Hop Redundancy Protocols (FHRPs): Hot Standby Router Protocol (HSRP), Virtual Router Redundancy Protocol (VRRP), and Gateway Load Balancing Protocol (GLBP).

The second set of protocols is IP routing protocols. These protocols provide path redundancy across different links. The most popular protocols used in data centers are Open Shortest Path First (OSPF) and Enhanced Interior Gateway Routing Protocol (EIGRP), with IBGP and Routing Information Protocol (RIP) for special applications.

The most common place to implement IP high availability is the data center aggregation layer.

In the last example, LISP is not directly a redundancy protocol, but can accommodate path selection and failover for the IP protocol. The primary use of LISP is to separate the IP endpoint identity information from the IP endpoint location information.

LISP facilitates a more robust high availability in situations where requirements go beyond a single data center. Server virtualization across geographically separated data centers requires location independence to allow for dynamically moving server resources from one data center to another. Dynamic workload requires route optimization when the virtual servers move while keeping the server IP address the same. LISP then enables IP endpoints to change location while keeping their assigned IP addresses.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

First Hop Redundancy Protocols

- Provide the IP default gateway service to devices in a subnet using these protocols:
  - HSRP
  - VRRP
  - GLBP
- Variants:
  - HSRP on Cisco Nexus 7000 Series Switches with vPC for access switches
- FHRP protocols are configurable and flexible:
  - Tuning possible for quicker switchover
  - Interface or object tracking to respond to topology changes: physical interface tracking, IP route tracking, IP reachability tracking, and so on

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

The FHRPs provide the default gateway service to devices in a subnet. The protocol is set up between (at least) two physical devices that otherwise host the default gateway IP address.

The use of FHRP may also depend on the amount of server-to-storage and server-to-server traffic, as well as where the storage is attached (aggregation or access layer). If the traffic does not remain local to the access switches, inter-VLAN traffic must be routed at the aggregation layer (Layer 3 FHRP gateways, because that is the demarcation between Layer 2 and Layer 3).

The most popular FHRP protocol is the HSRP. It has plenty of configurable options ranging from timer configuration to tracking objects, IPv6 support, and so on. This protocol is Cisco proprietary.

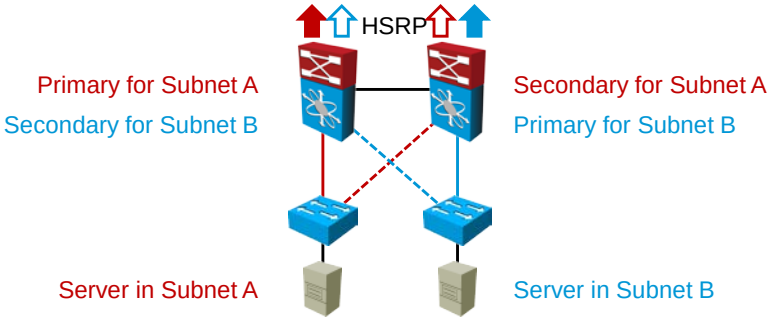
The VRRP protocol is open-standards-based and provides functionality that is similar to HSRP.

The GLBP protocol additionally provides load balancing between several default gateway devices. Servers in the same subnet can use multiple gateways to forward traffic upstream, utilizing all upstream links. However, traffic in the direction toward the servers usually travels only across one of these gateways.

https://t.me/learningnets

certcollecion.net

HSRP Within the Data Center



- Manual load balancing: primary default gateway on different aggregation switches for different subnets
- Only one default gateway. Two if combined with vPC.
- In case of failure, the surviving device takes 100 percent of the load.
- Same design applies when using vPC for the connection between access and aggregation layers
- Combine with tracking of upstream interfaces or routes from the core

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

The slide presents one of the most classic HSRP designs found in data center networks. HSRP is run on a pair of aggregation switches, where one is selected as the primary switch for selected VLANs, while the other is the primary default gateway for the remaining VLANs. Load balancing is achieved manually by equally distributing the VLANs among the switches.

Note To fully support this scenario, you need to use the Per-VLAN Spanning Tree (PVST) protocol or Multiple Spanning Tree (MST) protocol. The Spanning Tree Protocol (STP) primary root bridge and HSRP active router must be on the same device for the same network (VLAN). This setting allows the forwarding path to be aligned on both Layer 2 and Layer 3.

Within the same group, you have only one switch acting as the default gateway for servers in that subnet. In some cases, it is desirable to have multiple next-hop HSRP addresses that are active between different pairs of switches on the same subnet.

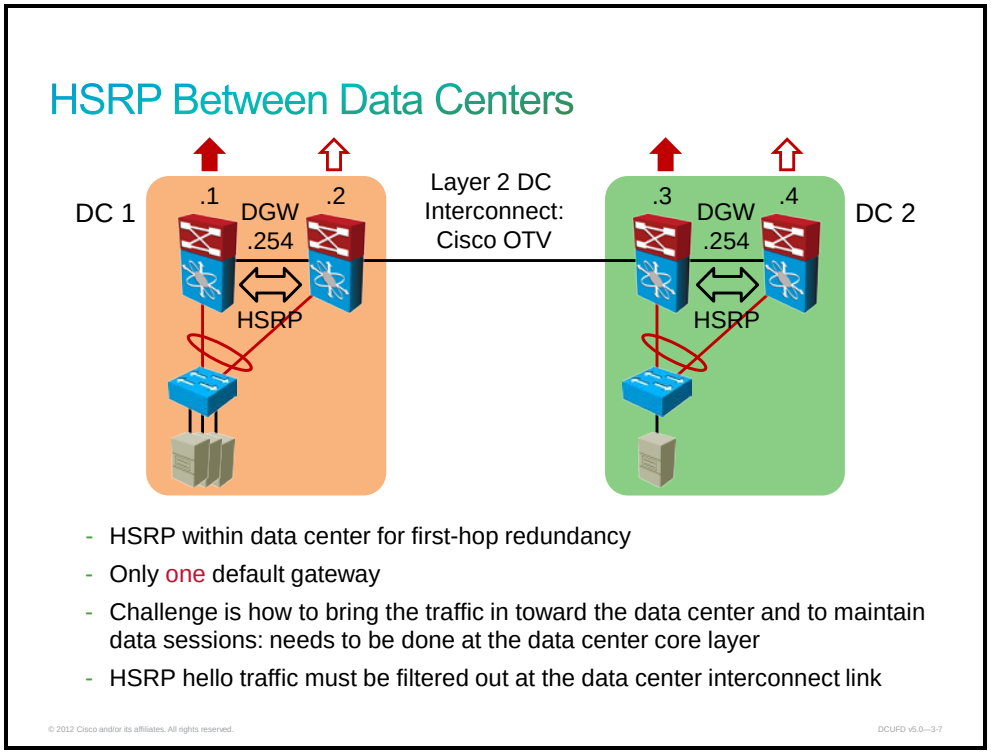
Interface tracking or object tracking can be used to have the active HSRP gateway running on the device that actually has functioning upstream connectivity. This is done by configuring object tracking that can monitor:

- Physical upstream interfaces
- IP routes
- General IP reachability, and so on

Note Given the prevalence of vPC, HSRP can leverage the capability to forward data on both the active and the standby switch. The data-plane optimization made by vPC allows Layer 3 forwarding at both the active HSRP peer and the standby HSRP peer. In effect, this provides an active-active FHRP behavior.
The aggregation switches need to have the same reachability for upstream routes for the HSRP secondary device to forward traffic.

https://t.me/learningnets

certcollecion.net



In this scenario, you have redundant data centers interconnected with Layer 2 transport. HSRP can run between the aggregation switches in both data centers. In the primary data center, there is the primary and secondary IP default gateway, while in the secondary data center there are local primary and local secondary IP default gateways.

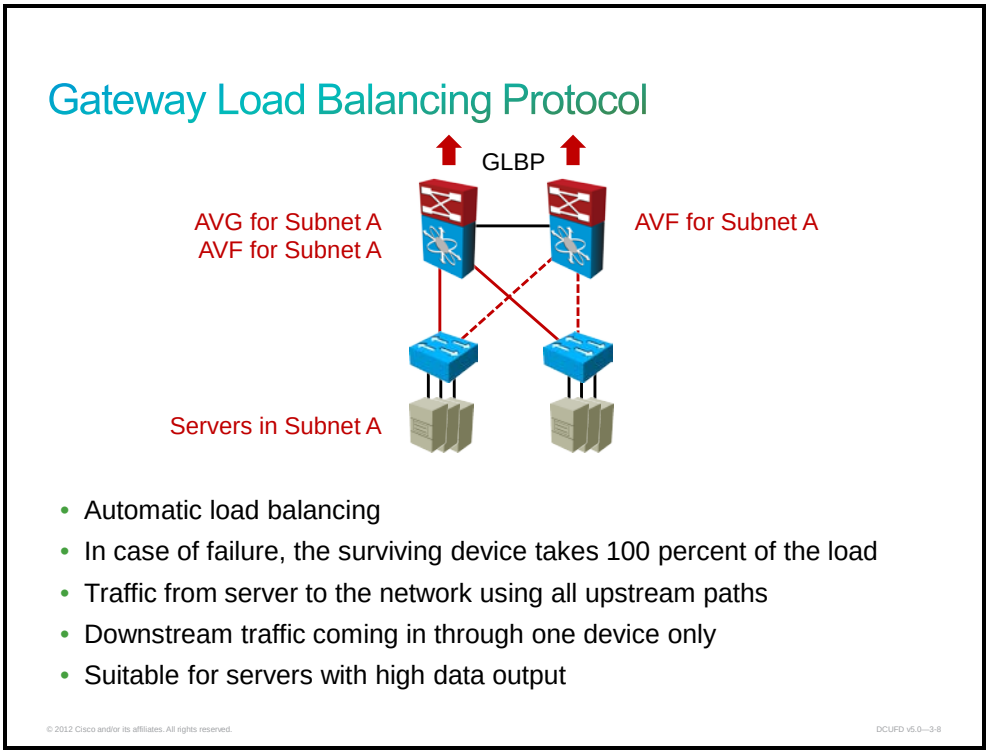
This approach uses the primary default gateway in the primary data center; servers in the secondary data center need to send the data across the data center interconnect.

This design is more suitable for active-standby load distribution, where most network traffic is concentrated in the primary data center.

Downstream traffic from the Internet to the servers goes only to the primary data center. If the primary site is down, IP routing must be adjusted to attract the traffic from the Internet to the secondary data center.

https://t.me/learningnets

certcollecion.net



The Cisco GLBP is another first-hop redundancy protocol that can be used in data center networks. The difference between GLBP and HSRP is that GLBP automatically provides for load balancing among multiple gateways.

GLBP distributes the load between gateways in the following way:

- When a host issues an Address Resolution Protocol (ARP) request for the MAC address for the configured default gateway IP address, the active virtual gateway (AVG) replies with an ARP reply and sends the MAC address of a chosen active virtual forwarder (AVF), which forwards traffic for that host.
- Different hosts receive different AVF MAC addresses.

If an AVF fails, another AVF assumes the MAC address of the failed AVF. Failure is detected by other gateways by lost hello packets.

GLBP is suitable for servers that produce much outgoing traffic. The return path may be asymmetrical at the last few hops, but this does not impose any problems.

Note Generally, the deployments of GLBP have a smaller footprint than HSRP. The most significant reason that GLBP does not have wider deployment is that it provides minimal value if you span VLANs between closet switches, for example, if there are any blocked uplinks from the access to the aggregation layer.
Because GLBP shares the load between two aggregation switches, it only makes sense to use GLBP if both uplinks are active for a specific VLAN. If both uplinks are not active, you send the traffic up to only one aggregation switch, and then forward data to the other aggregation switch across the interswitch link.
The other reason it is not used in more environments is that the Virtual Switching System (VSS) removes the need.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

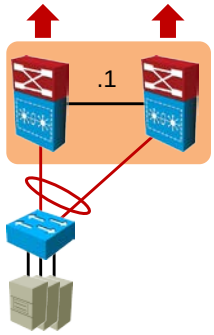
certcollecion.net

High Availability Using vPC and VSS

This topic describes how to design high availability by implementing link aggregation.

High Availability Using VSS

- No need for HSRP on VSS: default gateway redundancy coupled with control plane redundancy
- Upstream traffic exits either switch
- Downstream traffic enters either switch and prefers local path to server



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

The Cisco Catalyst 6500 Virtual Switching System (VSS) does not need an FHRP protocol because the default gateway IP address resides on the active chassis. The Cisco Catalyst 6500 VSS control plane manages gateway redundancy in this case. Because it is a single virtual device, an FHRP is not necessary.

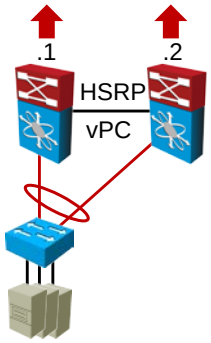
The Cisco Catalyst 6500 VSS also appears as one routing neighbor to all other devices. Typically it utilizes all available (Layer 2 or Layer 3) links at the same time.

https://t.me/learningnets

certcollecion.net

High Availability Using vPC

- Needs HSRP
- Secondary upstream Layer 3 forwarding with vPC if equal route found
- Downstream traffic through one device only unless one of these situations exists:
  - Double-sided vPC
  - Both devices advertise the same IP prefix for server subnet; switch will prefer local downstream path



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

The vPC scenario is different because you have two devices with two distinct control planes that are joined in the same vPC domain. Because the gateways remain two physical devices, an FHRP is necessary.

You need a VLAN trunk link between aggregation switches because it is necessary to transport HSRP hello packets between the switch virtual interfaces (SVIs) on both switches.

Normally, both aggregation switches will forward traffic upstream if both of them have the same route to the destination, with the same cost. This is done to optimize upstream network connectivity and to efficiently distribute the load between all links to the core layer.

Downstream traffic from the core to the servers will arrive on the primary aggregation switch unless Equal-Cost Multipath (ECMP) is used and IP routes for server subnets are advertised and tuned properly.

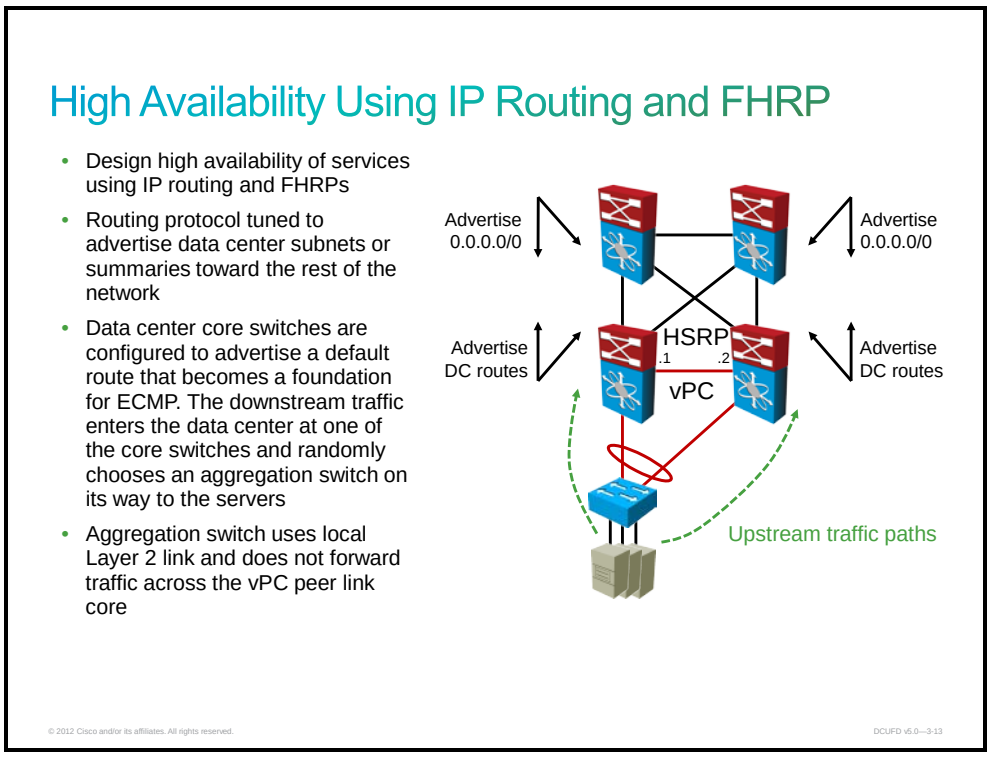
When IP routing is configured so that it also load-balances the links downstream (that is, when the aggregation switches advertise the server subnets with the same cost to all upstream switches), the packet destined for the server can arrive on any aggregation switch. In this case, the aggregation switch will receive the packet and use the local link to the access switch to forward traffic toward the server. It *will* avoid the vPC peer link to skip an unnecessary switched hop.

https://t.me/learningnets

certcollecion.net

High Availability Using IP Routing and FHRP

This topic describes how to design high availability of services using IP routing and FHRPs.



To design a fully redundant solution, you need to use a routing protocol in combination with an FHRP.

The routing protocol is tuned to advertise data center subnets and summary routes to the rest of the network, and data center core switches advertise default routes into the data center. This tuning allows core and aggregation switches to use ECMP for routed traffic.

When having core and aggregation switches connected with multiple links, all have same cost and are considered for distributing the load.

