



**642-891**

**Composite  
(BSCI and BCRAN)**

**Study Guide  
DEMO Version**

Copyright (c) 2003 Chinatag LLC. All rights reserved.

## **Important Note Please Read Carefully**

For demonstration purpose only, this free version Chinatag study guide contains **10** full length questions selected from our full version products which have more than **200** questions each.

This Study guide has been carefully written and compiled by Chinatag certification experts. It is designed to help you learn the concepts behind the questions rather than be a strict memorization tool. Repeated readings will increase your comprehension.

For promotion purposes, all PDF files are **not** encrypted. Feel free to distribute copies among your friends and let them know Chinatag website. Our IT certification products start at only **\$7.99**.

## **Study Tips**

This product will provide you questions and answers along with detailed explanations carefully compiled and written by our experts. Try to understand the concepts behind the questions instead of cramming the questions. Go through the entire document at least twice so that you make sure that you are not missing anything.

## **Latest Version**

We are constantly reviewing our products. New material is added and old material is revised. Free updates are available for 180 days after the purchase. You should check the products page on the <http://www.chinatag.com> website for an update 3-4 days before the scheduled exam date.

Please tell us what you think of our products. We appreciate both positive and critical comments as your feedback helps us improve future versions. Feedback on specific questions should be send to [feedback@chinatag.com](mailto:feedback@chinatag.com).

Thanks for purchasing our products and look forward to supplying you with all your Certification training needs.

Good studying!

Technical and Support Team  
Chinatag LLC.

## **TABLE OF CONTENTS**

### **List of Tables**

### **List of Acronyms**

### **Introduction**

## **1. The Campus Network**

- 1.1 The Traditional Shared Campus Network
  - 1.1.1 Collisions
  - 1.1.2 Bandwidth
  - 1.1.3 Broadcasts and Multicasts
- 1.2 The New Campus Network
- 1.3 The 80/20 Rule and the New 20/80 Rule
- 1.4 Characterizing Scalable Internetworks
  - 1.4.1 Reliability and Availability
  - 1.4.2 Responsiveness
  - 1.4.3 Efficiency
  - 1.4.4 Adaptability and Serviceability
  - 1.4.5 Accessibility and Security
- 1.5 Network Congestion
  - 1.5.1 Problems Created by Network Congestion
    - 1.5.1.1 Excessive Traffic
    - 1.5.1.2 Dropped Packets
    - 1.5.1.3 Retransmission of Packets
    - 1.5.1.4 Incomplete Routing Tables
    - 1.5.1.5 Incomplete Server Lists
    - 1.5.1.6 The Spanning-Tree Protocol Breaks
    - 1.5.1.7 Runaway Congestion
  - 1.5.2 Symptoms of Network Congestion
    - 1.5.2.1 Application Time Outs
    - 1.5.2.2. Clients Cannot Connect to Network Resources
    - 1.5.2.3 Network Death Results
- 1.6 Designing Scalable Networks
  - 1.6.1 Open Systems Interconnection Model
    - 1.6.1.1 Data Encapsulation
    - 1.6.1.2 Layer 2 Switching
    - 1.6.1.3 Layer 3 Switching
    - 1.6.1.4 Layer 4 Switching

## **CCNP/CCDP 642-891 (Composite)**

- 1.6.1.5 Multi-Layer Switching (MLS)
- 1.6.2 The Cisco Hierarchical Model
  - 1.6.2.1 Core Layer
  - 1.6.2.2 Distribution Layer
  - 1.6.2.3 Access Layer
- 1.6.3 Modular Network Design
  - 1.6.3.1 The Switch Block
  - 1.6.3.2 The Core Block
    - 1.6.3.2.1 The Collapsed Core
    - 1.6.3.2.2 Dual Core
    - 1.6.3.2.3 Core Size
    - 1.6.3.2.4 Core Scalability
    - 1.6.3.2.5 Layer 3 Core
  - 1.6.3.3 Additional Building Blocks
- 1.7 Alleviating Congestion
  - 1.7.1 Access Lists
  - 1.7.2 Extended Access Lists
  - 1.7.3 Distribution Lists
  - 1.7.4 Other Solutions to Traffic Control
  - 1.7.5 Prioritization
    - 1.7.5.1 First In, First Out (FIFO)
    - 1.7.5.2 Weighted Fair Queuing (WFQ)
    - 1.7.5.3 Priority Queuing (PQ)
    - 1.7.5.4 Custom Queuing
    - 1.7.5.5 Class-Based Weighted Fair Queuing (CBWFQ)
    - 1.7.5.6 Low-Latency Queuing (LLQ)
  - 1.7.6 Null Interface
  - 1.7.7 Fast, Autonomous, and Silicon Switching
  - 1.7.8 Cisco Express Forwarding (CEF)
  - 1.7.9 Enhanced Interior Gateway Routing Protocol (EIGRP)

## **2. IP/TCP**

- 2.1 The IP Address
  - 2.1.1 IP Address Classes
  - 2.1.2 Classless Interdomain Routing (CIDR) Notation
  - 2.1.3 Subnetting
  - 2.1.4 Variable-Length Subnet Masks
- 2.2 Summarization
  - 2.2.1 Automatic Summarization
  - 2.2.2 Manual Summarization
- 2.3 Implementing Private IP Addresses
  - 2.3.1 Private IP Addressing
  - 2.3.2 Network Address Translation

2.4 The Logical AND Operation

2.5 IP Routing

- 2.5.1 Routing Protocols
- 2.5.2 The show ip route Command
- 2.5.3 The clear ip route Command

**3. Basic Switching and Network Technologies**

3.1 Network Technologies

- 3.1.1 Ethernet
  - 3.1.1.1 Ethernet Switches
  - 3.1.1.2 Ethernet Media
- 3.1.2 Cisco Long Reach Ethernet (LRE)
- 3.1.3 Fast Ethernet
- 3.1.4 Gigabit Ethernet
- 3.1.5 10Gigabit Ethernet
- 3.1.6 Token Ring

3.2 Connecting Switches

- 3.2.1 Console Port Cables and Connectors
- 3.2.2 Ethernet Port Cables and Connectors
- 3.2.3 Gigabit Ethernet Port Cables and Connectors
- 3.2.4 Token Ring Port Cables and Connectors

3.3 Switch Management

- 3.3.1 Switch Naming
- 3.3.2 Password Protection
- 3.3.3 Remote Access
- 3.3.4 Inter-Switch Communication
- 3.3.5 Switch Clustering and Stacking

3.4 Switch File Management

- 3.4.1 OS Image Files
- 3.4.2 Configuration Files
- 3.4.3 More Catalyst Switch Files
- 3.4.4 Shifting Catalyst Switch Files About

3.5 Switch Port Configuration

- 3.5.1 Port Description
- 3.5.2 Port Speed
- 3.5.3 Ethernet Port Mode
- 3.5.4 Token Ring Port Mode

**4. Routing**

4.1 Routing Tables

- 4.1.1 Static Routing

- 4.1.2 Dynamic Routing
- 4.1.3 Routing Updates
- 4.1.4 Verifying Routing Tables

#### 4.2 Routing Protocols

- 4.2.1 Distance-Vector Routing
- 4.2.2 Link-State Routing
- 4.2.3 Classful Routing
- 4.2.4 Classless Routing
- 4.2.5 Multipath Routing

#### 4.3 Basic Switching Functions

#### 4.4 Convergence

- 4.4.1 Distance-Vector Routing Convergence
  - 4.4.1.1 RIP and IGRP Convergence
  - 4.4.1.2 EIGRP Convergence
- 4.4.2 Link-State Convergence

#### 4.5 Routing and Switching in a Cisco Router

#### 4.6 The Structure of a Routing Table

#### 4.7 Testing and Troubleshooting Routes

- 4.7.1 The ping Command
- 4.7.2 The traceroute Command

### **5. OSPF in a Single Area Network**

#### 5.1 OSPF Neighbors

- 5.1.1 Adjacent OSPF Neighbors

#### 5.2 The Designated Router (DR) and the Backup Designated Router (BDR)

#### 5.3 The OSPF Routing Table

- 5.3.1 Building the Routing Table on a New OSPF Router
- 5.3.2 The Topology Database
- 5.3.3 The Shortest Path First

#### 5.4 OSPF Across Nonbroadcast Multiaccess Networks (NBMA)

#### 5.5 Problems with OSPF in a Single Area

#### 5.6 Configuring OSPF in a Single Area

- 5.6.1 Configuring OSPF on an Internal Router
  - 5.6.1.1 The router ospf Command
  - 5.6.1.2 The network Command
- 5.6.2 Configuring OSPF on the External Router