Upstream Traffic Flow

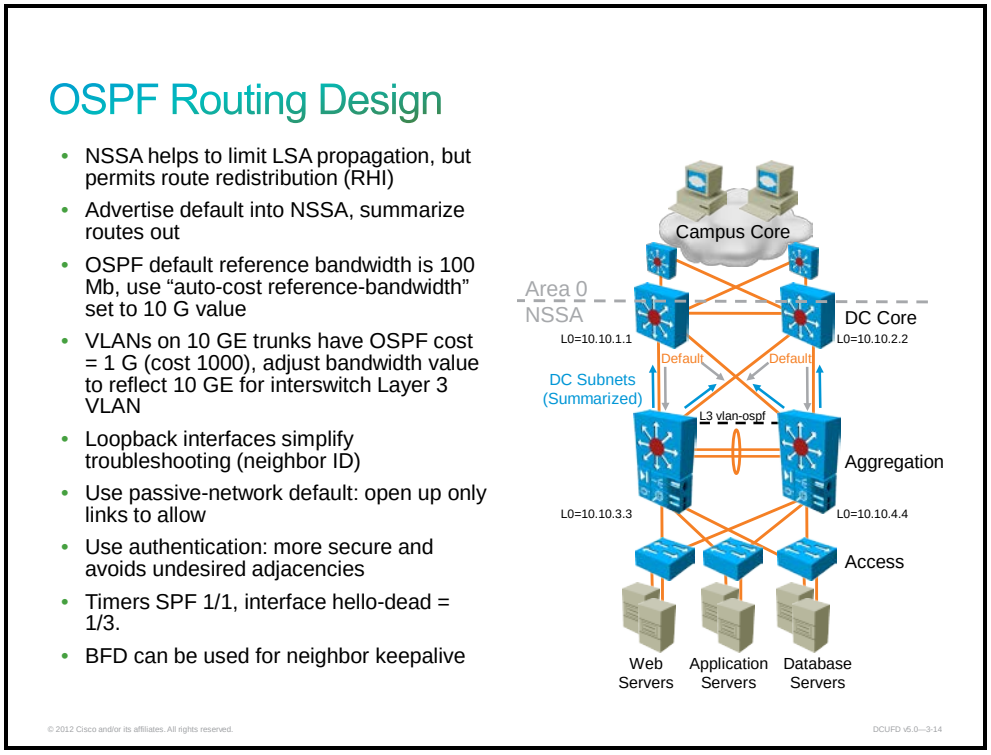
1. Traffic from the servers traveling upstream through the network is first forwarded on Layer 2.
2. The aggregation switch will forward the traffic upstream. Both aggregation switches will forward the packet upstream. The link between aggregation switches would not be used.
3. The core switches would forward the packets based on their routing table. If there are multiple paths to the same destination with the same cost, packets are forwarded using ECMP.

Downstream Traffic Flow

1. Traffic enters the data center core based on the routing information that the core advertises.
2. The core switch forwards the traffic to an aggregation switch using ECMP. There is no control at this point regarding which aggregation switch will get the traffic.
3. The aggregation switch uses the local link to send traffic in Layer 2 to the destination.

https://t.me/learningnets

certcollecion.net



IP Routing Protocols Deployment Design

OSPF Routing Protocol Design Recommendations

The data center aggregation and core layer can be an applied OSPF routing protocol, as shown in the figure:

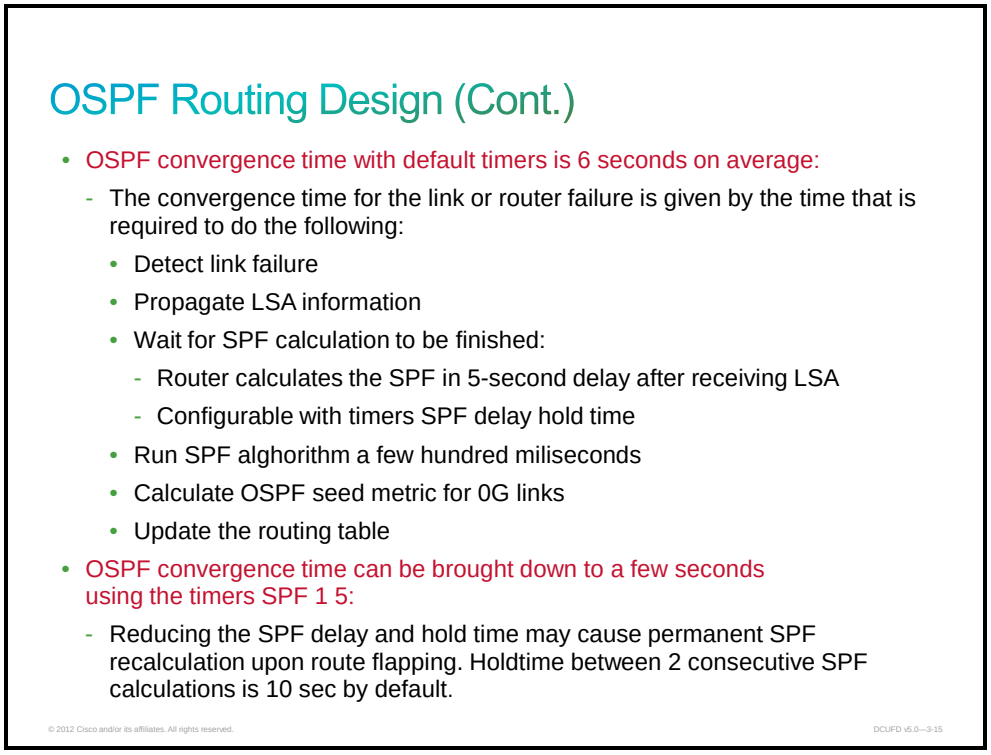
- The not-so-stubby area (NSSA) helps to limit link-state advertisement (LSA) propagation, but permits route redistribution if you use route health injection (RHI).
- OSPF advertises a default route into the NSSA area by itself, which makes the routing tables simple and easy to troubleshoot.
- For data center routes that are sent to the campus, summarization is advised.
- The OSPF default reference bandwidth is 100 Mb, so you need to use the **autocost reference-bandwidth** command with a value of 10 Gb, in order to distinguish between links that are faster than 100 Mb/s.
- Loopback interfaces simplify troubleshooting (OSPF router ID).
- Use the **passive-network default** command to prevent unwanted OSPF adjacencies and paths. Configure OSPF only on links for which you want to allow it.
- Establish a Layer 3 VLAN (SVI) on both switches that you use for OSPF adjacency and route exchange between them.
- Use routing protocol authentication. It is more secure and avoids undesired adjacencies.
- Reduce OSPF timers for faster convergence and neighbor loss detection.

Note	Because reduced OSPF timers impose an additional CPU load, you can use Bidirectional Forwarding Detection (BFD) instead to detect the presence of the neighbor, and link OSPF to the BFD instance.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net



OSPF convergence time with default timers is 6 seconds for an average topology. These steps can aid in reducing OSPF convergence time:

- Reduce the Shortest Path First (SPF) delay hold time.
- Enable incremental SPF to avoid recomputing the entire SPF.
- Run the SPF algorithm for a few hundred milliseconds.

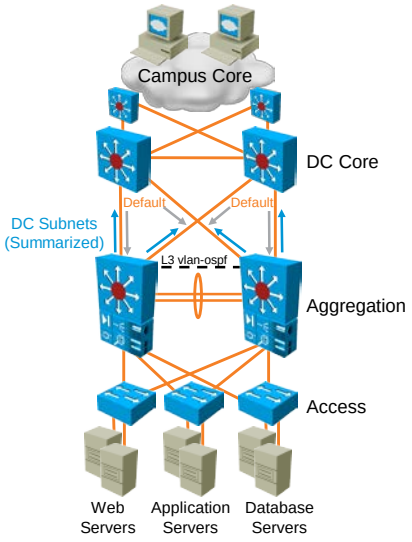
Note	The use of SPF timers is recommended only in an environment that uses a well-structured area, route summarization design, and link flap damping features.

https://t.me/learningnets

certcollecion.net

EIGRP Routing Design

- Advertise default into data center with interface command on core:
  - **ip summary-address eigrp 20 0.0.0.0 0.0.0.0 200**
  - Cost of 200 required to be preferred route over the NULL0 route installed by EIGRP
- If other default routes exist (from Internet edge, for example), may need to use distribute lists to filter out
- Use **passive-network** default
- Summarize data center subnets to core with interface command on aggregation switch:
  - **ip summary-address eigrp 20 10.20.0.0 255.255.0.0**



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

EIGRP Routing Protocol Design Recommendations

The data center aggregation and core layers can be applied EIGRP, as shown in the figure:

- You need to advertise the default into the data center with the interface command at the core:

  ip summary-address eigrp 20 0.0.0.0 0.0.0.0 200

  The cost (200) is required to make this route preferred over the Null0 summary route that is installed by EIGRP.
- If other default routes exist (from the Internet edge, for example), you may need to use distribute lists to filter them out.
- Use the **passive-network** default to prevent EIGRP from forming unwanted adjacencies.
- Summarize data center subnets to the core with the interface command on the aggregation switch:

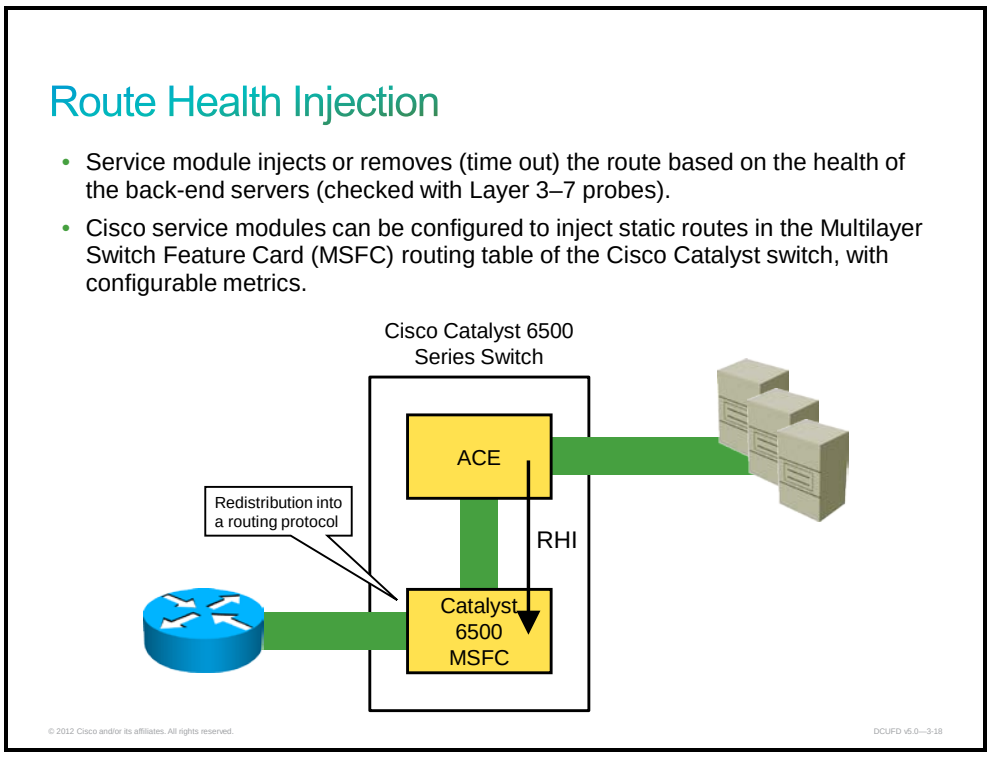
  ip summary-address eigrp 20 10.20.0.0 255.255.0.0

https://t.me/learningnets

certcollecion.net

High Availability Using RHI

This topic describes how to provide high availability with RHI.



RHI is a mechanism that can be used to advertise availability of a service into the routing domain.

The RHI feature is used to advertise a static host route to a particular server or service throughout the network. RHI enables active and standby services and anycast. The network finds the best way to reach a certain service, and the service must be uniform across multiple servers and across network sites.

The RHI feature gives the Cisco Application Control Engine (ACE) module the capability to inject static host routes into the routing table in the base Cisco IOS Software on the Cisco Catalyst 6500 Series chassis. These advertisements are sent out-of-band from the Cisco ACE Module directly to the Catalyst 6500 Series Switch supervisor. The Cisco IOS Software image on the supervisor takes the information from the RHI advertisement and creates a static route in its routing table. Both the Cisco ACE Module and the Cisco IOS Software are VRF-aware, and the routes advertised by RHI can therefore be put into appropriate VRF routing tables.

Note	Other platforms that do not have integrated service modules (for example, Cisco Nexus platforms) do not support RHI.

OSPF and RHI

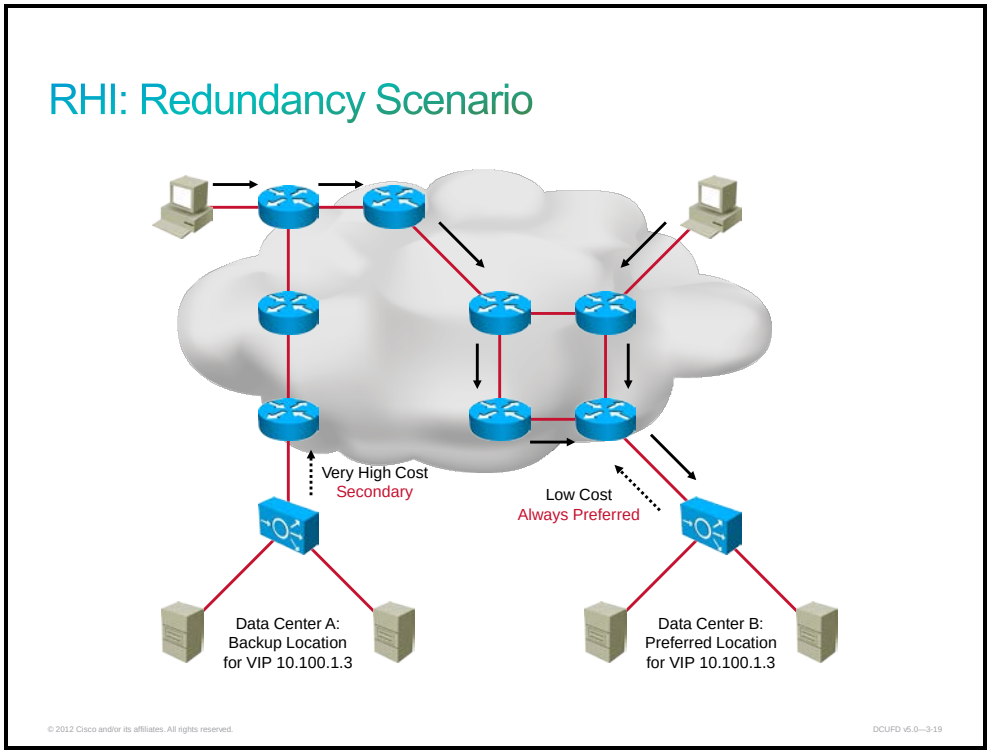
In the example involving the Cisco Catalyst 6500 Series Switch, service modules can be used to inject a static host route for a particular server or service in the routing table, and remove it if the service becomes unavailable.

This static route is then redistributed into the routing protocol and advertised to the rest of the network.

https://t.me/learningnets

certcollecion.net

Because it is a redistributed route, it will appear as an OSPF external route in the OSPF routing domain. If the data center routing domain is configured as an OSPF stub area, it cannot originate external routes. The solution to this problem is the OSPF NSSA area.



RHI can be used to provide additional redundancy for the service.

One of the ways to offer stateless services in a scalable way is anycast. Anycast can be used in global enterprise networks, where servers always answer the same type of requests with the same answers.

In this example, anycast is implemented in such way that you advertise a host route to the same IP address in multiple points of the network. Traffic to that destination IP address is then routed to the closest server.

Keep in mind that these scenarios are possible in enterprise data center networks. You must be able to advertise the same routing prefix at multiple sites (10.100.1.3/32 in this example), which typically is not allowed in the Internet.

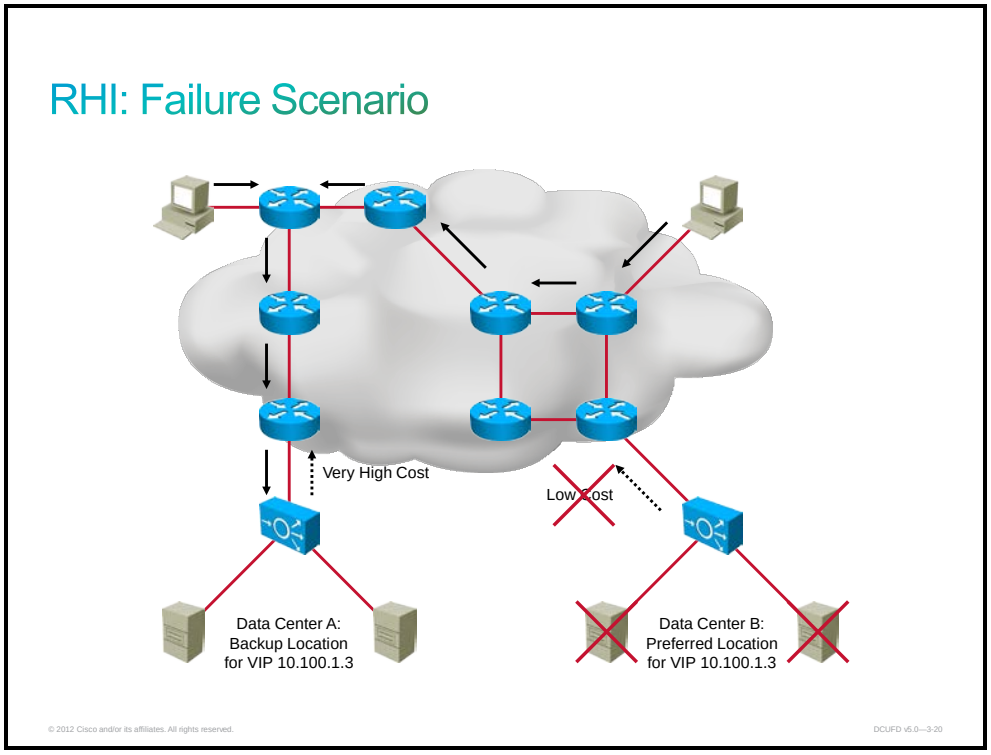
Note	An example of an anycast service deployed in the global Internet is the root Domain Name System (DNS) servers where public IP addresses are used. Root DNS servers are an allowed exception in the Internet.

Anycast can be used both for IPv4 and IPv6.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net



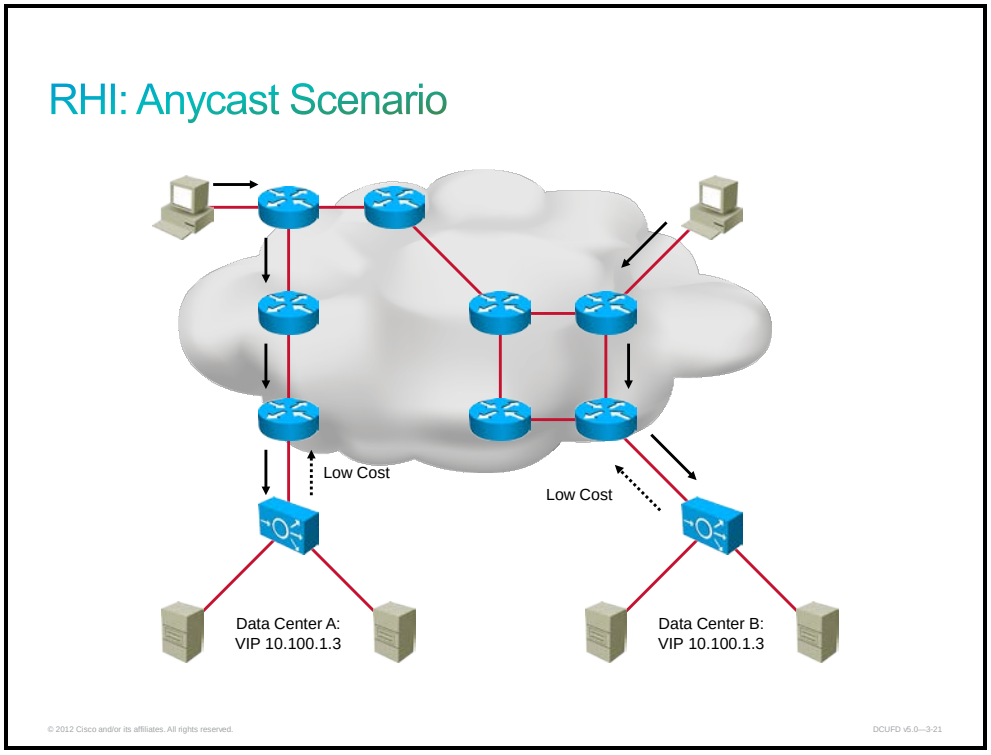
If the service fails, the host route is not advertised at that point in the network anymore, and a path to another host route (advertising the same IP address) is chosen. Service continuity is guaranteed by the routing protocol.

The figure shows how RHI manages a failure of the primary site. Clients are able to reach the backup virtual IP at Data Center A as soon as the routing protocols involved in the network converge. In general, this convergence happens very quickly.

Note	Connections that were in progress to Data Center B when it failed are lost. However, Data Center A accepts new connections very quickly.

https://t.me/learningnets

certcollecion.net



When RHI is used with the same costs, true anycast service is offered to the clients.

The load balancer injects the route to a particular server IP address in the routing table. This route is redistributed into the routing protocol and advertised at multiple sites.

RHI is also used to provide load-balanced service based on the proximity of the client to one of the server farms. The figure shows both locations advertising the same virtual IP via RHI. Routing functions in the network direct client requests to the server farm that is closest to the client. If either location fails, the routing protocols in the network quickly converge and the remaining location receives all client requests.

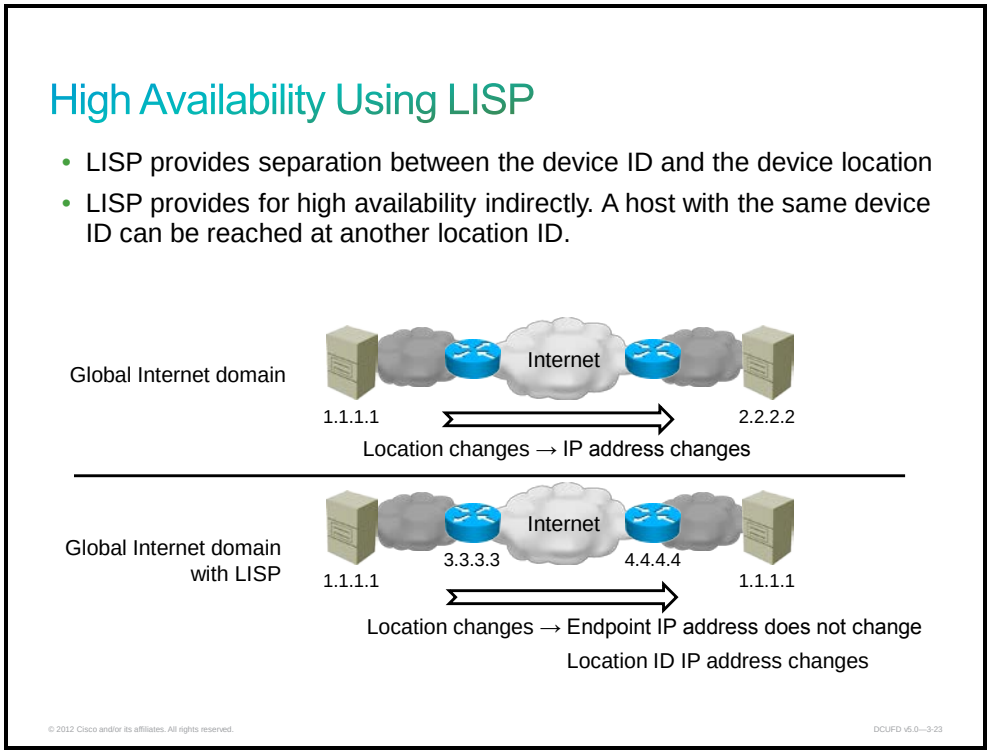
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

High Availability Using LISP

This topic describes how to design high availability of services using LISP.



LISP brings high availability indirectly. It is primarily a protocol that can provide separation between the *identity* and *location* for both IPv4 and IPv6.

Note	The server virtualization solution provides high availability. LISP helps to make it transparent.

LISP brings a whole new concept in IP routing that enables enterprises and service providers to simplify multihoming, facilitate scalable any-to-any WAN connectivity, support data center virtual machine mobility, and reduce operation complexities.

LISP implements a new semantic for IP addressing by creating two new namespaces:

- Endpoint identifiers (EIDs), which are assigned to end hosts
- Routing locators (RLOCs), which are assigned to devices (primarily routers) that make up the global routing system

In the current Internet routing and addressing architecture, the IP address is used as a single namespace that simultaneously expresses two functions about a device: its identity and how it is attached to the network. LISP uses a map-and-encapsulate routing model in which traffic destined for an EID is encapsulated and sent to an authoritative RLOC, rather than directly to the destination EID, based on the results of a lookup in a mapping database.

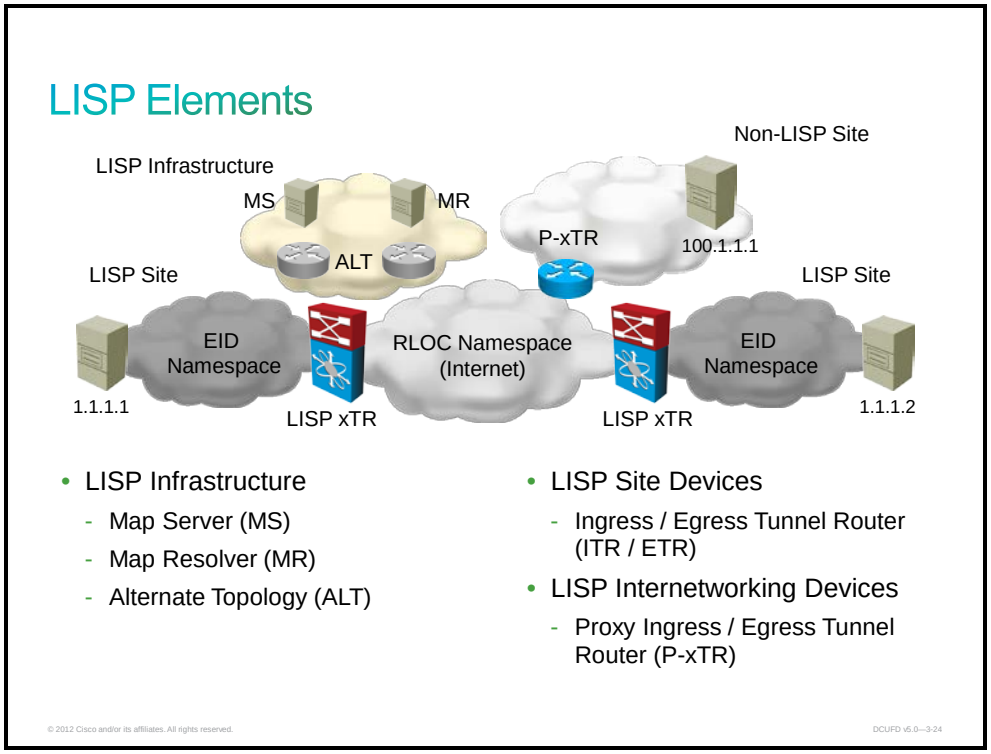
Services enabled by using LISP include the following:

- IP mobility with LISP for virtual machine mobility (Cisco LISP virtual machine [VM] mobility)
- IPv6 enablement

https://t.me/learningnets

certcollecion.net

- Multitenancy and large-scale VPNs
- Prefix portability and multihoming



The LISP site devices are as follows:

- **Ingress Tunnel Router (ITR):** This device is deployed as a LISP site edge device. It receives packets from site-facing interfaces (internal hosts). The ITR LISP encapsulates packets to remote LISP sites or the natively forwards packets to non-LISP sites.
- **Egress Tunnel Router (ETR):** This device is deployed as a LISP site edge device. It receives packets from core-facing interfaces (the Internet). The ETR de-encapsulates LISP packets or delivers them to local EIDs at the site.

Note Customer edge (CE) devices can implement both ITR and ETR functions. This type of CE device is referred to as a PxTR.

These are the LISP internetworking devices:

- **Proxy ITR (P-ITR):** This device is a LISP infrastructure device that provides connectivity between non-LISP sites and LISP sites. A P-ITR advertises coarse-aggregate prefixes for the LISP EID namespace into the Internet, which attracts non-LISP traffic that is destined to LISP sites. The PITR then encapsulates and forwards this traffic to LISP sites. This process not only facilitates internetworking between LISP and non-LISP sites, but also allows LISP sites to see LISP ingress traffic engineering benefits from non-LISP traffic.

Note The best location for an ITR or PITR is in the service provider environment.

- **Proxy ETR (P-ETR):** This device is a LISP infrastructure device that allows IPv6 LISP sites without native IPv6 RLOC connectivity to reach LISP sites that only have IPv6 RLOC connectivity. In addition, the P-ETR can also be used to allow LISP sites with Unicast Reverse Path Forwarding (uRPF) restrictions to reach non-LISP sites.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

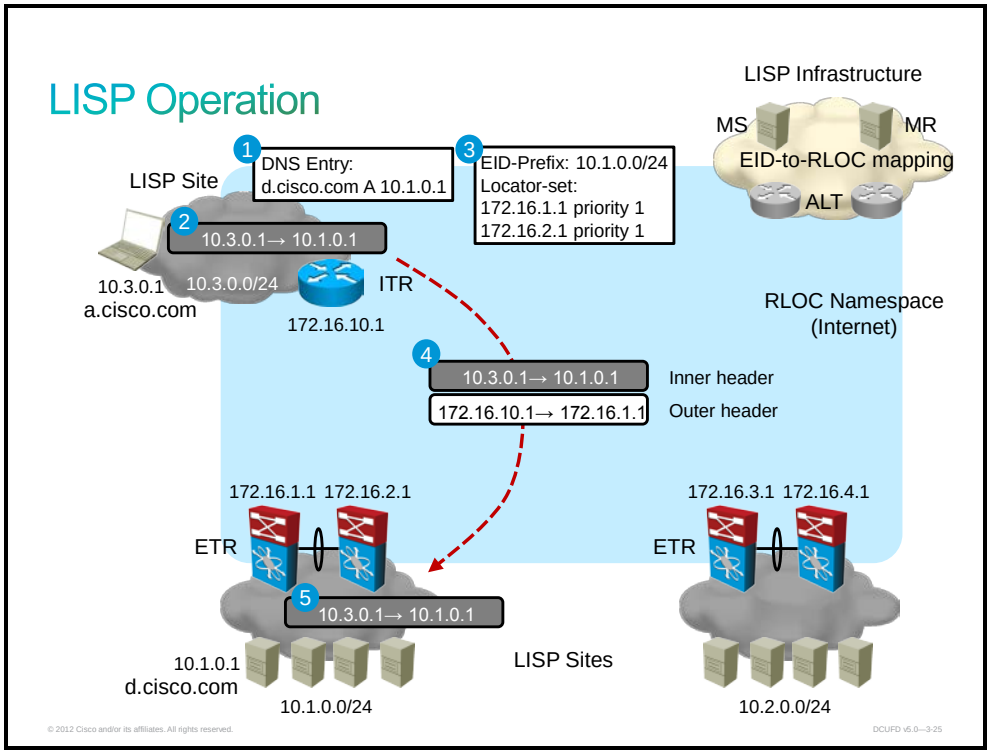
Note CE devices can implement both P-ITR and P-ETR functions. This type of CE device is referred to as an xTR.

These are the LISP infrastructure devices:

- **Map Server (MS):** This device is deployed as a LISP infrastructure component. It must be configured to permit a LISP site to register to it by specifying for each LISP site the EID prefixes for which registering ETRs are authoritative. An authentication key must match the key that is configured on the ETR. An MS receives Map-Register control packets from ETRs. When the MS is configured with a service interface to the LISP Alternate Topology (ALT), it injects aggregates for the EID prefixes for registered ETRs into the ALT. The MS also receives Map-Request control packets from the ALT, which it then encapsulates to the registered ETR that is authoritative for the EID prefix that is being queried.
- **Map Resolver (MR):** This device is deployed as a LISP infrastructure device. It receives encapsulated Map-Requests from ITRs. When configured with a service interface to the LISP ALT, it forwards Map-Requests to the ALT. The MR also sends Negative Map-Replies to ITRs in response to queries for non-LISP addresses.
- **Alternative Topology (ALT):** This is a logical topology and is deployed as part of the LISP infrastructure to provide scalable EID prefix aggregation. Because the ALT is deployed as a dual-stack (IPv4 and IPv6) Border Gateway Protocol (BGP) over Generic Routing Encapsulation (GRE) tunnels, you can use ALT-only devices with basic router hardware or other off-the-shelf devices that can support BGP and GRE.

https://t.me/learningnets

certcollecion.net



The figure describes the flow of a LISP packet:

1. The source endpoint performs a DNS lookup to find the destination: hostname *d.cisco.com*. The DNS server within the LISP domain replies with the IP address: 10.1.0.1.

2. Traffic is remote, so traffic is sent to the branch router.

   The IP packet has 10.3.0.1 as the source address, and 10.1.0.1 as the destination address.

3. The branch router does not know how to get to the specific address of the destination, but it is LISP-enabled so it performs a LISP lookup to find a locator address. The LISP mapping database informs the branch router how to get to the one (or more) available addresses that can get it to the destination. The LISP mapping database can return priority and weight as part of this lookup, to help with traffic engineering and shaping.

4. The branch router performs an IP-in-IP encapsulation and transmits the data out of the appropriate interface based on standard IP routing decisions.