- 5.6.2.1 The interface loopback Command
- 5.6.2.2 The cost Command
- 5.6.2.3 The auto-cost Command
- 5.6.2.4 The priority Command
- 5.6.3 Configuring OSPF over an NBMA Topology
  - 5.6.3.1 Configuring OSPF in NBMA Mode
  - 5.6.3.2 Configuring OSPF in Point-to-Multipoint Mode
  - 5.6.3.3 Configuring OSPF in Broadcast Mode
  - 5.6.3.4 Configuring OSPF in Point-to-Point Mode on a Frame Relay Subinterface

5.7 Verifying the OSPF Configuration on a Single Router

5.8 Differences between OSPF and RIP Routing Protocols

## **6. OSPF in a Multiple Area Network**

6.1 Different Router Types

6.2 The Link-State Advertisements

6.3 OSPF Path Selection Between Areas

6.3.1 The Path to Another Area

6.3.2 The Path to Another AS

6.4 Different Types of Areas

6.5 Design Considerations in Multiple Area OSPF

6.5.1 Cisco Design Guidelines

6.5.2 Summarization

6.5.3 The Virtual Link

6.5.4 OSPF over an NBMA Network

6.6 Configuring OSPF in a Multiple Area Network

6.6.1 The **network** Command

6.6.2 The **area range** Command for an ABR

6.6.3 The **summary-address** Command for an ASBR

6.6.4 The **area** Command

6.6.5 Configuring a Virtual Link

6.7 Verifying the OSPF Configuration a Multiple Area Network

## **7. EIGRP in Enterprise Networks**

7.1 Operation of EIGRP

7.1.1 The Neighbor Table

7.1.2 The Topology Table

7.1.3 EIGRP Metrics

7.2 Updating the Routing Table

7.2.1 Updating the Routing Table in Passive Mode

7.2.2 Updating the Routing Table in Active Mode

7.2.3 Adding a Network to the Topology Table

7.2.4 Removing a Path or Router from the Topology Table

7.3 Scaling EIGRP

7.4 Configuring EIGRP

7.5 Verifying the EIGRP Operation

**8. Using BGP-4 to Communicate with Other Autonomous Systems**

8.1 BGP-4 Overview

8.1.1 The BGP-4 Operation

8.1.2 Types of BGP-4

8.1.3 BGP-4 Synchronization

8.1.4 BGP-4 Policy-Based Routing

8.1.5 BGP-4 Attributes

8.2 Basic BGP-4 Configuration Commands

8.2.1 Starting the Routing Process

8.2.2 Defining the Networks to Be Advertised

8.2.3 Identifying Neighbors and Defining Peer Groups

8.2.4 Forcing the Next-Hop Address

8.2.5 Disabling Synchronization

8.2.6 Aggregating Routes

8.3 Effecting BGP-4 Configuration Changes

8.4 Verifying the Basic BGP-4 Configuration

8.5 Advanced BGP-4 Configuration

8.5.1 Configuring Route Reflectors

8.5.2 Controlling BGP-4 Traffic

8.5.3 Redundant Connections into the Internet

8.5.4 Determining the BGP-4 Path by Configuring the Attributes

8.6 Verifying the Advanced BGP-4 Configuration

**9. Using Integrated IS-IS in Connectionless Networks**

9.1. IS-IS Overview

9.1.1 The OSI Connectionless Network Service (CLNS)

9.1.2 Integrated IS-IS

9.2 IS-IS Operations

9.2.1 IS-IS Data-Flow Diagram

9.2.2 Adjacency Building

9.2.3 The Link-State Database and Reliable Flooding

9.2.4 DIS and Pseudonodes

9.2.5 IS-IS Metrics

9.3 IS-IS Routing

9.3.1 IP Routing with IS-IS

9.4 Security

9.5 Configuring Integrated IS-IS

9.5.1 Enabling IS-IS and Assigning Areas

9.5.2 Enabling IP Routing for an Area on an Interface

9.5.3 Configuring Optional Interface Parameters

9.5.4 Configuring IS-IS Authentication Passwords

9.5.5 Monitoring IS-IS

**10. Controlling Routing Updates Across the Network**

10.1 Features of Redistribution

10.2 Problems of Configuring Multiple Routing Protocols

10.2.1 Path Selection

10.2.2 Routing Loops

10.2.3 Redistribution and Network Convergence

10.3 Configuring Redistribution

10.4 The Default or Seed Metric

10.4.1 Configuring the Default Metric for OSPF, RIP, EGP or BGP-4

10.4.2 Configuration for EIGRP or IGRP

10.5 Configure the Administrative Distance

10.5.1 Configuring the Administrative Distance in EIGRP

10.5.2 Configuring the Administrative Distance in Other Protocols

10.6 The Passive Interface

10.7 Static Routes

10.8 Controlling Routing Updates with Filtering

10.9 Policy-Based Routing Using Route Maps

10.10 Managing the Redistribution

- 10.10.1 Trouble shooting Redistribution
- 10.10.2 Monitoring Policy-Routing Configurations

## **11. Virtual LANs (VLANs) and Trunking**

- 11.1 VLAN Membership
- 11.2 Extent of VLANs
- 11.3 VLAN Trunks
  - 3.3.1 VLAN Frame Identification
  - 3.3.2 Dynamic Trunking Protocol
  - 3.3.3 VLAN Trunk Configuration
- 11.4 Service Provider Tunneling
  - 3.4.1 IEEE 802.1Q Tunnels
  - 3.4.2 Layer 2 Protocol Tunnels
  - 3.4.3 Ethernet Over Multiprotocol Label Switching (MPLS) Tunneling
- 11.5 VLAN Trunking Protocol (VTP)
  - 11.5.1 VTP Modes
    - 11.5.1.1 Server Mode
    - 11.5.1.2 Client Mode
    - 11.5.1.3 Transparent Mode
  - 11.5.2 VTP Advertisements
    - 11.5.2.1 Summary Advertisements
    - 11.5.2.2 Subset Advertisements
    - 11.5.2.3 Client Request Advertisements
  - 11.5.3 VTP Configuration
    - 11.5.3.1 Configuring a VTP Management Domain
    - 11.5.3.2 Configuring the VTP Mode
    - 11.5.3.3 Configuring the VTP Version
  - 11.5.4 VTP Pruning
- 11.6 Token Ring VLANs
  - 11.6.1 TrBRF
  - 11.6.2 TrCRF
  - 11.6.3 VTP and Token Ring VLANs
  - 11.6.4 Duplicate Ring Protocol (DRiP)

## **12. Redundant Switch Links**

- 12.1 Switch Port Aggregation with EtherChannel
  - 12.1.1 Bundling Ports with EtherChannel
  - 12.1.2 Distributing Traffic in EtherChannel
  - 12.1.3 Port Aggregation Protocol (PAgP)
  - 12.1.4 Link Aggregation Control Protocol (LACP)
  - 12.1.5 EtherChannel Configuration

- 12.2 Spanning-Tree Protocol (STP)
- 12.3 Spanning-Tree Communication
  - 12.3.1 Root Bridge Election
  - 12.3.2 Root Ports Election
  - 12.3.3 Designated Ports Election
- 12.4 STP States
- 12.5 STP Timers
- 12.6 Convergence
  - 12.6.1 PortFast: Access Layer Nodes
  - 12.6.2 UplinkFast: Access Layer Uplinks
  - 12.6.3 BackboneFast: Redundant Backbone Paths
- 12.7 Spanning-Tree Design
- 12.8 STP Types
  - 12.8.1 Common Spanning Tree (CST)
  - 12.8.2 Per-VLAN Spanning Tree (PVST)
  - 12.8.3 Per-VLAN Spanning Tree Plus (PVST+)
- 12.9 Protecting Against Unforeseen Bridge Protocol Data Units (BPDU)
  - 12.9.1 The Root Guard Feature
  - 12.9.2 The BPDU Guard Feature
- 12.10 Protecting Against the Sudden Loss of BPDUs
  - 12.10.1 BPDU Skew Detection Feature
  - 12.10.2 Loop Guard Feature
  - 12.10.3 Unidirectional link detection (UDLD) STP Feature
- 12.11 Advanced Spanning-Tree Protocol
  - 12.11.1 Rapid Spanning Tree Protocol (RSTP)
    - 12.11.1.1 RSTP Port Performance
    - 12.11.1.2 BPDUs and RSTP
    - 12.11.1.3. RSTP Convergence
    - 12.11.1.4. RSTP and Topology Changes
    - 12.11.1.5 Configuring RSTP
  - 12.11.2 The Multiple Spanning Tree Protocol (MSTP or MST)
    - 12.11.2.1 MST Regions
    - 12.11.2.2 Spanning Tree Instances in MST
    - 12.11.2.3 Configuring MST