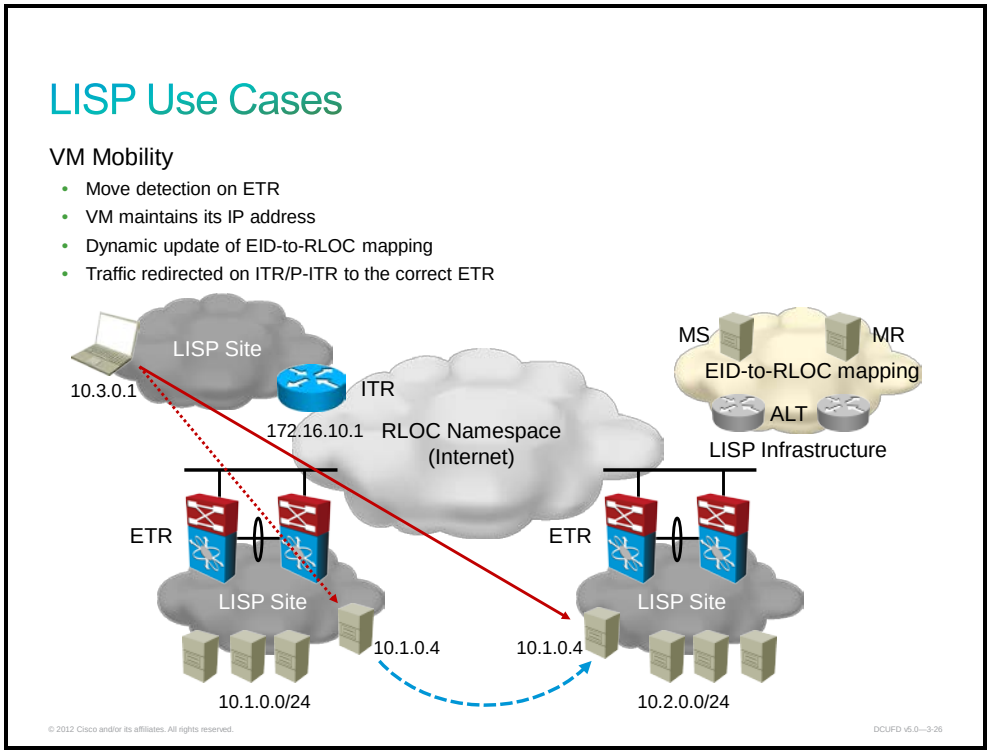
   The 10.1.0.1 host is located behind the 172.16.1.1 RLOC, so the packet is encapsulated into another IP header, with the 172.16.10.1 source address (ITR), and 172.16.1.1 destination address (ETR).

5. The receiving LISP-enabled router receives the packet, removes the encapsulation, and forwards the packet to the final destination (10.1.0.1).

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net



There are a couple of LISP use cases that are relevant for data centers. By using the LISP routing infrastructure, you can deploy virtual machine mobility between data centers, multiple virtualized infrastructures, and so on.

VM Mobility with LISP

The first use of LISP is to support virtual machine mobility.

The scenario involves providing mobility of virtual machines between data centers. Moving virtual machines between data centers is done over a dedicated Layer 2 network (VLAN).

Note	When a VM is moved, its access to its storage (disks) is also managed by the system so that the VM accesses the storage locally.

Another challenge is how to manage incoming [production] data flows from the network and how to preserve existing open data sessions. The VM is, after all, running in another data center, in another network, and you must route IP traffic correctly to reach the VM at its new location.

With LISP, when the VM changes its location, its RLOC changes, but not its EID. LISP updates the Map Servers and indicates that the EID of that particular virtual machine has a new RLOC (location).

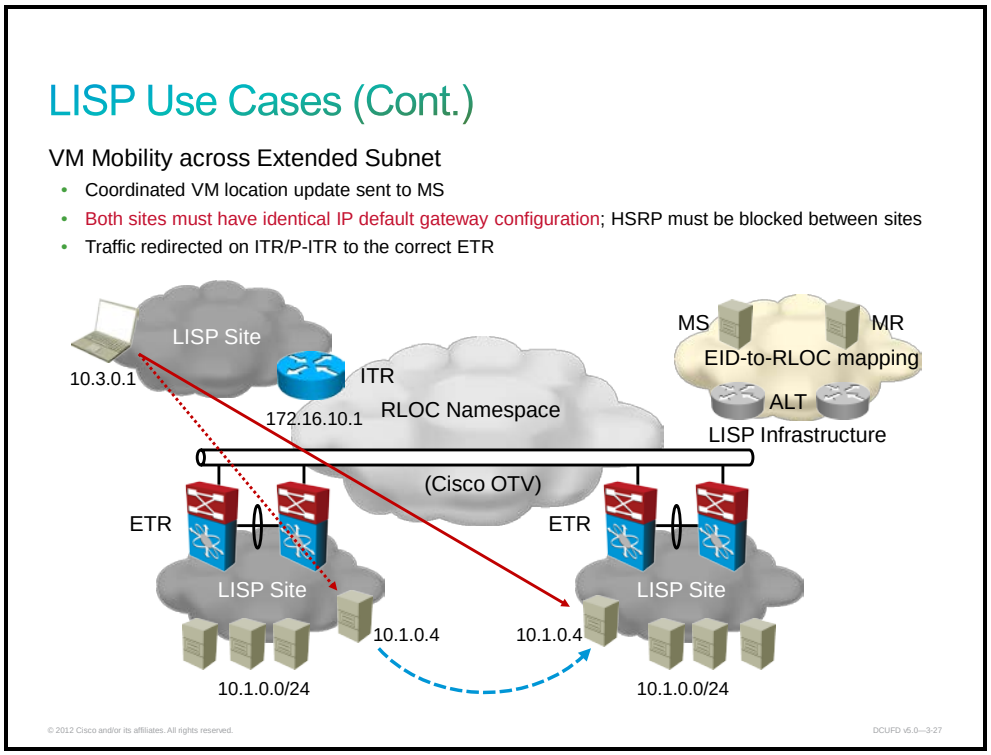
The LISP Tunnel Router (xTR) dynamically detects VM moves based on data plane events. LISP VM mobility compares the source IP address of host traffic received at the LISP router against a range of prefixes that are allowed to roam. IP prefixes of roaming devices within the range of allowed prefixes are referred to as the dynamic EIDs. When a new xTR detects a move, it updates the mappings between EIDs and RLOCs, which redirects traffic to the new locations without causing any disruption to the underlying routing. When deployed at the first hop router, LISP VM, mobility provides adaptable and comprehensive first hop router functionality to service the IP gateway needs of the roaming devices that relocate.

https://t.me/learningnets

certcollecion.net

Note	The LISP ETR implementation on the Cisco Nexus 7000 monitors the MAC addresses in the local subnets and is able to detect that a MAC address of a VM that has moved is not available at this location anymore. The Cisco Nexus Operating System (NX-OS) software resynchronizes with the MS/MR servers so that they are aware that a VM is available behind another ETR.

The figure shows LISP VM mobility across subnets between two enterprise-class data centers. In this case, two different subnets exist, one in each data center, and subnet and VLAN extension techniques such as Cisco Overlay Transport Virtualization (OTV) and Virtual Private LAN Services (VPLS) are not deployed. This mode can be used when an enterprise IT department needs to quickly start disaster recovery facilities when the network is not provisioned for virtual server subnets, or in case of cloud bursting, relocate EIDs across organization boundaries.



The figure shows LISP VM mobility in an extended subnet between two enterprise-class data centers. The subnets and VLANs are extended from the West data center (West DC) to the East data center (East DC) using Cisco OTV or VPLS, or any other LAN extension technology.

In traditional routing, this approach poses the challenge of ingress path optimization. LISP VM mobility provides transparent ingress path optimization by detecting the mobile EIDs (virtual servers) dynamically, and it updates the LISP mapping system with its current EID- RLOC mapping, which allows the virtual servers to be mobile between the data centers with ingress path optimization.

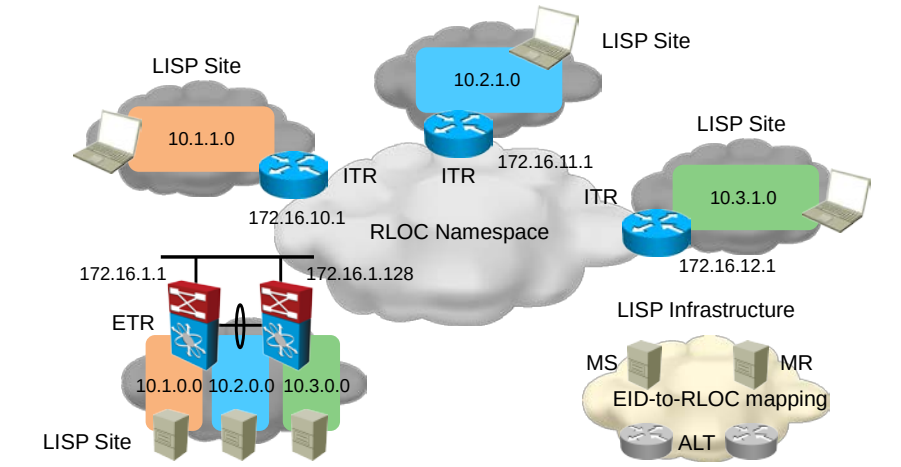
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

LISP Use Cases (Cont.)

Multitenant Environments



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

As a map-and-encapsulate mechanism, LISP is well suited to manage multiple virtual parallel address spaces. LISP mappings can be “color coded” to give VPN and tenant semantics to each prefix managed by LISP.

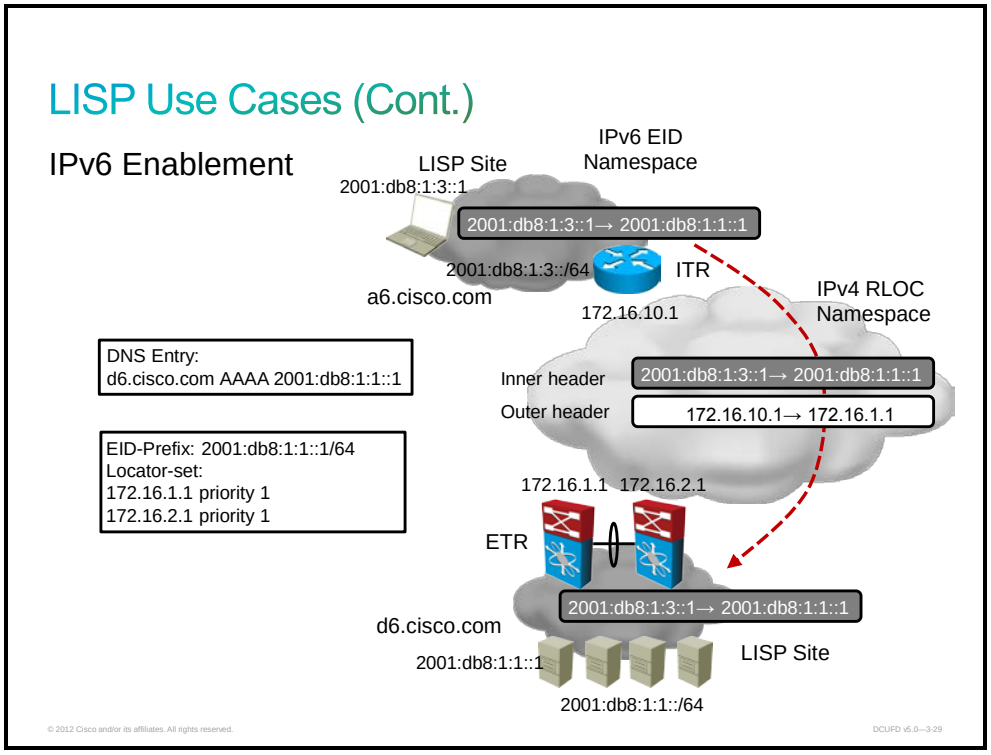
Color coding is encoded in the LISP control plane as stipulated in the standard definition of the protocol, and the LISP data plane has the necessary fields to support the segmentation of traffic into multiple VPNs. The LISP multitenancy solution is particularly attractive because it is natively integrated with the mobility, scalability, and IPv6 enablement functions that LISP offers, allowing all the various services to be enabled with the deployment of a single protocol.

The LISP multitenancy solution is not constrained by organizational boundaries, allowing users to deploy VPNs that can cut across multiple organizations to effectively reach any location and extend the network segmentation ubiquitously.

Virtual routing and forwarding (VRF) instances are used as containers to cache mapping entries and to provide transparent interoperability between the LISP segmentation solution and more traditional VRF interconnection techniques such as Multiprotocol Label Switching (MPLS) VPNs, VRF-Lite, and Easy Virtual Network (EVN).

https://t.me/learningnets

certcollecion.net



LISP can be used as a technology to extend your IPv6 “islands” across a commodity IPv4 network. The EID namespace runs IPv6, while the RLOC namespace runs IPv4. IPv6 traffic is encapsulated in IPv4 packets that can travel across the IPv4 Internet.

At the same time, LISP can be used to provide other benefits, such as VM mobility. LISP provides seamless connectivity for IPv6-enabled data centers across IPv4-enabled networks.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

Summary

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

Summary

- High availability for the IP protocol is achieved by using several technologies: IP routing protocols and first-hop redundancy protocols.
- Clustered devices, such as the Cisco Catalyst 6500 VSS, do not need first-hop redundancy protocols. Semicoupled nonclustered devices, such as Cisco Nexus switches with vPC, need first-hop redundancy protocols to provide high availability, but their behavior is slightly modified. Both devices forward traffic upstream.
- When core and aggregation switches are connected with multiple links, all have the same cost and are considered to distribute the load.
- RHI is a mechanism that can be used to advertise availability of a service into the routing domain.
- LISP is an emerging protocol that has many applications within data centers, including providing for easy virtual machine mobility, IPv6 implementation, and support for multitenant environments.

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

https://t.me/learningnets

certcollecion.net

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

Lesson 6

Designing Data Center Interconnects

Overview

This lesson explains transport options for data center interconnections (DCIs) and main reasons to implement DCIs. These interconnects are crucial for enterprises who want to have globally available data centers to provide continuous services. The purpose of interconnects is to have a link for data replication and for workload mobility. These links are typically of high bandwidth.

Objectives

Upon completing this lesson, you will be able to design data center interconnects for both data traffic and storage traffic, over various underlying technologies. This ability includes being able to meet these objectives:

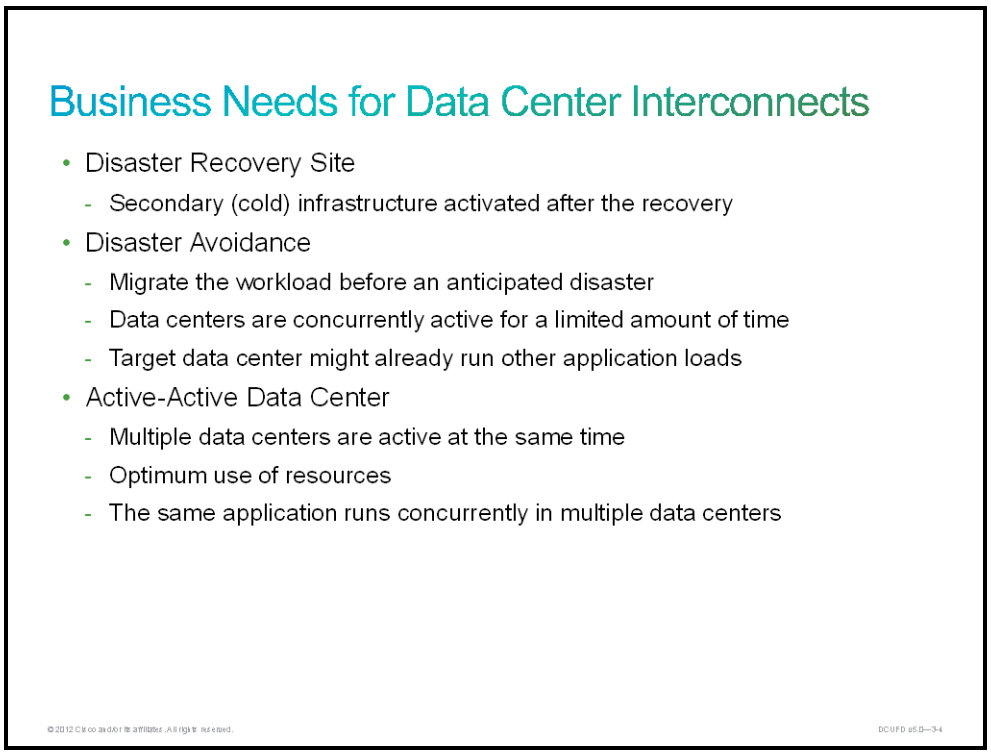
- Identify the reasons for data center interconnects
- Describe data center interconnect technologies
- Design data center interconnects using Cisco OTV
- Describe storage replication technologies

https://t.me/learningnets

certcollecion.net

Reasons for Data Center Interconnects

This topic describes how to identify the reasons for data center interconnects.



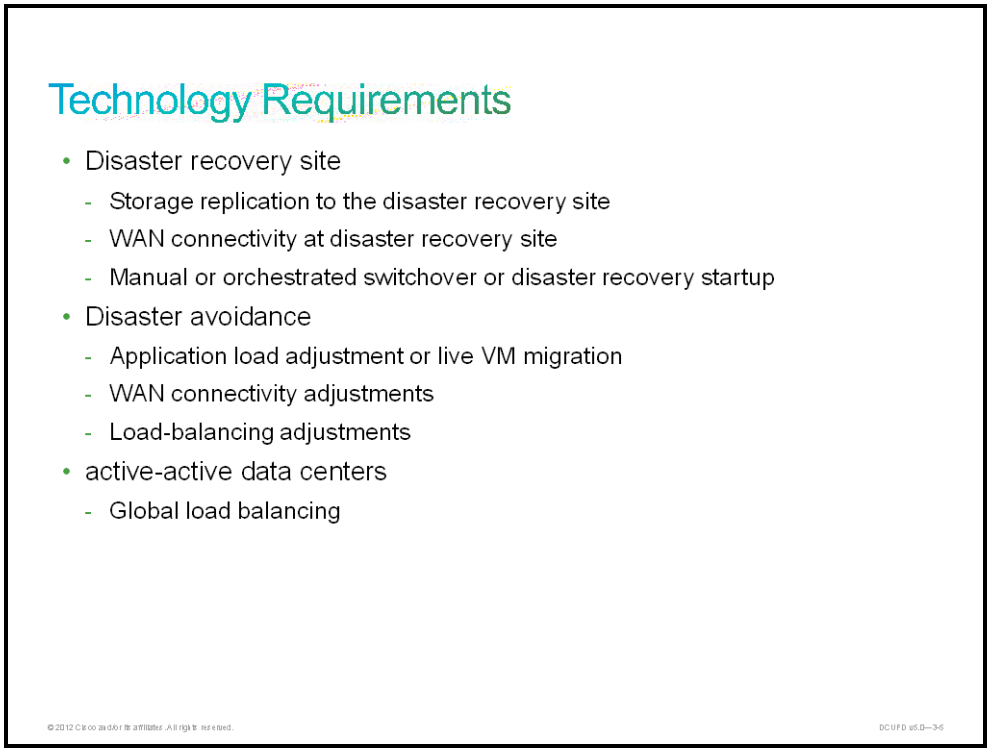
One of the main reasons to implement data center interconnections are business needs, which may require that you use a disaster recovery site that is activated after the recovery. You should always try to lower the probability of a disaster scenario by migration of the workload before an anticipated disaster. In such a case, data centers are concurrently active for a limited amount of time.

Business needs may also dictate that you to use an active-active data center, where multiple data centers are active at the same time. The same application runs concurrently in multiple data centers. This situation represents the optimum use of resources.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net



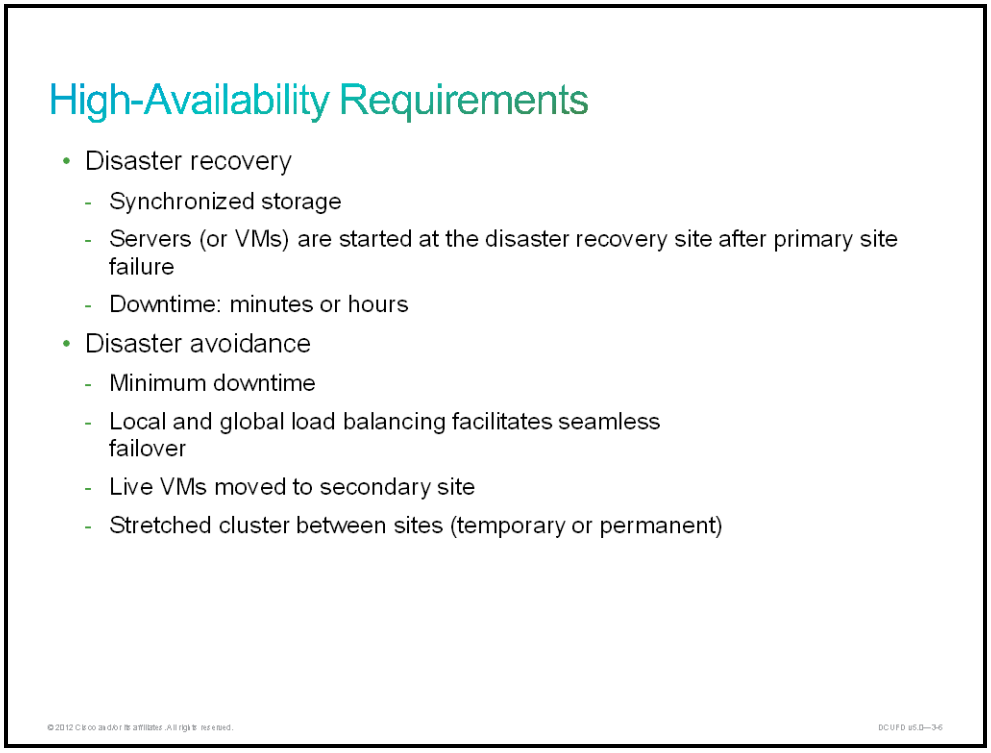
Interconnection of data centers may require replication of storage to the disaster recovery site. For this replication to be possible, you may need WAN connectivity at the disaster recovery site. You should always try to lower the probability of a disaster scenario by adjusting the application load, WAN connectivity, and load balancing.

In the case of an active-active data center, use global load balancing to manage requests and traffic flows between data centers.

Note	The Cisco global server load balancing (GSLB) solution is the Cisco Application Control Engine (ACE) Global Site Selector.

https://t.me/learningnets

certcollecion.net

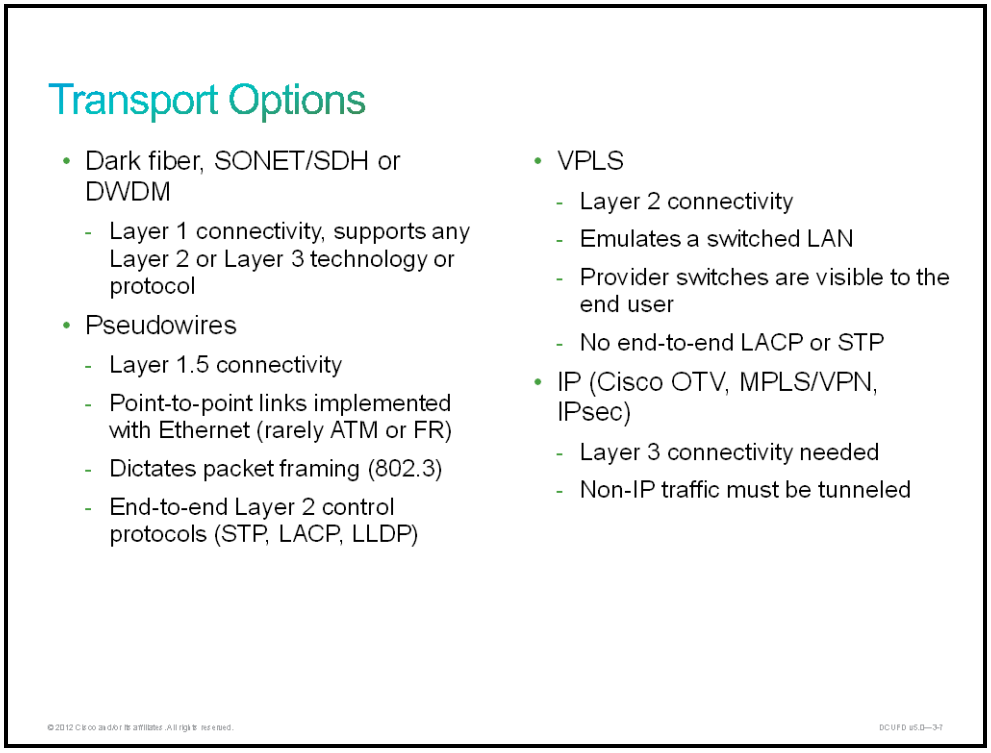


An important aspect of designing data center interconnections is the requirement for high availability. An example of high availability in such a case is duplicated storage. Servers at the disaster recovery site are started after primary site failure. You should always try to lower the probability of a disaster scenario so that you experience minimum downtime. Local and global load balancing facilitates seamless failover. You can also use a temporary or permanent stretched cluster between sites.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

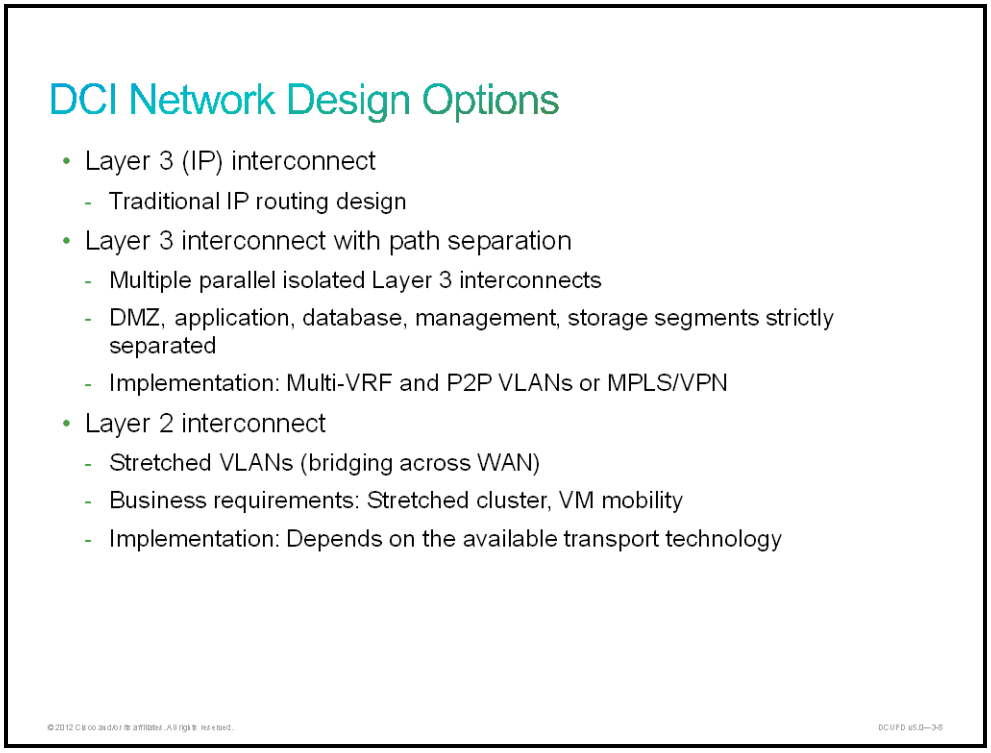


There are several options to interconnect data centers:

- Dark fiber, SONET or SDH, or dense wavelength division multiplexing (DWDM):
  - Layer 1 connectivity supports any Layer 2 and Layer 3 technology or protocol
- Pseudowires:
  - A mechanism providing connectivity above Layer 1 and below Layer 2 that performs emulation and adaptation of Layer 1 mechanisms to transport Layer 2 payload
  - Point-to-point links implemented with Ethernet (rarely ATM or Frame Relay)
  - Dictates packet framing
  - End-to-end Layer 2 control protocols (Spanning Tree Protocol [STP], Link Aggregation Control Protocol [LACP], Link Layer Discovery Protocol [LLDP])
- Virtual Private LAN Services (VPLS):
  - Layer 2 connectivity
  - Emulates a switched LAN
  - Service provider switches are visible to the end user
  - No end-to-end STP or LACP
- IP-based solutions (Cisco Overlay Transport Virtualization [OTV], Multiprotocol Label Switching [MPLS] VPN, IP Security [IPsec]):
  - Plain Layer 3 connectivity is needed
  - Non-IP traffic must be tunneled

https://t.me/learningnets

certcollecion.net



There are several DCI network-side design options:

- Layer 3 (IP) interconnect:
  - Traditional IP routing design
- Layer 3 interconnect with path separation:
  - Multiple parallel isolated Layer 3 interconnects
  - Segments that are strictly separated, such as a demilitarized zone (DMZ), application, database, management, storage
  - Implementation: Multiple virtual routing and forwarding (VRF) and point-to-point VLANs, or MPLS VPN
- Layer 2 interconnect:
  - Stretched VLANs (bridging across WAN)
  - Business requirements: stretched cluster, virtual machine (VM) mobility
  - Implementation: Depends on the available transport technology

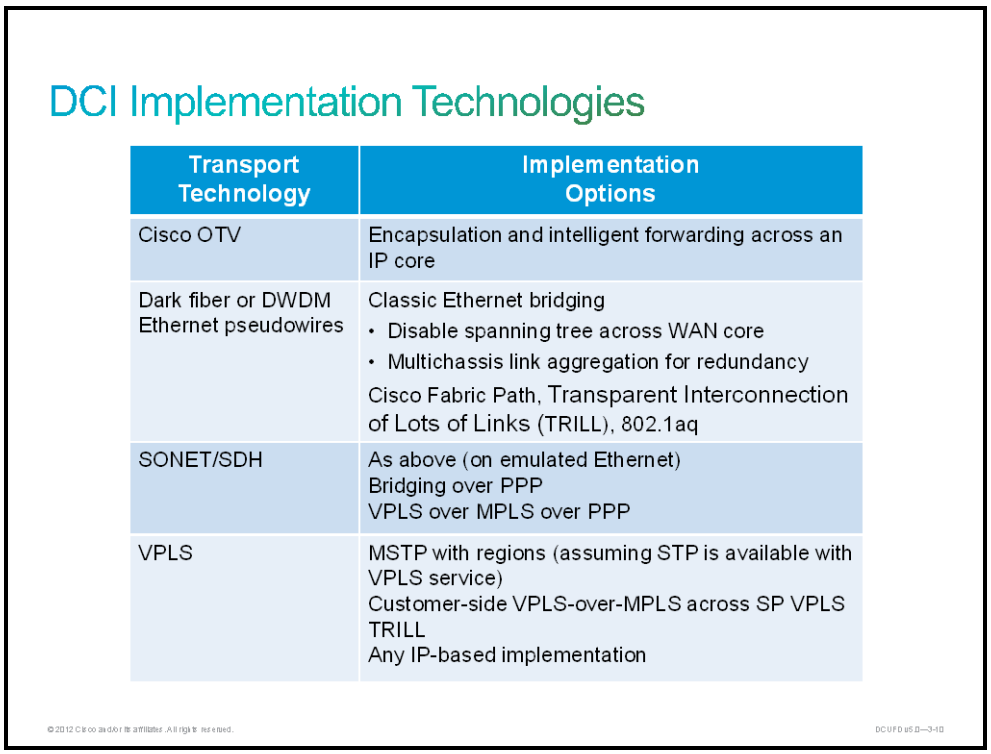
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

Data Center Interconnect Technologies

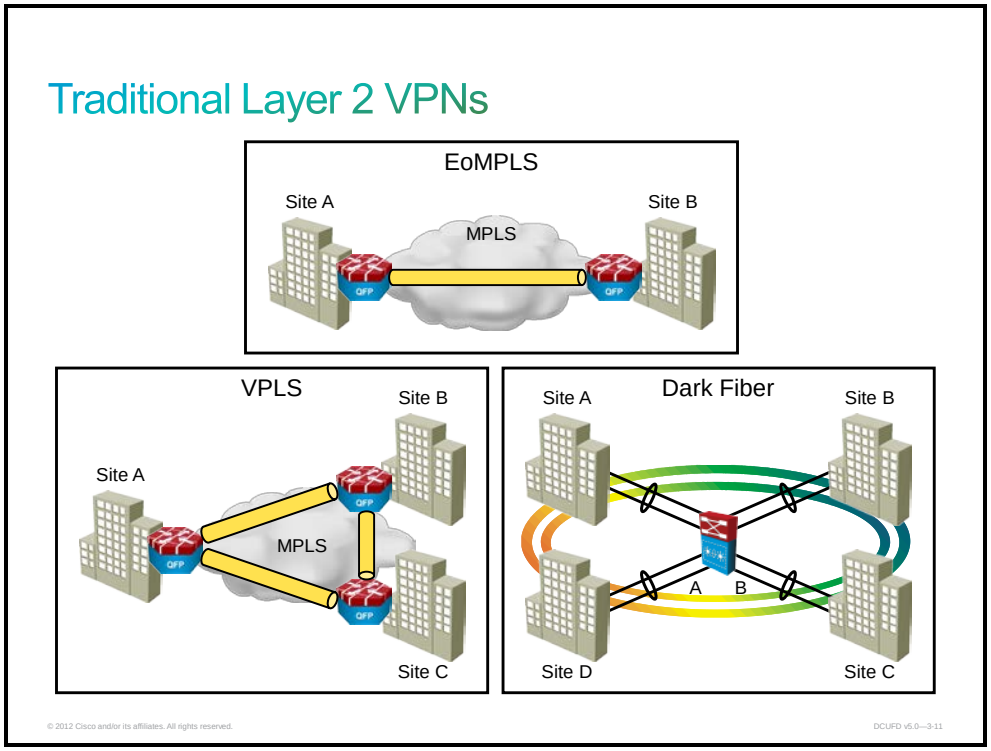
This topic describes data center interconnect technologies.



The table presents Layer 2 DCI transport technologies and their implementation options.

https://t.me/learningnets

certcollecion.net



Traditional Layer 2 VPNs use either tunnels or pseudowires for DCIs. The main disadvantage of those topologies is often complex adding or removing of sites.

In order to facilitate DCIs, several different technologies must be combined, which creates complex configuration.