## **13. Trunking with ATM LAN Emulation (LANE)**

- 13.1 ATM

- 13.1.1 The ATM Model
- 13.1.2 Virtual Circuits
- 13.1.3 ATM Addressing
  - 13.1.3.1 VPI/VCI Addresses
  - 13.1.3.2 NSAP Addresses
- 13.1.4 ATM Protocols

- 13.2 LAN Emulation (LANE)
  - 13.2.1 LANE Components
  - 13.2.2 LANE Operation
  - 13.2.3 Address Resolution
  - 13.2.4 LANE Component Placement
  - 13.2.5 LANE Component Redundancy (SSRP)

- 13.3 LANE Configuration
  - 13.3.1 Configuring the LES and BUS
  - 13.3.2 Configuring the LECS
  - 13.3.3 Configuring Each LEC
  - 13.3.4 Viewing the LANE Configuration

## **14. InterVLAN Routing**

- 14.1 InterVLAN Routing Design
  - 14.1.1 Routing with Multiple Physical Links
  - 14.1.2 Routing over Trunk Links
    - 14.1.2.1 802.1Q and ISL Trunks
    - 14.1.2.2 ATM LANE
- 14.2 Routing with an Integrated Router
- 14.3 InterVLAN Routing Configuration
  - 14.3.1 Accessing the Route Processor
  - 14.3.2 Establishing VLAN Connectivity
    - 14.3.2.1 Establishing VLAN Connectivity with Physical Interfaces
    - 14.3.2.2 Establishing VLAN Connectivity with Trunk Links
    - 14.3.2.3 Establishing VLAN Connectivity with LANE
    - 14.3.2.4 Establishing VLAN Connectivity with Integrated Routing Processors
  - 14.3.3 Configure Routing Processes
  - 14.3.4 Additional InterVLAN Routing Configurations

## **15. Multilayer Switching (MLS)**

- 15.1 Multilayer Switching Components
- 15.2 MLS-RP Advertisements
- 15.3 Configuring Multilayer Switching

15.4 Flow Masks

15.5 Configuring the MLS-SE

- 15.5.1 MLS Caching
- 15.5.2 Verifying MLS Configurations
- 15.5.3 External Router Support
- 15.5.4 Switch Inclusion Lists
- 15.5.5 Displaying MLS Cache Entries

## **16. Cisco Express Forwarding (CEF)**

16.1 CEF Components

- 16.1.1 Forwarding Information Base (FIB)
- 16.1.2 Adjacency Tables

16.2 CEF Operation Modes

16.3 Configuring Cisco Express Forwarding

- 16.3.1 Configuring Load Balancing for CEF
  - 16.3.1.1 Per-Destination Load Balancing
  - 16.3.1.2 Per-Packet Load Balancing
- 16.3.2 Configuring Network Accounting for CEF

## **17. The Hot Standby Router Protocol (HSRP)**

17.1 Traditional Redundancy Methods

- 17.1.1 Default Gateways
- 17.1.2 Proxy ARP
- 17.1.3 Routing Information Protocol (RIP)
- 17.1.4 ICMP Router Discovery Protocol (IRDP)

17.2 Hot Standby Router Protocol

- 17.2.1 HSRP Group Members
- 17.2.2 Addressing HSRP Groups Across ISL Links

17.3 HSRP Operations

- 17.3.1 The Active Router
- 17.3.2 Locating the Virtual Router MAC Address
- 17.3.3 Standby Router Behavior
- 17.3.4 HSRP Messages
- 17.3.5 HSRP States

17.4 Configuring HSRP

- 17.4.1 Configuring an HSRP Standby Interface
- 17.4.2 Configuring HSRP Standby Priority
- 17.4.3 Configuring HSRP Standby Preempt
- 17.4.4 Configuring the Hello Message Timers
- 17.4.5 HSRP Interface Tracking

17.4.6 Configuring HSRP Tracking

17.4.7 HSRP Status

17.5 Troubleshooting HSRP

## **18. Multicasts**

18.1 Unicast Traffic

18.2 Broadcast Traffic

18.3 Multicast Traffic

18.4 Multicast Addressing

18.4.1 Multicast Address Structure

18.4.2 Mapping IP Multicast Addresses to Ethernet

18.4.3 Managing Multicast Traffic

18.4.4 Subscribing and Maintaining Groups

18.4.4.1 IGMP Version 1

18.4.4.2 IGMP Version 2

18.4.5 Switching Multicast Traffic

18.5 Routing Multicast Traffic

18.5.1 Distribution Trees

18.5.2 Multicast Routing Protocols

18.5.2.1 Dense Mode Routing Protocols

18.5.2.2 Sparse Mode Routing Protocols

18.6 Configuring IP Multicast

18.6.1 Enabling IP Multicast Routing

18.6.2 Enabling PIM on an Interface

18.6.2.1 Enabling PIM in Dense Mode

18.6.2.2 Enabling PIM in Sparse Mode

18.6.2.3 Enabling PIM in Sparse-Dense Mode

18.6.2.4 Selecting a Designated Router

18.6.3 Configuring a Rendezvous Point

18.6.4 Configuring Time-To-Live

18.6.5 Debugging Multicast

18.6.6 Configuring Internet Group Management Protocol (IGMP)

18.6.7 Configuring Cisco Group Management Protocol (CGMP)

## **19. Quality of Service**

19.1 Understanding the Need for Quality of Service

19.2 QoS Types

19.2.1 Best Efforts Delivery

19.2.2 Integrated Services Model

19.2.3 Differentiated Services Model

19.3 Differentiated Services QoS

19.3.1 IEEE 802.1p

19.3.2 Using the QoS Model

19.3.3 Prioritizing the Traffic Classes

19.3.4 Queuing Methods

19.3.4.1. Auto-QoS

19.4 Configuring QoS

19.4.1 Per-interface QoS Trust

19.4.2 Defining a QoS Policy

19.4.3 Configuring and Tuning Egress Scheduling

19.4.4 Congestion Prevention

## **20. IP Telephony**

201 Inline Power

201.1 Inline Power Configuration and Verification

202 Voice VLANs

202.1 Voice VLANs Configuration and Verification

203 Voice QoS

203.1 QoS Trust

203.1.1 QoS Trust Configuration and Verification

203.2 Voice Packet Classification

## **21. Controlling Access in the Campus Environment**

21.1 Access Policies

21.2 Managing Network Devices

21.2.1 Physical Access

21.2.2 Passwords

21.2.3 Privilege Levels

21.2.4 Virtual Terminal Access

21.3 Access Layer Policy

21.4 Distribution Layer Policy

21.4.1 Filtering Traffic at the Distribution Layer

21.4.2 Controlling Routing Update Traffic

21.4.3 Configuring Route Filtering

21.5 Core Layer Policy

## **22. Monitoring and Troubleshooting**

## **CCNP/CCDP 642-891 (Composite)**

### 22.1 Monitoring Cisco Switches

#### 22.1.1 Out-of-Band Management

##### 22.1.1.1 Console Port Connection

##### 22.1.1.2 Serial Line Internet Protocol (SLIP)

#### 22.1.2 In-Band Management

##### 22.1.2.1 SNMP

##### 22.1.2.2 Telnet Client Access

##### 22.1.2.3 Cisco Discovery Protocol (CDP)

#### 22.1.3 Embedded Remote Monitoring

#### 22.1.4 Switched Port Analyzer

#### 22.1.5 CiscoWorks 2000

### 22.2 General Troubleshooting Model

#### 22.2.1 Troubleshooting with show Commands

#### 22.2.2 Physical Layer Troubleshooting

#### 22.2.3 Troubleshooting Ethernet

##### 22.2.3.1 Network Testing

##### 22.2.3.2 The Traceroute Command

##### 22.2.3.3 Network Media Test Equipment

## **LIST OF TABLES**

TABLE 1.1:	Network Service Types
TABLE 1.2:	OSI Encapsulation
TABLE 1.3:	Differences between Switches and Bridges
TABLE 1.4:	Parameters for the Extended access-list Command
TABLE 2.1:	Private IP Address Ranges
TABLE 2.2:	The Metrics used by Different Routing Protocols
TABLE 3.1:	Coaxial Cable for Ethernet
TABLE 3.2:	Twisted-Pair and Fiber Optic Cable for Ethernet
TABLE 3.3:	Fast Ethernet Cabling and Distance Limitations
TABLE 3.4:	Gigabit Ethernet Cabling and Distance Limitations
TABLE 3.5:	Catalyst Switch File Locations
TABLE 3.6:	File Management Commands
TABLE 4.1:	Parameters for the ping Command
TABLE 4.2:	Parameters for the traceroute Command
TABLE 5.1:	OSPF Terminology
TABLE 5.2:	The Default Hello and Dead Time Intervals
TABLE 5.3:	Default Costs in OSPF
TABLE 7.1:	EIGRP Terminology
TABLE 8.1:	The Categories of BGP-4 Attributes
TABLE 8.2:	The BGP-4 Attributes supported by Cisco
TABLE 10.1:	The Default Administrative Distance
TABLE 12.1:	MST Configuration Commands
TABLE 13.1:	Automatic NSAP Address Generation for LANE Components
TABLE 15.1:	Displaying Specific MLS Cache Entries
TABLE 16.1:	Adjacency Types for Exception Processing
TABLE 18.1:	Well-Known Class D Addresses
TABLE 19.1:	Differentiated Services Types of Traffic
TABLE 21.1:	Access Policy Guidelines
TABLE 22.1:	Keywords and Arguments for the set snmp trap Command
TABLE 22.2:	CiscoWorks 2000 LAN Management Features
TABLE 22.3:	Ethernet Media Problems
TABLE 22.4:	Parameters for the ping Command
TABLE 22.5:	Parameters for the traceroute Command