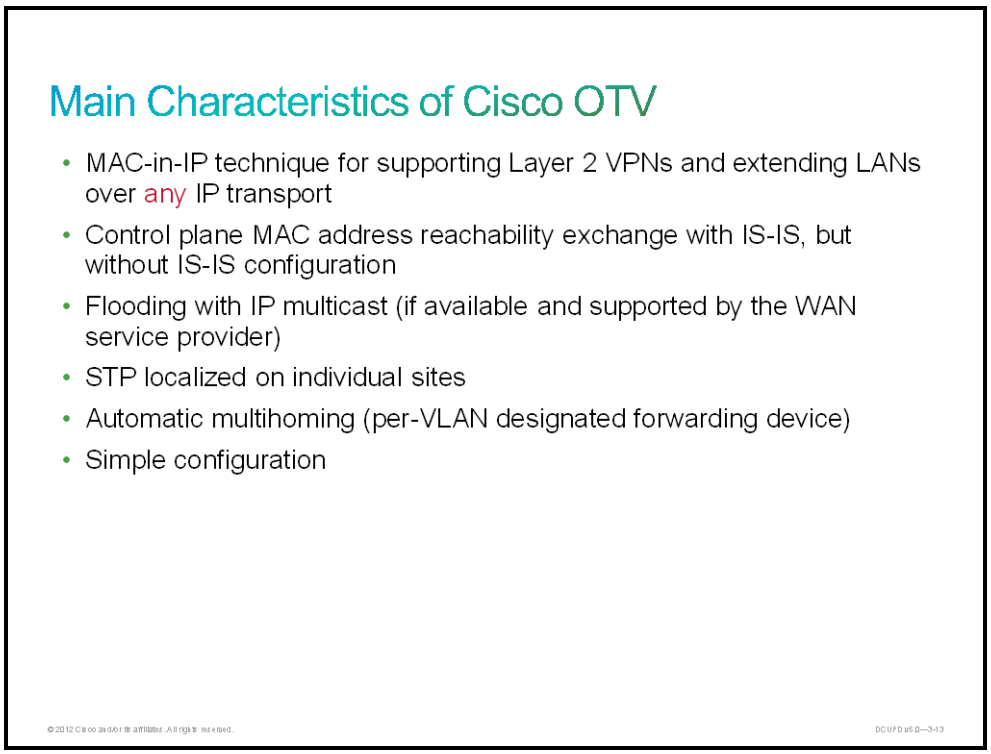
© 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

Cisco OTV

This topic describes how to design data center interconnects using Cisco OTV.



Cisco OTV is a MAC-in-IP Layer 2 tunneling technique that is used to extend selected VLANs between different sites across the IP core.

Robust control plane operation is achieved using Intermediate System-to-Intermediate System (IS-IS) as the underlying protocol. No IS-IS configuration is needed during Cisco OTV configuration.

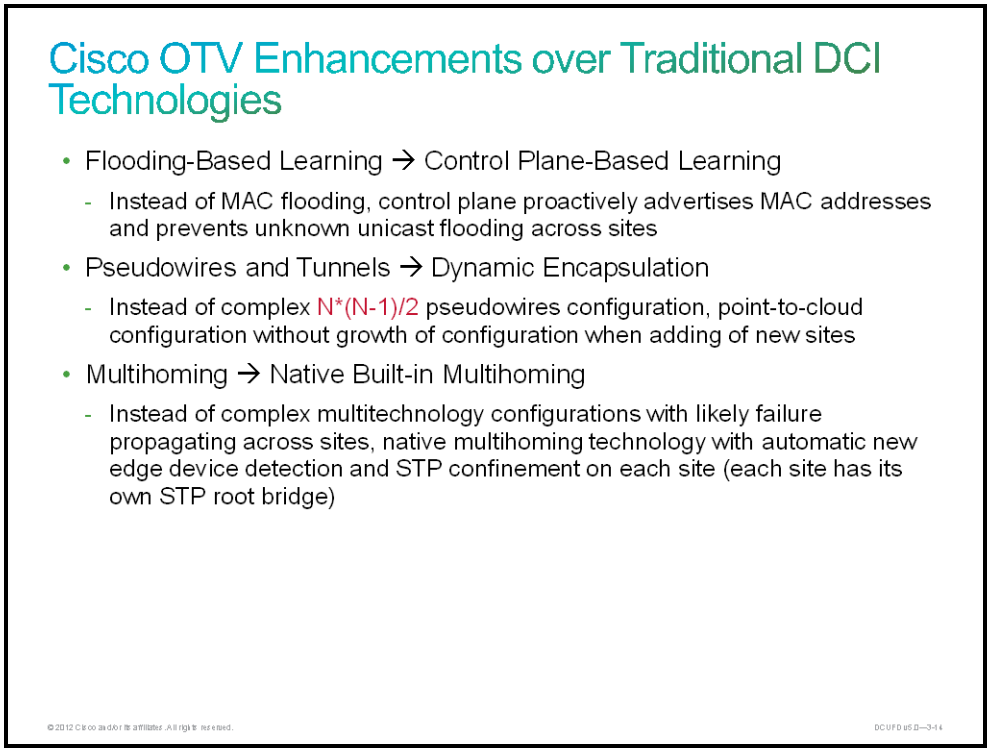
STP is confined to each site with bridge protocol data unit (BPDU) filtered and with it preventing any failure flooding. Each site has its own STP root bridge.

Multihoming is natively supported and used without any additional configuration.

Cisco OTV configuration is composed of several configuration lines on each participating Cisco OTV device and no additional configuration is needed when adding new sites.

https://t.me/learningnets

certcollecion.net



Cisco OTV technology features several enhancements of traditional, Layer 2 VPN data center interconnect technologies:

- **Control plane-based MAC address learning:** Control plane-based learning replaces MAC address learning, which is based on flooding of traffic with unknown or unlearned destinations. The control plane uses IS-IS to send reachability information for MAC addresses, reduce flooding traffic, and improve the efficiency of the DCI.
- **Dynamic encapsulation:** Dynamic encapsulation replaces complex full-mesh topologies. A packet for which the destination is known is encapsulated and sent as unicast.
- **Native built-in multihoming:** Native built-in multihoming greatly simplifies Layer 2 DCI designs and eliminates complex configurations that need to take into account all possible equipment and link failures. Cisco OTV also splits the STP domain so that every site has its local STP domain, and no high-level loops are possible through the DCI.

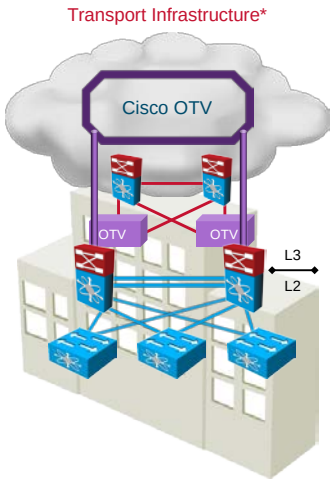
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

Cisco OTV Terminology

- Edge device: Device performing Ethernet-to-IP encapsulation
- Internal interface: Data center-facing interface on edge device
  - Regular Layer 2 interface
- Join interface: WAN-facing uplink interface on edge device
  - Routed interface
  - Edge device is an IP host
- Overlay interface: Virtual interface with Cisco OTV configuration
  - Logical multiaccess multicast-capable interface
  - No spanning tree on overlay interface
- ARP ND cache: ARP snooping reduces intersite ARP traffic
- Site VLAN: VLAN used for edge device discovery
  - Must be configured on internal interfaces
- Authoritative edge device: Edge device performing internal-to-overlay forwarding for a VLAN



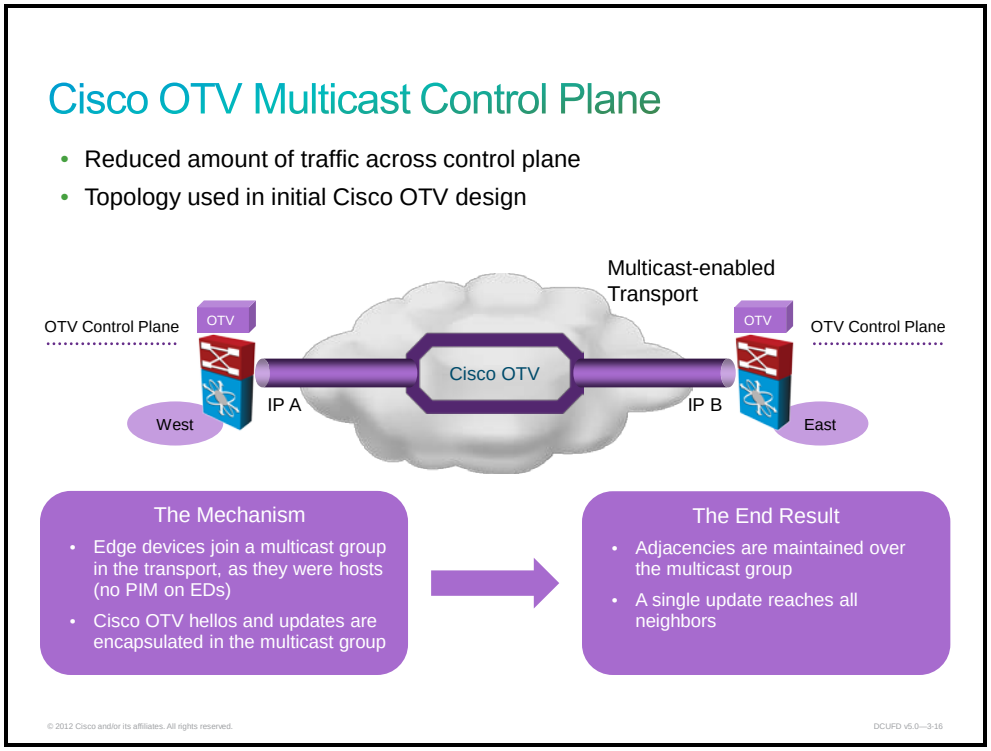
© 2012 Cisco and/or its affiliates. All rights reserved.

The Cisco OTV terminology is as follows:

- **Edge device:** This device performs Ethernet-to-IP encapsulation.
- **Internal interface:** This interface is a data center-facing interface on an edge device. It is a regular Layer 2 interface, such as a VLAN.
- **Join interface:** This interface is a WAN-facing uplink interface on an edge device and is a routed interface.
- **Edge device:** This device is an IP host.
- **Overlay interface:** This interface is a virtual interface with Cisco OTV configuration. This is a logical multiple access and multicast-capable interface. There is no spanning tree on the overlay interface.
- **ARP neighbor discover (ND) cache:** Address Resolution protocol (ARP) snooping reduces intersite ARP traffic.
- **Site VLAN:** This VLAN is used for edge device discovery, and must be configured on internal interfaces. This allows a Cisco OTV device to find another Cisco OTV device on the same site.
- **Authoritative edge device:** This edge device performs internal-to-overlay forwarding for a VLAN.

https://t.me/learningnets

certcollecion.net



The Cisco OTV multicast control plane uses IP multicast on transport infrastructure.

Cisco OTV site adjacency is established by exchanging multicast messages between edge devices on different sites. Once adjacency is established, all control traffic continues to be exchanged between sites using multicast packages. A single Protocol Independent Multicast sparse mode (PIM-SM) or Bidirectional PIM (BIDIR-PIM) group is used for control plane traffic, while data traffic uses other multicast groups.

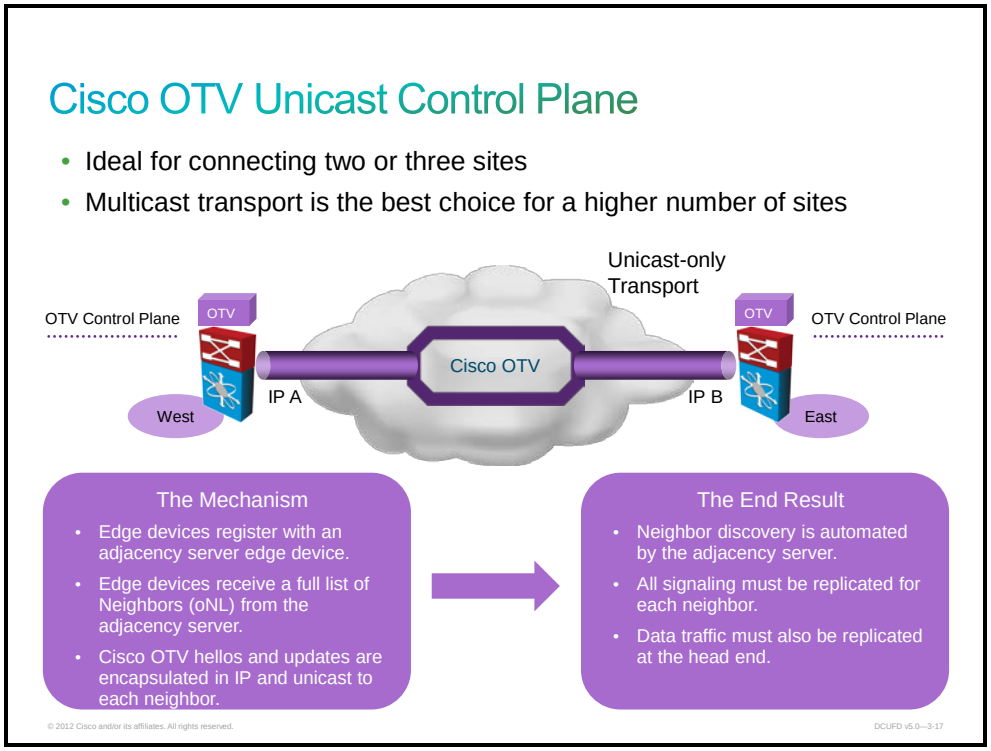
From a multicast perspective, edge devices are multicast hosts. There is no PIM configuration on Cisco OTV edge devices.

Multicast must be supported by transport infrastructure and it is configured by the infrastructure owner (either an enterprise or service provider).

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net



When the transport network does not support multicast or the number of Cisco OTV connected sites is low, the Cisco OTV unicast control plane can be used.

Note	When using Cisco OTV over a unicast control plane (that is, when the service provider does not support relaying multicast traffic), this approach has a cost. Each Cisco OTV device needs to replicate each control plane packet and unicast it to each remote Cisco OTV device that is part of the same logical overlay.

Instead of announcing themselves across a multicast group, edge devices announce their presence to a configured adjacency server. Neighbor discovery is achieved by querying the adjacency server on the Cisco OTV cloud using unicast packets.

The adjacency server is not a separate device on the network, but a service that can run on any edge device.

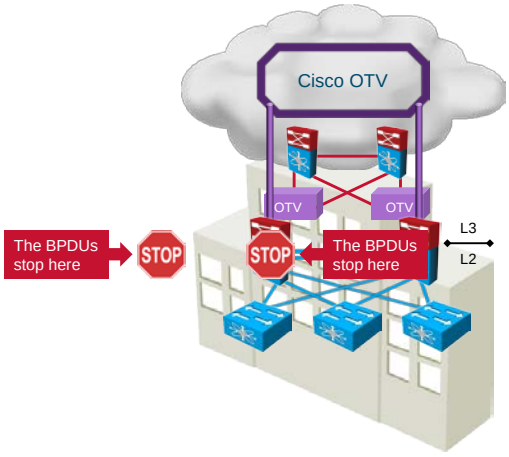
In the Cisco OTV unicast control plane, all traffic between edge devices is IP unicast, so there is no additional configuration required by the transport network owner.

https://t.me/learningnets

certcollecion.net

Spanning Tree and Cisco OTV

- Cisco OTV is site transparent. There are no changes to the STP topology.
- Each site keeps its own STP domain.
- This functionality is built into Cisco OTV and no additional configuration is required.
- An edge device will send and receive BPDUs only on the Cisco OTV internal interfaces.



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

The Cisco OTV cloud prevents STP traffic from flowing between edge devices. Edge devices only send and receives BPDUs on Cisco OTV internal interfaces. Consequently, each Cisco OTV site is a separate STP domain with its own root bridge switch.

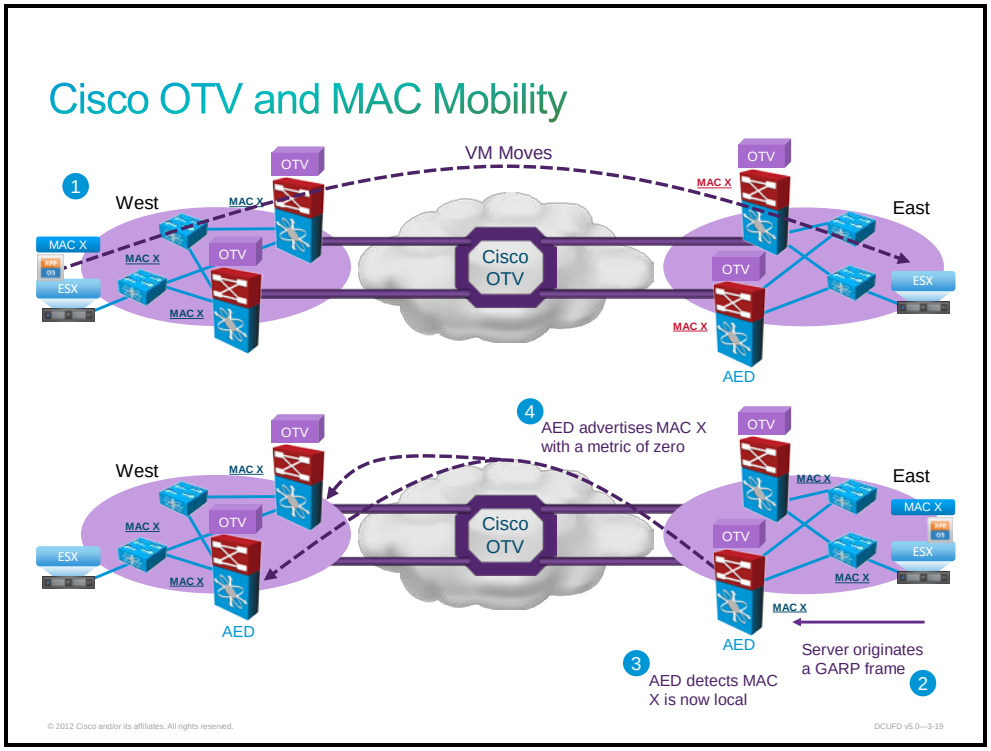
Because STP domains are separated, a failure on any STP site does not influence traffic on any other site, so possible damage is contained within single site.