## **LIST OF ACRONYMS**

AAA	Authentication, Authorization, and Accounting
ABR	Area Border Router
ACF	Advanced Communications Function
ACK	Acknowledgment bit (in a TCP segment)
ACL	Access Control List
ACS	Access Control Server
AD	Advertised Distance
ADSL	Asymmetric Digital Subscriber Line
ANSI	American National Standards Institute
API	Application Programming Interface
APPC	Advanced Program-to-Program Communications
ARAP	AppleTalk Remote Access Protocol
ARE	All Routes Explorer
ARP	Address Resolution Protocol
ARPA	Advanced Research Projects Agency
ARPANET	Advanced Research Projects Agency Network
AS	Autonomous System
ASA	Adaptive Security Algorithm
ASBR	Autonomous System Boundary Router
ASCII	American Standard Code for Information Interchange
ASIC	Application Specific Integrated Circuits
ATM	Asynchronous Transfer Mode
AUI	Attachment Unit Interface
Bc	Committed burst (Frame Relay)
B channel	Bearer channel ( ISDN)
BDR	Backup Designated Router
Be	Excess burst (Frame Relay)
BECN	Backward Explicit Congestion Notification (Frame Relay)
BGP	Border Gateway Protocol
BGP-4	BGP version 4
BIA	Burned-in Address (another name for a MAC address)

## CCNP/CCDP 642-891 (Composite)

BOD	Bandwidth on Demand.
BPDU	Bridge Protocol Data Unit
BRF	Bridge Relay Function
BRI	Basic Rate Interface (ISDN)
BSD	Berkeley Standard Distribution (UNIX)
CBT	Core Based Trees
CBWFQ	Class-Based Weighted Fair Queuing
CCITT	Consultative Committee for International Telegraph and Telephone
CCO	Cisco Connection Online
CDDI	Copper Distribution Data Interface
CDP	Cisco Discovery Protocol
CEF	Cisco Express Forwarding
CHAP	Challenge Handshake Authentication Protocol
CIDR	Classless Interdomain Routing
CIR	Committed Information Rate. (Frame Relay)
CLI	Command-Line Interface
CLNP	Connectionless Network Protocol (IS-IS)
CLNS	Connectionless Network Service (IS-IS)
CPE	Customer Premises Equipment
CPU	Central Processing Unit
CR	Carriage Return.
CRC	Cyclic Redundancy Check (error)
CRF	Concentrator Relay Function
CSNP	Complete Sequence Number PDU (IS-IS)
CST	Common Spanning Tree
CSU	Channel Service Unit
DB	Data Bus (connector)
DCE	Data Circuit-Terminating Equipment
dCEF	Distributed CEF
DDR	Dial-on-Demand Routing
DE	Discard Eligible Indicator
DECnet	Digital Equipment Corporation Protocols
DES	Data Encryption Standard
DHCP	Dynamic Host Control Protocol

## CCNP/CCDP 642-891 (Composite)

DIS	Designated Intermediate System (IS-IS)
DLCI	Data-Link Connection Identifier
DNIC	Data Network Identification Code. (X.121 addressing)
DNS	Domain Name System
DoD	Department of Defense (US)
DRiP	Duplicate Ring Protocol
DR	Designated Router
DS	Digital Signal
DS0	Digital Signal level 0
DS1	Digital Signal level 1
DS3	Digital Signal level 3
DSL	Digital Subscriber Line
DSU	Data Service Unit
DTE	Data Terminal Equipment
DTP	Dynamic Trunking Protocol
DUAL	Diffusing Update Algorithm
DVMRP	Distance Vector Multicast Routing Protocol
EBC	Ethernet Bundling Controller
EGP	Exterior Gateway Protocol
EIA/TIA	Electronic Industries Association/Telecommunications Industry Association
EIGRP	Enhanced IGRP
ES	End System (IS-IS)
ES-IS	End System-to-Intermediate System Protocol (IS-IS)
ESH	End System Hello message (IS-IS)
FCC	Federal Communications Commission
FCS	Frame Check Sequence
FC	Feasible Condition (Routing)
FD	Feasible Distance (Routing)
FDDI	Fiber Distributed Data Interface
FEC	Fast EtherChannel
FECN	Forward Explicit Congestion Notification
FIB	Forwarding Information Base
FIFO	First-In, First-Out (Queuing)
FLSM	Fixed-Length Subnet Masks

## CCNP/CCDP 642-891 (Composite)

FR	Frame Relay
FS	Feasible Successor (Routing)
FSSRP	Fast Simple Server Redundancy Protocol
FTP	File Transfer Protocol
GBIC	Gigabit Interface Converters
GBPT	Generic Bridge PDU Tunneling
GEC	Gigabit EtherChannel
GSR	Gigabit Switch Router
HDLC	High-Level Data Link Control
HDSL	High data-rate digital subscriber line
HSRP	Hot Standby Router Protocol
HSSI	High-Speed Serial Interface
HTTP	Hypertext Transfer Protocol
I/O	Input/Output
IANA	Internet Assigned Numbers Authority
ICMP	Internet Control Message Protocol
IDRP	Interdomain Routing Protocol
IDN	International Data Number
IEEE	Institute of Electrical and Electronic Engineers
IETF	Internet Engineering Task Force
IGP	Interior Gateway Protocol
IGRP	Interior Gateway Routing Protocol
ILMI	Integrated Local Management Interface
IOS	Internetwork Operating System
IP	Internet Protocol
IPSec	IP Security
IPv6	IP version 6
IPX	Internetwork Packet Exchange (Novell)
IRDP	ICMP Router Discovery Protocol
IS	Information Systems
	<i>or</i>
	Intermediate System
IS-IS	Intermediate System-to-Intermediate System.
ISDN	Integrated Services Digital Network

## CCNP/CCDP 642-891 (Composite)

ISH	Intermediate System Hello message (IS-IS)
ISO	International Organization for Standardization
ISOC	Internet Society
ISP	Internet Service Provider
IST	Internal Spanning Tree
ITU-T	International Telecommunication Union–Telecommunication Standardization Sector
kbps	kilobits per second (bandwidth)
LACP	Link Aggregation Control Protocol
LAN	Local Area Network
LANE	LAN Emulation
LAPB	Link Access Procedure, Balanced
LAPD	Link Access Procedure on the D channel
LEC	LAN Emulation Client
LECS	LAN Emulation Configuration Server
LED	Light Emitting Diode
LES	LAN Emulation Server
LLC	Logic Link Control (OSI Layer 2 sublayer)
LLQ	Low-Latency Queuing
LMI	Local Management Interface
LSA	Link-State Advertisement
LSP	Link-State PDU
MAC	Media Access Control (OSI Layer 2 sublayer)
MAN	Metropolitan-Area Network
MD5	Message Digest Algorithm 5
MLS	Multilayer Switching
MLS-RP	Multilayer Switching Route Processor
MLS-SE	Multilayer Switching Switch Engine
MLSP	Multilayer Switching Protocol
MOSPF	Multicast Open Shortest Path First
MPLS	Multiprotocol Label Switching
MSAU	Multistation Access Unit
MSFC	Multilayer Switch Feature Card
MST	Multiple Spanning Tree
MTU	Maximum Transmission Unit

## CCNP/CCDP 642-891 (Composite)

NAK	Negative Acknowledgment
NAS	Network Access Server
NAT	Network Address Translation
NBMA	Nonbroadcast Multiaccess
NetBEUI	NetBIOS Extended User Interface
NetBIOS	Network Basic Input/Output System
NFFC	NetFlow Feature Card
NMS	Network Management System
NNI	Network-to-Network Interface
NPDU	Network Protocol Data Unit
NVRAM	Nonvolatile Random Access Memory
OC	Optical Carrier
ODBC	Open Database Connectivity
OLE	Object Linking and Embedding
OSI	Open Systems Interconnection (Model)
	<i>or</i>
	Open System Interconnection (IS-IS)
OSPF	Open Shortest Path First
OTDR	Optical Time Domain Reflectometer
OUI	Organizationally Unique Identifier
PAgP	Port Aggregation Protocol
PAP	Password Authentication Protocol
PAT	Port Address Translation
PDN	Public Data Network
PDU	Protocol Data Unit (i.e., a data packet)
PIM	Protocol Independent Multicast
PIM	SM Protocol Independent Multicast Sparse Mode
PIMDM	Protocol Independent Multicast Mode
PIX	Private Internet Exchange (Cisco Firewall)
PNNI	Private Network-to-Network Interface
POP	Point of Presence
POTS	Plain Old Telephone Service
PPP	Point-to-Point Protocol
PQ	Priority Queuing

## CCNP/CCDP 642-891 (Composite)

PRI	Primary Rate Interface (ISDN)
PSNP	Partial Sequence Number PDU (IS-IS)
PSTN	Public Switched Telephone Network
PTT	Poste, Telephone, Telegramme
PVC	Permanent Virtual Circuit (ATM)
PVST	Per-VLAN Spanning Tree
PVST+	Per-VLAN Spanning Tree Plus
QoS	Quality of Service
RADIUS	Remote Authentication Dial-In User Service
RAS	Remote Access Service
RIF	Routing Information Field
RIP	Routing Information Protocol
RJ	Registered Jack (connector)
RMON	Embedded Remote Monitoring
RP	Rendezvous Point
RPF	Reverse Path Forwarding
RSFC	Route Switch Feature Card
RSM	Route Switch Module
RSP	Route Switch Processor
RSTP	Rapid Spanning-Tree Protocol
RTP	Reliable Transport Protocol
RTO	Retransmission Timeout
SA	Source Address
SAID	Security Association Identifier
SAP	Service Access Point
	<i>or</i>
	Service Advertising Protocol (Novell)
SAPI	Service Access Point Identifier.
SAR	Segmentation and Reassembly
SDLC	Synchronous Data Link Control (SNA)
SIA	Stuck in Active (EIGRP)
SIN	Ships-in-the-Night (Routing)
SLIP	Serial Line Internet Protocol
SMDS	Switched Multimegabit Data Service

## CCNP/CCDP 642-891 (Composite)

SMTP	Simple Mail Transfer Protocol
SNA	Systems Network Architecture (IBM)
SNAP	SubNetwork Access Protocol
SNMP	Simple Network Management Protocol
SOF	Start of Frame
SOHO	Small Office, Home Office
SONET	Synchronous Optical Network
SONET/SDH	Synchronous Optical Network/Synchronous Digital Hierarchy
SPAN	Switched Port Analyzer
SPF	Shortest Path First
SPID	Service Profile Identifier
SPP	Sequenced Packet Protocol (Vines)
SPT	Shortest Path Tree (IS-IS)
SPX	Sequenced Packet Exchange (Novell)
SQL	Structured Query Language
SRAM	Static RAM
SRB	Source-Route Bridge
SRT	Source-Route Transparent (Bridging)
SRTT	Smooth Round-Trip Timer (EIGRP)
SS7	Signaling System 7
SSAP	Source service access point (LLC)
SSE	Silicon Switching Engine.
SSP	Silicon Switch Processor
SSRP	Simple Server Redundancy Protocol
STA	Spanning-Tree Algorithm
STP	Spanning-Tree Protocol; also Shielded Twisted-Pair (cable)
SVC	Switched Virtual Circuit (ATM)
SYN	Synchronize (TCP segment)
TA	Terminal Adapter (ISDN)
TAC	Technical Assistance Center (Cisco)
TACACS	Terminal Access Controller Access Control System
TCI	Tag Control Information
TCP	Transmission Control Protocol
TCP/IP	Transmission Control Protocol/Internet Protocol

## CCNP/CCDP 642-891 (Composite)

TCN	Topology Change Notification
TDM	Time-Division Multiplexing
TDR	Time Domain Reflectometers
TFTP	Trivial File Transfer Protocol
TIA	Telecommunications Industry Association
TLV	Type-Length-Value
ToS	Type of Service
TPID	Tag Protocol Identifier
TrBRF	Token Ring Bridge Relay Function
TrCRF	Token Ring Concentrator Relay Function
TTL	Time To Live
UDLD	Unidirectional Link Detection
UDP	User Datagram Protocol
UNC	Universal Naming Convention or Uniform Naming Convention
UNI	User-Network Interface
URL	Uniform Resource Locator
UTC	Coordinated Universal Time (same as Greenwich Mean Time)
UTL	Utilization
UTP	Unshielded Twisted-Pair (cable)
VBR	Variable Bit Rate
VC	Virtual Circuit (ATM)
VID	VLAN Identifier
VIP	Versatile Interface Processor
VLAN	Virtual LAN
VLSM	Variable-Length Subnet Mask
VMPS	VLAN Membership Policy Server
VPN	Virtual Private Network
VTP	VLAN Trunking Protocol
vt	Virtual terminal line
WAIS	Wide Area Information Server
WAN	Wide Area Network
WFQ	Weighted Fair Queuing
WWW	World Wide Web
XNS	Xerox Network Systems

**CCNP/CCDP 642-891 (Composite)**

XOR	Exclusive-OR
XOT	X.25 over TCP
ZIP	Zone Information Protocol (AppleTalk)

# Building Scalable Cisco Internetworks (BSCI) and Building Cisco Multilayer Switched Networks (BCMSN)

**Exam Code: 642-891**

## **Certifications:**

<b>Cisco Certified Internetwork Professional (CCIP)</b>	<b>Core</b>
<b>Cisco Certified Network Professional (CCNP)</b>	<b>Core and Recertification</b>
<b>Cisco Certified Design Professional (CCDP)</b>	<b>Core and Recertification</b>

## **Prerequisites:**

Cisco CCNA 640-801 – Cisco Certified Network Associate, or  
Cisco CCNA 640-811 – Interconnecting Cisco Network Devices (ICND) AND  
Cisco CCNA 640-821 – Introduction to Cisco Network Devices (INTRO), or  
Cisco CCDA 640-861 – Designing Cisco Internetwork Solutions.

## **About This Study Guide**

This Study Guide is based on the current pool of exam questions for the 642-891 exam. As such it provides all the information required to pass the Cisco 642-801 exam and is organized around the specific skills that are tested in that exam. Thus, the information contained in this Study Guide is specific to the 642-891 exam and does not represent a complete reference work on the subject of Building Scalable Cisco Internetworks. Topics covered in this Study Guide includes: List the key information routers needs to route data; Describe classful and classless routing protocols; Describe link-state router protocol operation; Compare classful and classless routing protocols; Compare distance vector and link state routing protocols; Describe concepts relating to extending IP addresses and the use of VLSMs to extend IP addresses; Describe the features and operation of EIGRP; Describe the features and operation of single area OSPF; Describe the features and operation of multi-area OSPF; Explain basic OSI terminology and network layer protocols used in OSI; Identify similarities and differences between Integrated IS-IS and OSPF; List the types of IS-IS routers and their role in IS-IS area design; Describe the hierarchical structure of IS-IS areas; Describe the concept of establishing adjacencies; Describe the features and operation of BGP; Explain how BGP policy-based routing functions within an autonomous system; Explain the use of redistribution between BGP and Interior Gateway Protocols (IGPs); Implementation and Configuration of OSPF in a single-area an in a multiple area network, Enhanced IGRP, and BGP and verifying their proper operation; Identifying the steps to select and configure the different ways to control routing update traffic; Identifying the steps to configure router redistribution in a network; Identifying the steps to configure policy-based routing using route maps; Identifying the steps to configure a router for Network Address Translation with overload, static translations, and route maps; Describing the three-layer hierarchical design model and explain the function of each layer; Choosing the correct routing protocol to meet the requirements; Identifying the correct IP addressing scheme; Describing the concepts relating to route summarization and apply them to hypothetical scenarios; Troubleshooting the OSPF operation in a single area, the OSPF operation in multiple areas, the Enhanced

IGRP operation, and the BGP operation; Identifying verification methods which ensure proper operation of Integrated IS-IS on Cisco routers; Identifying the steps to verify route redistribution; Describing the scalability problems associated with internal BGP; Interpreting the output of various show and debug commands to determine the cause of route selection errors and configuration problems; Describing the functionality of CGMP, Enabling CGMP on the distribution layer devices, Identifying the correct Cisco Systems product solution given a set of network switching requirements; Describing how switches facilitate Multicast Traffic; Translating Multicast Addresses into MAC addresses; Identifying the components necessary to effect multilayer switching; Applying flow masks to influence the type of MLS cache; Describing layer 2, 3, 4 and multilayer switching; Verifying existing flow entries in the MLS cache; Describing how MLS functions on a switch; Configuring a switch to participate in multilayer switching; Describing Spanning Tree; Configuring the switch devices to improve Spanning Tree Convergence in the network; Identifying Cisco Enhancement that improve Spanning Tree Convergence; Configuring a switch device to Distribute Traffic on Parallel Links; Providing physical connectivity between two devices within a switch block; Providing connectivity from an end user station to an access layer device; Providing connectivity between two network devices; Configuring a switch for initial operation; Applying IOS command set to diagnose and troubleshoot a switched network problems; Describing the different Trunking Protocols; Configuring Trunking on a switch; Maintaining VLAN configuration consistency in a switched network; Configuring the VLAN Trunking Protocol; Describing the VTP Trunking Protocol; Describing LAN segmentation using switches; Configuring a VLAN; Ensuring broadcast domain integrity by establishing VLANs; Facilitating InterVLAN Routing in a network containing both switches and routers; and Identify the network devices required to effect InterVLAN routing; Quality of Service; and IP Telephony.