Loop prevention on the Cisco OTV cloud itself is performed using IS-IS loop prevention.

© 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

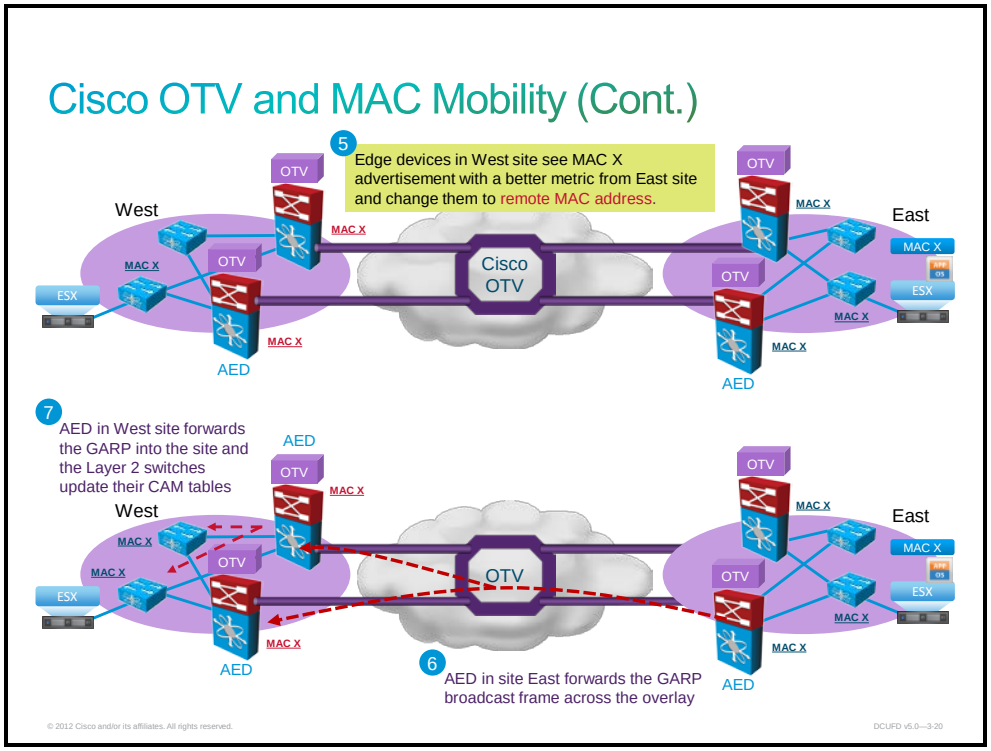


Cisco OTV is ideal for solving issues with MAC mobility that might occur during migration of live VMs between sites:

1. VM is moved from the West site to the East site due to VMware VMotion activity.
2. Once on the new site, the East VM will send a Gratuitous ARP (GARP) frame to authoritative edge device (AED) for its VLAN.
3. The AED detects that the VM MAC is now local and sends a GARP frame to all Layer 2 switches on the local site.
4. The AED advertises the VM MAC with metric 0 to all edge devices on the Cisco OTV overlay.

https://t.me/learningnets

certcollecion.net



5. The edge device on the West site will receive an advertisement with a better metric for the VM MAC and change the VM MAC from a local to a remote address.

6. The AED on the East site will forward a GARP broadcast across the Cisco OTV overlay to other edge devices.

7. The AED on the West site will forward a GARP broadcast into the site and the Layer 2 switches will update their content-addressable memory (CAM) tables with the local AED as the target for the VM MAC.

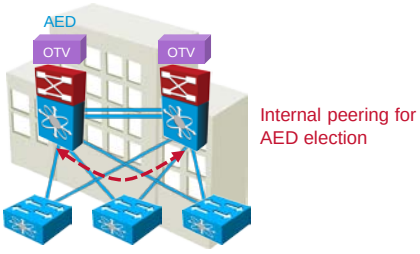
 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

Cisco OTV Automated Multihoming

- The detection of multihoming is fully automated and it does not require additional protocols and configuration.
- Cisco OTV elects one of the edge devices to be the authoritative edge device (AED) for a subset of the OTV extended VLANs:
  - Election is done within the site using the Cisco OTV site VLAN.
- VLANs are split between the Cisco OTV edge device, and each edge device is responsible for its own VLAN traffic:
  - Lower system ID manages even-numbered VLANs.
  - Higher system ID manages odd-numbered VLANs.



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

Multihoming is configured automatically at sites where more than one edge device is connected to the Cisco OTV overlay. Cisco OTV will elect one of edge device to be the AED for a set of VLANs. This election is done automatically over the Cisco OTV site VLAN, which is the VLAN that is configured for communication between edge devices on site.

VLANs are split between edge devices and each edge device is responsible for its own VLAN traffic across the Cisco OTV overlay.

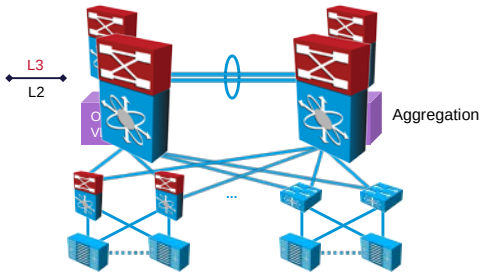
Even-numbered VLANs are managed by the edge device with the lower IS-IS system ID. Odd-numbered VLANs are managed by the edge device with the higher IS-IS system ID. VLAN allocation is currently not configurable, but will be in future Cisco OTV releases.

https://t.me/learningnets

certcollecion.net

Cisco OTV and SVI Separation

- Guideline: The current Cisco OTV implementation on the Cisco Nexus 7000 enforces the separation between SVI routing and OTV encapsulation for a given VLAN.
- This separation can be achieved by having two separate devices to perform these two functions.
- An alternative, cleaner, and less intrusive solution is the use of VDCs, which are available with the Cisco Nexus 7000 platform:
  - A dedicated Cisco OTV VDC to perform the OTV functionalities
  - The aggregation VDC used to provide SVI routing support



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

Current Cisco OTV implementation requires that the Cisco OTV join interfaces are physical interfaces on an M-family module. Due to that restriction for VLANs that are extended across the Cisco OTV overlay, switch virtual interface (SVI) interface is not available.

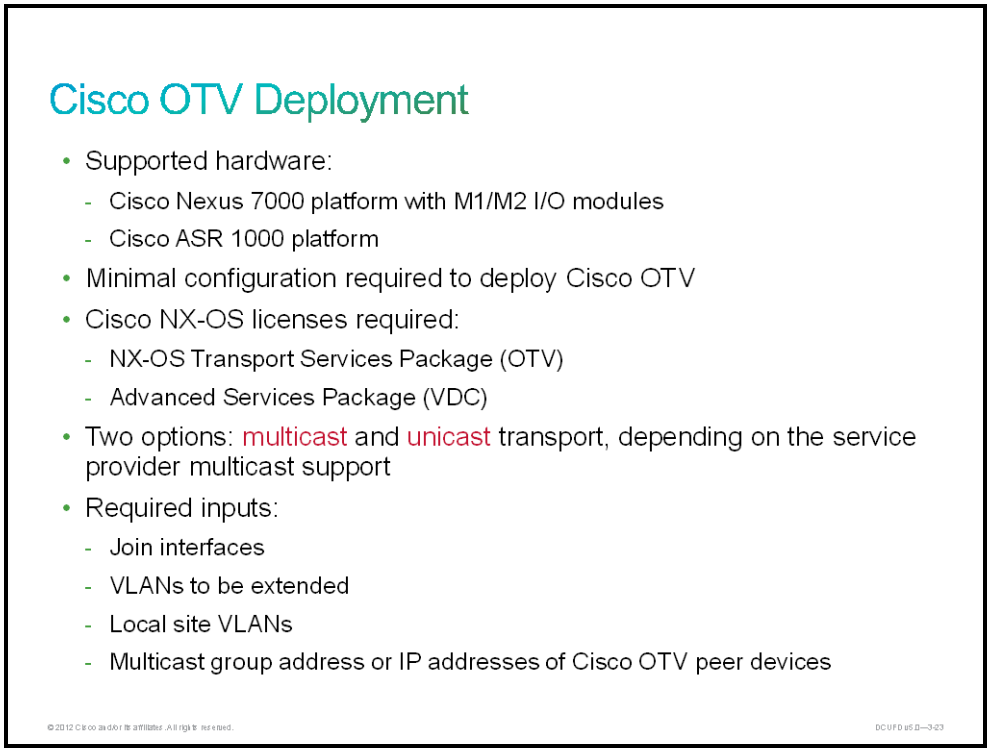
If network design requires an SVI interface on a VLAN that is extended across Cisco OTV, following solutions can be used:

- The SVI interface can be on a separate physical device that is connected to that VLAN.
- The SVI interface can be in a separate virtual device context (VDC), which is available with the Cisco Nexus 7000 platform:
  - A dedicated VDC should be configured for Cisco OTV.
  - Another VDC should provide SVI routing service.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net



Cisco OTV functionality is delivered on different hardware platforms:

- Cisco Nexus 7000 Series Switches:
  - Initial hardware platform with Cisco OTV support
  - Cisco OTV supported on M family line cards
  - Licensed feature
- Cisco ASR 1000 Series Aggregation Services Routers:
  - Cisco OTV (Phase 1) supported on all platforms with Cisco IOS XE 3.5S
  - Advanced switch image

Note This allows you to terminate Cisco OTV on different hardware in the primary and in the secondary data center, in case you do not have the same equipment available.

The following licenses are required to deploy Cisco OTV using Cisco Nexus 7000 Series Switches:

- Transport Services Package (LAN_TRANSPORT_SERVICES_PKG) to enable Cisco IOS XE 3.5S OTV functionality
- Advanced Services Package (LAN_ADVANCED_SERVICES_PKG) to enable VDCs

Cisco OTV configuration on each site consists of only a few commands.

On the switch global level, in addition to enabling Cisco OTV functionality, only Cisco OTV site VLAN configuration is needed. In addition, only the overlay interface needs configuration.

Note These configuration examples are listed here only to show how simple it is to design and deploy Cisco OTV.

https://t.me/learningnets

certcollecion.net

The required input parameters to design a Cisco OTV deployment are the following:

- VLANs to be extended between sites
- Join interfaces: Interfaces that join the Cisco OTV overlay
- Site VLANs: VLANs local to a site to control multihoming
- Multicast group address for the Cisco OTV control plane or unicast addresses of the Cisco OTV peers if the service provider does not support multicast.

Cisco OTV Multicast Configuration Example

- West device:

```
feature otv
otv site-vlan 99
interface Overlay1
  description WEST-DC
  otv join-interface e1/1
  otv control-group 239.1.1.1
  otv data-group 232.192.1.0/24
  otv extend-vlan 100-15
```

- South device:

```
feature otv
otv site-vlan 99
interface Overlay1
  description SOUTH-DC
  otv join-interface Po16
  otv control-group 239.1.1.1
  otv data-group 232.192.1.0/24
  otv extend-vlan 100-150
```

- East device:

```
feature otv
otv site-vlan 99
interface Overlay1
  description EAST-DC
  otv join-interface e1/1.10
  otv control-group 239.1.1.1
  otv data-group 232.192.1.0/24
  otv extend-vlan 100-150
```

Cisco OTV Unicast Configuration Example

On the switch global level, in addition to enabling Cisco OTV functionality, Cisco OTV site VLAN configuration is needed. In addition, only the overlay interface needs configuration.

- West device:

```
feature otv
otv site-vlan 99
interface Overlay1
  description WEST-DC
  otv join-interface e1/1
  otv adjacency-server local
```

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

```
  otv extend-vlan 100-150
```

- South device:

```
feature otv
otv site-vlan 99
interface Overlay1
  description SOUTH-DC
  otv join-interface Po16
  otv adjacency-server 10.1.1.1
  otv extend-vlan 100-150
```

- East device:

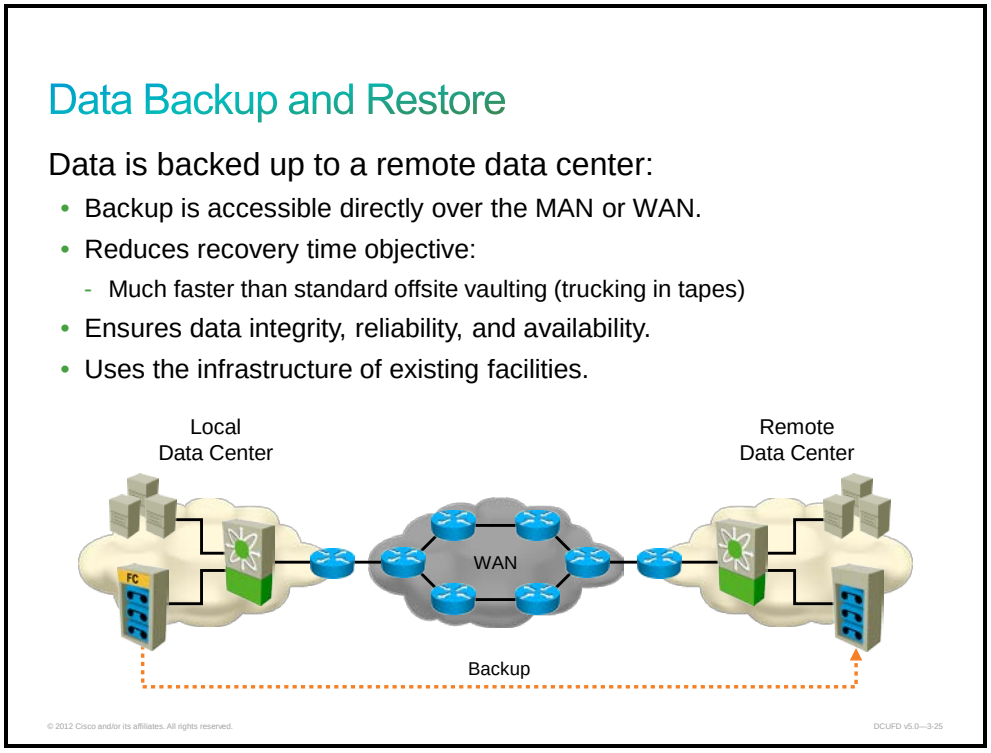
```
feature otv
otv site-vlan 99
interface Overlay1
  description EAST-DC
  otv join-interface e1/1.10
  otv adjacency-server 10.1.1.1
  otv extend-vlan 100-150
```


https://t.me/learningnets

certcollecion.net

Storage Replication Technologies and Interconnects

This topic describes storage replication technologies.



Remote backup is a core application for Fibre Channel over IP (FCIP). It is sometimes known as remote vaulting. Backup is accessible directly over the WAN or metropolitan-area network (MAN). In this approach, data is backed up with the use of standard backup applications, such as Veritas NetBackup or Legato Celestra Power, but the backup site is located at a remote location. FCIP is an ideal solution for remote backup applications for several reasons:

- FCIP is relatively inexpensive, compared to optical storage networking.
- Enterprises and storage service providers can provide remote vaulting services by using existing IP WAN infrastructures.
- Backup applications are sensitive to high latency, but in a properly designed SAN, the application can be protected from problems with the backup process by the use of techniques such as snapshots and split mirrors.

 © 2012 Cisco Systems, Inc.

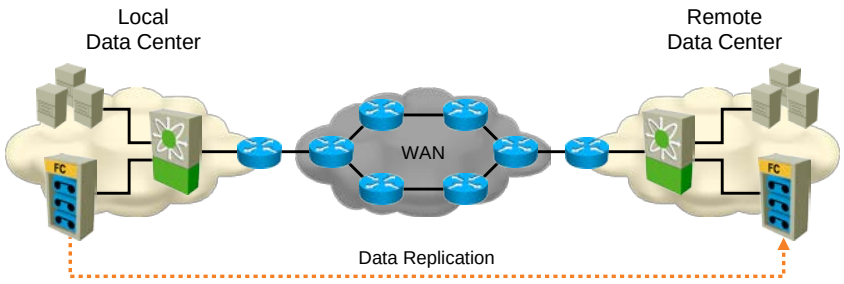
https://t.me/learningnets

certcollecion.net

Data Replication

Data is continuously synchronized across the network:

- Data can be mirrored for multiple points of access.
- Data replication enables rapid failover to a remote data center and business continuity for critical applications.
- Data replication reduces the recovery time objective.
- Data replication reduces the recovery point objective.