### Intended Audience

This Study Guide is targeted specifically at people who wish to take the Cisco 642-891 exam. This information in this Study Guide is specific to the exam. It is not a complete reference work. Although our Study Guides are aimed at new comers to the world of IT, the concepts dealt with in this Study Guide are complex and require an understanding of material provided for the Cisco CCNA 640-801 - Routing and Switching Certification Exam or the Cisco CCDA 640-861 - Designing for Cisco Internetwork Solutions Exam. Knowledge of CompTIA's Network+ course would also be advantageous.

**Note:** There is a fair amount of overlap between this Study Guide, and the 640-801, the 642-801, 642-811 and 642-831 Study Guides. We would, however not advise skimming over the information that seems familiar as this Study Guide expands on the information in the 642-831 and 640-801 Study Guides.

### How To Use This Study Guide

To benefit from this Study Guide we recommend that you:

- Although there is a fair amount of overlap between this Study Guide and the 640-801 Study Guide, the 642-801 Study Guide, the 642-811 Study Guide and the 642-831 Study Guide, the relevant information from those Study Guides is included in this Study Guide. This is thus the only Study Guide you will require to pass the 642-801 exam.
- Study each chapter carefully until you fully understand the information. This will require regular and disciplined work. Where possible, attempt to implement the information in a lab setup.

## CCNP/CCDP 642-891 (Composite)

- Be sure that you have studied and understand the entire Study Guide before you take the exam.

**Note:** Remember to pay special attention to these note boxes as they contain important additional information that is specific to the exam.

**Note:** A large portion of the 642-891 exam is based on IP addressing. For this reason, you must thoroughly understand IP addressing. IP addressing is discussed in detail in [Section 2](#) of this Study Guide.

Good luck!

## **1. The Campus Network**

A campus network is a building or group of buildings that connects to one network that is typically owned by one company. This local area network (LAN) typically uses Ethernet, 802.11 wireless LANs, Fast Ethernet, Fast EtherChannel, Gigabit Ethernet LANs, Token Ring, Fiber Distributed Data Interface (FDDI), or Asynchronous Transfer Mode (ATM) technologies. The task for network administrators is to ensure that the campus network run effectively and efficiently. This requires an understanding current and new emerging campus networks and equipment such as Cisco switches, which can be used to maximize network performance. Understanding how to design for the emerging campus networks is critical for implementing production networks.

### **1.1 The Traditional Shared Campus Network**

In the 1990s, the traditional campus network started as one LAN and grew until segmentation needed to take place to keep the network up and running. In this era of rapid expansion, response time was secondary to ensure the network functionality. Typical campus networks ran on 10BaseT or 10Base2, which was prone to collisions, and were, in effect, collision domains. Ethernet was used because it was scalable, effective, and comparatively inexpensive. Because a campus network can easily span many buildings, bridges were used to connect the buildings together. As more users were attached to the hubs used in the Ethernet network, performance of the network became extremely slow.

Availability and performance are the major problems with traditional campus networks. Bandwidth helps compound these problems. The three performance problems in traditional campus networks were:

#### **1.1.1 Collisions**

Because all devices could see each other, they could also collide with each other. If a host had to broadcast, then all other devices had to listen, even though they themselves were trying to transmit. And if a device were to malfunction, it could bring the entire network down. Bridges were used to break these networks into subnetworks, but broadcast problems remained. Bridges also solved distance-limitation problems because they usually had repeater functions built into the electronics.

#### **1.1.2 Bandwidth**

The bandwidth of a segment is measured by the amount of data that can be transmitted at any given time. However, the amount of data that can be transmitted at any given time is dependent on the medium, i.e. its carrier line: on its quality and length. All lines suffer from attenuation, which is the progressive degradation of the signal as it travels along the line and is due to energy loss and energy abortion. For the remote end to understand digital signaling, the signal must stay above a critical value. If it drops below this critical, the remote end will not be able to receive the data. The solution to bandwidth issues is maintaining the distance limitations and designing the network with proper segmentation of switches and routers.

Another problem is congestion, which happens on a segment when too many devices are trying to use the same bandwidth. By properly segmenting the network, you can eliminate some of these bandwidth issues.

#### **1.1.3 Broadcasts and Multicasts**

All protocols have broadcasts built in as a feature, but some protocols, such as Internet Protocol (IP), Address Resolution Protocol (ARP), Network Basic Input Output System (NetBIOS), Internetworking Packet eXchange (IPX), Service Advertising Protocol (SAP), and Routing Information Protocol (RIP), need to be configured correctly. However, there are features, such as packet filtering and queuing, that are built into the Cisco router Internetworking Operating System (IOS) that, if correctly designed and implemented, can alleviate these problems.

Multicasts are broadcasts that are destined for a specific or defined group of users. If you have large multicast groups or a bandwidth-intensive application, such as Cisco's IPTV application, multicast traffic can consume most of the network bandwidth and resources.

To solve broadcast issues, create network segmentation with bridges, routers, and switches. Another solution is Virtual LANs (VLANs). A VLAN is a group of devices on different network segments defined as a broadcast domain by the network administrator. The benefit of VLANs is that physical location is no longer a factor for determining the port into which you would plug a device into the network. You can plug a device into any switch port, and the network administrator gives that port a VLAN assignment. However, routers or layer 3 switches must be used for different VLANs to communicate. VLANs are discussed in more detail in [Section 11](#).

### 1.2 The New Campus Network

The problems with collision, bandwidth, and broadcasts, together with the changes in customer network requirements have necessitated a new network campus design. Higher user demands and complex applications force the network designers to think more about traffic patterns instead of solving a typical isolated department issue. Now network administrators need to create a network that makes everyone capable of reaching all network services easily. They therefore need to must pay attention to traffic patterns and how to solve bandwidth issues. This can be accomplished with higher-end routing and switching techniques. Because of the new bandwidth-intensive applications, video and audio to the desktop, as well as more and more work being performed on the Internet, the new campus model must be able to perform:

- **Fast Convergence**, i.e., when a network change takes place, the network must be able to adapt very quickly to new changes and keep data moving quickly.
- **Deterministic paths**, i.e., users must be able to gain access to a certain area of the network without fail.
- **Deterministic failover**, i.e., the network design must have provisions which ensure that the network stays up and running even if a link fails.
- **Scalable size and throughput**, i.e., the network infrastructure must be able to handle the new increase in traffic as users and new devices are added to the network.
- **Centralized applications**, i.e., enterprise applications accessed by all users must be available to support all users on the internetwork.
- **The new 20/80 rule**, i.e., instead of 80 percent of the users' traffic staying on the local network, 80 percent of the traffic will now cross the backbone and only 20 percent will stay on the local network. (The new 20/80 rule is discussed below in [Section 1.3](#).)
- **Multiprotocol support**, i.e., networks must support multiple protocols, some of which are routed protocols used to send user data through the internetwork, such as IP or IPX; and some of which are routing protocols used to send network updates between routers, such as RIP, Enhanced Interior Gateway Routing Protocol (EIGRP), and Open Shortest Path First (OSPF).

- **Multicasting**, which is sending a broadcast to a defined subnet or group of users who can be placed in multicast groups.

### 1.3 The 80/20 Rule and the New 20/80 Rule

The traditional campus network followed what is called the **80/20 rule** because 80% of the users' traffic was supposed to remain on the local network segment and only 20% or less was supposed to cross the routers or bridges to the other network segments. If more than 20% of the traffic crossed the network segmentation devices, performance was compromised. Because of this, users and groups were placed in the same physical location. In other words, users who required a connection to one physical network segment in order to share network resources, such as network servers, printers, shared directories, software programs, and applications, had to be placed in the same physical location. Therefore, network administrators designed and implemented networks to ensure that all of the network resources for the users were contained within their own network segment, thus ensuring acceptable performance levels.

With new Web-based applications and computing, any computer can be a subscriber or a publisher at any time. Furthermore, because businesses are pulling servers from remote locations and creating server farms to centralize network services for security, reduced cost, and administration, the old 80/20 rule cannot work in this environment and, hence, is obsolete. All traffic must now traverse the campus backbone, effectively replacing the 80/20 rule with a **20/80 rule**. Approximately 20% of user activity is performed on the local network segment while up to 80% percent of user traffic crosses the network segmentation points to access network services.

The problem that the 20/80 rule has is that the routers must be able to handle an enormous amount of network traffic quickly and efficiently. More and more users need to cross broadcast domains, which are also called Virtual LANs (VLANs). This puts the burden on routing, or layer 3 switching. By using VLANs within the new campus model, you can control traffic patterns and control user access easier than in the traditional campus network. VLANs break up the network by using either a router or switch that can perform layer 3 functions. VLANs are discussed in more detail in [Section 11](#).

The network should be designed around traffic flow and not a specific type of traffic. Each network service type is determined by the situation of the network service in relation to the user. The three types of traffic flow within a campus network are illustrated below.

TABLE 1.1: Network Service Types

Service Type	Service Location	Traffic flow
Local	Same sector (VLAN user)	Access Layer Access
Remote	Different sector (VLAN user)	Distribution Layer Access
Enterprise	Central (Campus users)	Core Layer Access

### 1.4 Characterizing Scalable Internetworks

The key requirements for scalable internetworks are:

- They must be **reliable and available**. This includes being dependable and available 24 hours, 7 days a week. In addition, failures need to be isolated and recovery must be nonvisible to the end user.
- They must be **responsive**. This includes managing the quality of service (QoS) needs for the different protocols being used without affecting response at the desktop.
- They must be **efficient**. Large internetworks must optimize the use of resources, especially bandwidth. Reducing the amount of overhead traffic such as unnecessary broadcasts, service location, and routing updates results in an increase in data throughput without increasing the cost of hardware or the need for additional WAN services.
- They must be **adaptable and serviceable**. This includes being able to accommodate disparate networks and interconnect independent network clusters, as well as to integrate legacy technologies, such as those running Systems Network Architecture (SNA).
- They must be **accessible and secure**. This includes the ability to enable connections into the internetwork using dedicated, dialup, and switched services while maintaining network integrity.

#### 1.4.1 Reliability and Availability

The internetwork should be reliable and available at all layers, especially at the core layer. Core routers are reliable when they can accommodate failures by rerouting traffic and respond quickly to changes in the network topology. The protocols that enhance network reliability and availability that the Cisco IOS supports are scalable protocols such as Open Shortest Path First (OSPF) and Enhanced IGRP (EIGRP). These protocols provide reachability and fast convergence times.

- Scalable networks, including those using a hierarchical design, can have a large number of subnetworks. These networks can be subject to reachability problems due to metric limitations of distance vector routing protocols. Scalable routing protocols such as OSPF and EIGRP use metrics that expand the reachability potential for routing updates because they use cost, rather than hop count, as a metric.
- Scalable protocols can converge quickly because of the router's ability to detect failure rapidly and because each router maintains a network topology map.

#### 1.4.2 Responsiveness

In addition to improving network reachability and reliability, scalable protocols also improve responsiveness because they support alternate paths and load balancing.

- Scalable protocols enable a router to maintain a map of the entire network topology. When a failure is detected the router can reroute traffic by looking at the network topology and finding an alternate path. Enhanced IGRP also keeps a record of alternate routes in case the preferred route goes down.

#### Route Metrics

In a routed network, the routing process relies on the routing protocol to locate the best path to the destination network. Different routing protocols in the TCP/IP environment use different measuring mechanisms, or metrics, to locate the best path to a destination network. In addition, routers advertise the path to a network in terms of a metric value. Some examples of metrics are: hop count and cost. If the destination network is not local to the router, then the path is represented by the total of metric values defined for all of the links that must be traversed to reach the destination network.

Once the routing process knows the metric values associated with the different paths, and then the routing decision can be made. The routing process will select the path that has the smallest metric value.

- Because scalable protocols have a map of the entire network topology, and because of how they maintain their routing tables, they are able to transport data across multiple paths simultaneously to a given location.

In addition, you can configure backup links on WAN connections when you need to make the primary WAN connection more reliable; and when you need to increase availability by configuring the backup connections to be used when a primary connection is experiencing congestion.

### **1.4.3 Efficiency**

Optimizing your network at all layers of an internetwork hierarchy is critical because it can reduce potential costs in additional WAN services. Bandwidth optimization is normally done by reducing the amount of update traffic over a WAN connection, without dropping essential routing information, to increase data traffic throughput. The Cisco IOS has a number of features that help optimize bandwidth use. These are:

- Access lists, which can be used to allow or prevent protocol update traffic, data traffic, and broadcast traffic. Access lists are available for IP and other protocols and can be tailored to meet the needs for each protocol.
- Reduce the number of routing table entries, which reduce the number of router processing cycles. This can be done using route summarization, or incremental updates.
- Dial-on-demand routing (DDR), which can be used to create connections as required for infrequent traffic flow after interesting traffic is detected by the router.
- Switched access, through packet-switched networks such as X.25 and Frame Relay, which offer the advantage of providing global connectivity through a large number of service providers with established circuits to most major cities.
- Snapshot routing, which allows peer routers to exchange full distance vector routing information upon initial connection, then on a predefined interval. This is usually used on ISDN connections and can reduce WAN costs when using distance vector protocols because routing information is exchanged at an interval you define. Between update exchanges, the routing tables for the distance vector protocols are kept frozen.
- Compression, which can be used to reduce traffic that is crossing a WAN connection. Cisco supports TCP/IP header compression and payload compression.

### **1.4.4 Adaptability and Serviceability**

Because scalable internetworks experience change frequently, they must be able to adapt to possible changes. It is difficult to anticipate every change that your company may make in terms of mergers and organizational structure. Therefore, building an adaptable network protects capital investment. It also increases the reliability of the network. It is essential that attention be given to the interoperability of both products and applications when designing the network. Serviceability is related to adaptability, but it is more focused toward being able to make changes to production systems without disrupting normal operations.

### **1.4.5 Accessibility and Security**

Security is a major consideration, particularly as more companies connect to the Internet and thereby increase the chance of access to the network. You must weigh the needs of users to access the network, particularly when remote access is required, against the need to secure the company's network resources. It

is important to consider security as part of the initial design because it is very difficult to address this issue as an afterthought.

## **1.5 Network Congestion**

The consequence of having a network that is incapable of scaling is that as it grows it becomes constricted, resulting in network congestion.

### **1.5.1 Problems Created by Network Congestion**

Network congestion results when too many packets are competing for limited bandwidth. The problems caused by network congestion can be easily identified using network-monitoring tools, such as Cisco's TrafficDirector or a standard protocol analyzer. An understanding of the traffic volumes within the network can also be gained by issuing commands, such as `show interface`, `show buffers`, and `show queuing`, at the Cisco router.

Problems created by network congestion are:

- Excessive traffic;
- Dropped packets;
- Retransmission of packets;
- Incomplete routing tables
- Incomplete server lists;
- The Spanning-Tree Protocol breaks; and
- Runaway congestion.

#### **1.5.1.1 Excessive Traffic**

If the traffic volume outgrows the network, the result is congestion. When this occurs on a single segment, it results in the dropping of packets.

Ethernet has strict rules about accessing the medium. Physical problems, such as extraneous noise or electromagnetic interference, can result in excessive traffic and can cause collisions. A collision requires all transmitting devices to stop sending data and to wait a random amount of time before attempting to send the original packet. Only the nodes involved in the collision are required to wait during the backoff period. Other nodes must wait until the end of the jam signal and the interframe gap. If after 16 attempts the device fails to transmit, it reports an error to the calling process. If for this or any other reason the device fails to transmit and drops the packet from its buffer, the application typically retransmits the original packet. This may result in increased congestion that grows exponentially. The latter is referred to as runaway congestion.

#### **1.5.1.2 Dropped Packets**

When congestion occurs, not all the packets can get through the network. The queues and buffers in the intermediate forwarding devices, such as routers, overflow and must drop packets, causing an OSI higher-layer process on either end device to time out. Typically, the transport or application layers have the responsibility to ensure the arrival of every piece of data. Maintaining the integrity of the transmission

requires the communication to be connection oriented, giving the end devices the mechanisms to perform error detection and correction through windowing, sequencing, and acknowledgments.