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

The primary type of application for an FCIP implementation is a disk replication application that is used for business continuance or disaster recovery.

Here are some examples of these types of applications:

- Array-based replication schemes such as EMC Symmetrix Remote Data Facility (SRDF), Hitachi TrueCopy, IBM Peer-to-Peer Remote Copy (PPRC), or HP-Compaq Data Replication Manager (DRM)
- Host-based application schemes such as Veritas Volume Replicator (VVR)

https://t.me/learningnets

certcollecion.net

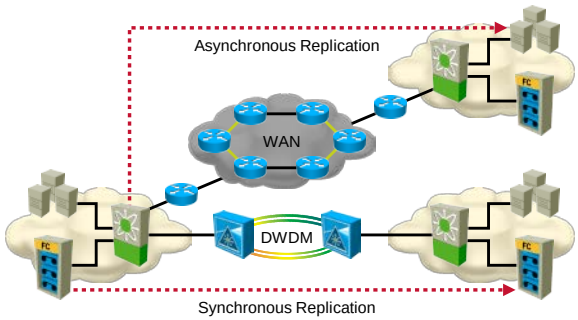
Synchronous and Asynchronous Replication

Synchronous replication:

- Data must be written to both arrays before I/O operation is complete
- Data in both arrays is always fully synchronized
- Distance between sites influences application performance

Asynchronous replication:

- Data for the remote site is cached and replicated later
- Trade-off between performance and business continuity
- Can extend over greater distances and use high-latency transport



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

Replication applications can be run in synchronous mode or asynchronous mode.

Synchronous Replication

In synchronous mode, an acknowledgment of a disk write is not sent until copying to the remote site is completed. Applications that use synchronous copy replication are very sensitive to latency delays and might be subject to unacceptable performance. This is why synchronous replication is supported only up to a certain distance.

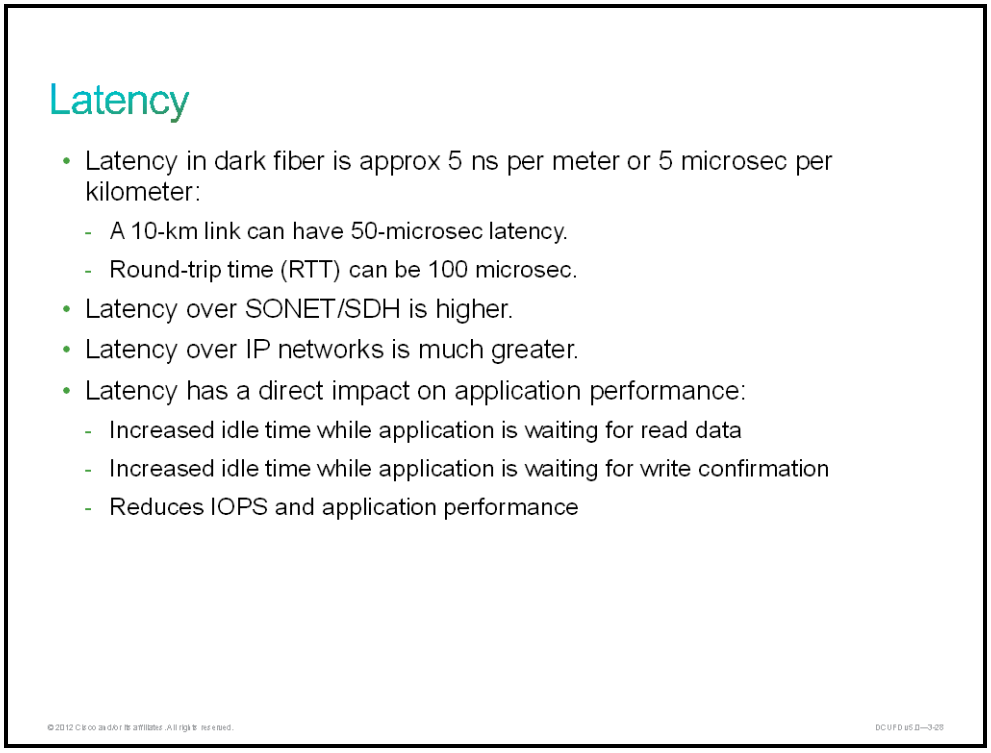
The local array does not acknowledge the data to be written until it has been written to the remote array as well. Consequently, the data on both storage arrays is always up to date.

Asynchronous Replication

In asynchronous mode, disk writes are acknowledged before the remote copy is completed. The data to the remote storage array is cached and written afterward. The application response time is shortened because the application does not need to wait for confirmation that the data has been written on the remote location as well.

https://t.me/learningnets

certcollecion.net



Here are some characteristics of latency:

- Latency in dark fiber is approximately 5 ns per meter or 5 microsec per kilometer (1 kilometer equals 0.62 miles):
  - Therefore, a 10-km (6.2-mile) fiber channel link can have a 5 * 10 = 50-microsec latency and a 2 * 5 * 10 = 100-microsec round-trip latency.
- Latency over SONET/SDH is higher because of the added latency of the infrastructure.
- Latency over IP networks is much greater because of the added latency of the infrastructure and delays due to TCP/IP processing.
- Latency has a direct impact on application performance:
  - **Read operation:** The application is idle, waiting for data to arrive.
  - **Write operation:** The application is idle, waiting for write confirmation before it can proceed.

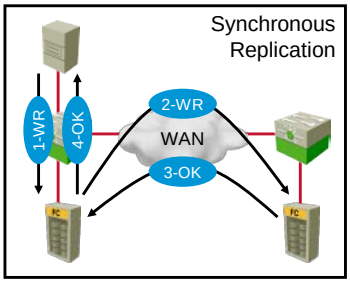
Note	The added idle time can significantly reduce the I/O operations per second (IOPS) that the server can achieve.

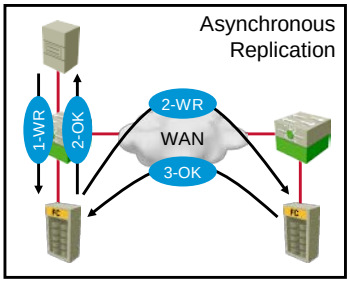
https://t.me/learningnets

certcollecion.net

Back-End and Storage Connectivity

- FC
- FCIP
- FCoE
- iSCSI (use checksum)





Alternative: Distributed file system with NFS

Transport	FC	iSCSI
DWDM/fiber	✓	✓
Pseudowires	FCIP	✓
VPLS	FCIP	✓
IP	FCIP	✓

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

There are several options for back-end and storage connectivity: Fibre Channel, FCIP, Fibre Channel over Ethernet (FCoE), and Internet Small Computer Systems Interface (iSCSI). Next-generation Nexus equipment supports distances for FCoE of a few kilometers.

SAN Extension Solutions

SAN Extension Solution	Dark Fiber	CWDM	DWDM	SDH/ SONET	FCIP
Distance	Up to 10 km	Up to 100 km	Up to 200 km	Thousands of km	Ten thousands of km
Latency	5 microsec/ km	5 microsec/ km	Low	Medium	High
Bandwidth	1000 MB/s on 10 Gb/s FC	8 channels x 4 Gb/s	80 or 96 channels x 10 Gb/s	OC-192 ~ 10 Gb/s	1 Gb/s per tunnel
Reliability	BER $1x\ 10^{12}$	BER $1x\ 10^{12}$	High	High	Medium
Cost	Low	Low	High	Medium	Low

BER = Bit error rate

CWDM = coarse wavelength-division multiplexing

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

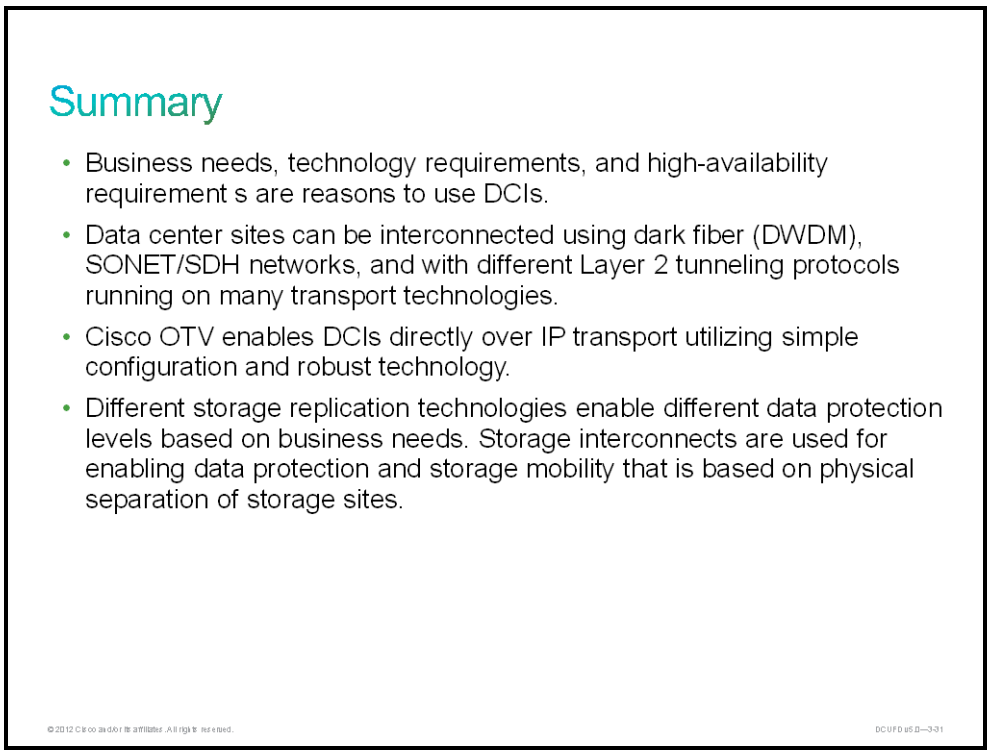
The figure summarizes SAN extension solutions and compares capabilities like maximum distance, latency, bandwidth, reliability, and relative cost.

https://t.me/learningnets

certcollecion.net

Summary

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



References

For additional information, refer to these resources:

- Cisco Data Interconnect

  http://www.cisco.com/go/dci

- Overlay Transport Virtualization

  http://www.cisco.com/go/otv

https://t.me/learningnets

certcollecion.net

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

Module Summary

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

Module Summary

- The data center core layer interconnects several aggregation blocks with the campus core or enterprise edge networks. The main activity performed in a data center core is fast packet switching and load balancing between links, without oversubscription. IP routing protocols and ECMP are used in the core.
- The data center aggregation layer aggregates connections from access switches. This layer is typically the boundary between the Layer 2 and Layer 3 network topologies. IP Services, which include firewalling and server load balancing, are installed in this layer. The aggregation layer can be combined with the core layer for smaller designs, forming a collapsed core layer. Some of the best designs involve virtualization with VDCs.
- The data center access layer is used to connect servers to the network and is the largest layer considering the number of devices. The focus is on technologies such as Cisco Unified Fabric (to save on cabling and equipment) and FEXs for improved ease of management.

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

Module Summary (Cont.)

- The virtual access layer interconnects virtual machines with the physical network. Various products and technologies are available for the virtualized access layer, including the Cisco Nexus 1000V switch, virtual appliances, and hardware-assisted switching between virtual machines using Cisco VM-FEX.
- To provide high availability on the IP level, several routing protocols and technologies are available, including OSPF, EIGRP, BGP, and LISP. All these protocols allow for highly available designs.
- Data center interconnection technologies can extend Layer 2 domains between data centers. This is a foundation for workload mobility in disaster avoidance or disaster recovery scenarios. The main technologies in this area are Cisco OTV, Layer 2 MPLS VPN, VPLS, dark fiber, and CWDM and DWDM.

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

https://t.me/learningnets

certcollecion.net

In this module, you learned about various components of data center networks (the core, aggregation, access, and virtual access layers), highly available IP designs, and data center interconnections. Understanding the role of every component of the data center allows you to design data center networks functionally and optimally. A number of technologies are used in combination to provide the best efficiency, ease of management, and utilization of equipment.

For example, virtual device contexts (VDCs) are used to virtualize equipment, Cisco Unified Fabric is used to optimize the number of links and device utilization, virtual access layer devices are used to provide efficient management for virtual networks inside server virtualization hosts, and data center interconnect solutions are used to provide workload mobility.

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

Module Self-Check

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

Q1) What is the classic division of a hierarchical network? (Source: Designing the Data Center Core Layer Network)

A) access—aggregation—core
B) management—policy control—policy enforcement
C) routing—switching—inspection
D) hypervisor—kernel

Q2) Under which circumstances would you implement a collapsed core layer? (Source: Designing the Data Center Core Layer Network)

Q3) Which option is a reason to implement a Layer 2 core? (Source: Designing the Data Center Core Layer Network)

A) when IP routing cannot scale to the required level
B) when a very large Layer 2 domain is required
C) when the data center is used as a web server farm
D) when performing server load balancing at the access layer

Q4) Where is the common Layer 2 termination point in data center networks? (Source: Designing the Data Center Aggregation Layer)

A) data center core layer
B) data center aggregation layer
C) data center access layer
D) data center virtual access layer

Q5) Which two technologies are used to optimize bandwidth utilization between the access and aggregation layers? (Choose two.) (Source: Designing the Data Center Aggregation Layer)

A) per-VLAN RSTP
B) MEC
C) vPC
D) OSPF

Q6) Which three combinations can be done and make sense with VDCs at the aggregation layer? (Choose three.) (Source: Designing the Data Center Aggregation Layer)

A) core and aggregation VDC
B) aggregation and storage VDC
C) multiple aggregation VDCs
D) multiple access layer VDCs
E) multiple core layer VDCs

https://t.me/learningnets

certcollecion.net

Q7) When using a storage VDC on the Cisco Nexus 7000 Series Switch in the aggregation layer, in which Fibre Channel mode must the storage VDC operate? (Source: Designing the Data Center Aggregation Layer)

A) pinning mode
B) FCoE NPV mode
C) Fibre Channel switch mode
D) Fibre Channel transparent mode

Q8) Which FEX does not support multihoming to two managing switches? (Source: Designing the Data Center Access Layer)

A) Cisco Nexus 2248P
B) Cisco Nexus 2148T
C) Cisco Nexus 2224TP
D) Cisco Nexus 2232PP

Q9) Which kind of wiring needs to be installed to support migration of the access layer from the spanning tree design to the vPC design? (Source: Designing the Data Center Access Layer)

A) loop-free inverted-U wiring
B) triangle-loop wiring
C) square-loop wiring
D) loop-free U wiring

Q10) What is the recommended mode for an access switch when designing a Cisco Unified Fabric deployment? (Source: Designing the Data Center Access Layer)

A) FCoE FCF mode
B) Fibre Channel switch mode
C) FCoE NPV mode
D) domain manager mode

Q11) What are the two purposes of a virtual access layer? (Choose two.) (Source: Designing the Data Center Virtualized Access Layer)

A) provides network communication between virtual machines
B) provides firewalling for network edge
C) provides access for the virtual machines to the physical network
D) provides server load-balancing capabilities, as in physical access layer
E) provides console management access for virtual machines

Q12) Which three Cisco technologies or solutions are used in the virtual access layer? (Choose three.) (Source: Designing the Data Center Virtualized Access Layer)

A) Cisco Distributed Virtual Switch
B) Cisco Nexus 1000V switch
C) Cisco UCS Pass-through switching with VM-FEX
D) Cisco Adapter-FEX when using Cisco UCS C-Series servers
E) Cisco Virtual Services Appliance

Q13) Which protocol or solution provides default gateway redundancy? (Source: Designing High Availability)

A) HSRP
B) OSPF
C) RIP
D) NHRP

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets

certcollecion.net

Q14) Which OSPF area type is most suitable for data center usage? (Source: Designing High Availability)

A) OSPF totally stubby area
B) OSPF not-so-stubby area
C) OSPF Area 0
D) OSPF virtual link

Q15) Which two mechanisms are required to implement anycast? (Choose two.) (Source: Designing High Availability)

A) route redistribution
B) gratuitous ARP
C) RHI
D) transparent mode firewall
E) rendezvous points

Q16) What are the two possible uses of LISP? (Choose two.) (Source: Designing High Availability)

A) inter-AS multicast routing
B) Virtual Machine Mobility
C) IPv6 transition
D) low latency server failover

Q17) Which underlying technology does Cisco OTV require to establish an overlay link between two data centers? (Source: Designing Data Center Interconnects)

A) IP multicast
B) IP
C) Ethernet
D) DWDM

https://t.me/learningnets

certcollecion.net

Module Self-Check Answer Key

Q1) A

Q2) when a separate core is not needed because of small data center size

Q3) B

Q4) B

Q5) B, C

Q6) A, B, C

Q7) C

Q8) B

Q9) B

Q10) C

Q11) A, C

Q12) B, C, D

Q13) A

Q14) B

Q15) A, C

Q16) B, C

Q17) B

 © 2012 Cisco Systems, Inc.

https://t.me/learningnets