#### **1.5.1.3 Retransmission of Packets**

If packets are dropped, the layer responsible for the integrity of the transmission will retransmit the lost packets. If the session or application layer does not receive the packets that were resent in time, the result will be either incomplete information or timeouts.

#### **1.5.1.4 Incomplete Routing Tables**

If a connection congested, packets may be dropped, possibly resulting in the receipt of partial routing updates. If the routing table of an intermediate forwarding device such as a router is incomplete, it may make inaccurate forwarding decisions, resulting in loss of connectivity or even the dreaded routing loop.

#### **1.5.1.5 Incomplete Server Lists**

Congestion results in the random loss of packets. Under extreme circumstances, packet loss may result in incomplete routing tables and server lists. Entries may ghost in and out of these tables. Users may find that their favorite service is sometimes unavailable. The intermittent nature of this type of network problem makes it difficult to troubleshoot.

#### **1.5.1.6 The Spanning-Tree Protocol Breaks**

The Spanning-Tree Protocol is maintained in each Layer 2 device, a switch or a bridge, allowing the device to ensure that it has only one path back to the root bridge. Any redundant paths will be blocked, as long as the Layer 2 device continues to see the primary path. The health of this primary path is ensured by the receipt of spanning-tree updates. As soon as the Layer 2 device fails to see the updates, the device removes the block on the redundant path. The block on the redundant path is removed after several updates have been missed, after the MaxAge timer has been exceeded. This ensures some stability in the network. However, if this problem occurs, in a short time, spanning-tree loops and broadcast storms will cause the network to seize up and die.

#### **1.5.1.7 Runaway Congestion**

When packets are dropped, requiring retransmission, the congestion will inevitably increase. In some instances, this may increase the traffic exponentially; this is often called runaway congestion. In relatively unsophisticated protocols, such as Spanning-Tree Protocol, it is almost unavoidable, although others may have methods of tracking the delays in the network and throttling back on transmission. Both TCP and AppleTalk's DDP use flow control to prevent runaway congestion.

### **1.5.2 Symptoms of Network Congestion**

The symptoms of congestion are intermittent. However, some of these symptoms can be due to other underlying problem within the network. Furthermore, the symptoms of network congestion are difficult to troubleshoot because some protocols are more sensitive than others and will time out after very short delays are experienced. The three symptoms of network congestion are: application time outs; clients cannot connect to network resources; and network death results.

### 1.5.2.1 Application Time Outs

The session layer of the OSI model (Layer 5) is responsible for maintaining the communication flow between the two end devices. This includes assigning resources to incoming requests to connect to an application. To allocate resources adequately, idle timers disconnect sessions after a set time, releasing those resources for other requests. Although the OSI model assigns these duties to the session layer, many protocol stacks, such as TCP/IP, include the upper layers of the stack in the application.

### 1.5.2.2. Clients Cannot Connect to Network Resources

In a client/server environment, the available resources are communicated throughout the network. The dynamic nature of the resource tables gives an up-to-date and accurate picture of the network. NetWare, AppleTalk, Vines, and Windows NT all work on this principle. If these tables are inaccurate as a result of the loss of packets in your network, errors will be introduced because decisions were made with incorrect information. Some network systems are moving more toward a peer-to-peer system in which the end user requests a service identified not by the network, but by the administrator.

### 1.5.2.3 Network Death Results

The most common problems arising from network congestion are intermittent connectivity and excessive delays, users are disconnected from applications, print jobs fail, and errors result when trying to write files to remote servers. If the response of the applications is to retransmit, congestion could reach a point of no recovery. Likewise, if routing or spanning-tree loops are introduced as a result of packet loss, the excessive looping traffic could bring your network down.

## 1.6 Designing Scalable Networks

It is important to know how to reduce network congestion when it occurs, but it is more important that you build a network that is scalable and can accommodate future growth. When used properly in network design, a hierarchical model makes networks more predictable. It helps to define and expect at which levels of the hierarchy we should perform certain functions. The hierarchy requires that you use tools like access lists at certain levels in hierarchical networks and must avoid them at others. In short, a hierarchical model helps us to summarize a complex collection of details into an understandable model. Then, as specific configurations are needed, the model dictates the appropriate manner for in which they are to be applied.

Switching technologies are crucial to the new network design. To understand switching

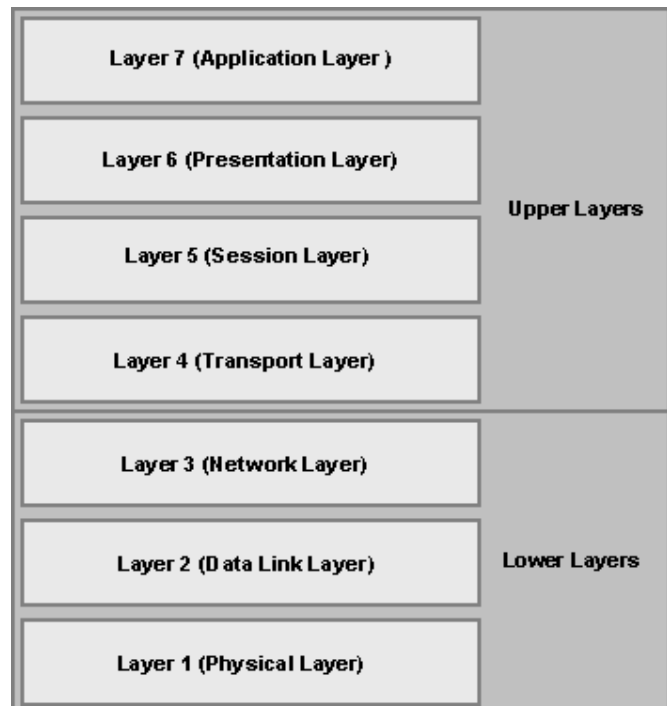


FIGURE 1.1: The Open System Interconnection (OSI Model)

technologies and how routers and switches work together, you must understand the Open Systems Interconnection (OSI) model.

### **1.6.1 Open Systems Interconnection Model**

The OSI model has seven layers (see Figure 1.1), each of which specifies functions that allow data to be transmitted from one host to another on an internetwork. The OSI model is the cornerstone for application developers to write and create networked applications that run on an internetwork. What is important to network engineers and technicians is the encapsulation of data as it is transmitted on a network.

#### **1.6.1.1 Data Encapsulation**

Data encapsulation is the process by which the information in a protocol is wrapped, in the data section of another protocol. In the OSI reference model, each layer encapsulates the layer immediately above it as the data flows down the protocol stack. The logical communication that happens at each layer of the OSI reference model does not involve many physical connections because the information each protocol needs to send is encapsulated in the layer of protocol information beneath it. This encapsulation produces a set of data called a packet.

Each layer communicates only with its peer layer on the receiving host, and they exchange Protocol Data Units (PDUs). The PDUs are attached to the data at each layer as it traverses down the model and is read only by its peer on the receiving side.

*TABLE 1.2: OSI Encapsulation*

<b>OSI Layer</b>	<b>Name of Protocol Data Units (PDUs)</b>	<b>Device to Process PDUs</b>
Transport	TCP Segment	TCP Port
Network	Packet	Router
Data Link	Frame	Bridge/switch

Starting at the Application layer, data is converted for transmission on the network, and then encapsulated in Presentation layer information. The Presentation layer receives this information, and hands the data to the Session layer, which is responsible for synchronizing the session with the destination host. The Session layer then passes this data to the Transport layer, which transports the data from the source host to the destination host. However, before this happens, the Network layer adds routing information to the packet. It then passes the packet on to the Data Link layer for framing and for connection to the Physical layer. The Physical layer sends the data as bits (1s and 0s) to the destination host across fiber or copper wiring. When the destination host receives the bits, the data passes back up through the model, one layer at a time. The data is de-encapsulated at each of the OSI model's peer layers.

The Network layer of the OSI model defines a logical network address. Hosts and routers use these addresses to send information from host to host within an internetwork. Every network interface must have a logical address, typically an IP address.

#### **1.6.1.2 Layer 2 Switching**

Layer 2 (Data Link) switching is hardware based, which means it uses the Media Access Control (MAC) address from the host's network interface cards (NICs) to filter the network. Switches use Application-Specific Integrated Circuits (ASICs) to build and maintain filter tables. Layer 2 switching provides hardware-based bridging; wire speed; high speed; low latency; and low cost. It is efficient because there is no modification to the data packet, only to the frame encapsulation of the packet, and only when the data packet is passing through dissimilar media, such as from Ethernet to FDDI.

Layer 2 switching has helped develop new components in the network infrastructure. These are:

- **Server farms** - servers are no longer distributed to physical locations because virtual LANs can be created to create broadcast domains in a switched internetwork. This means that all servers can be placed in a central location, yet a certain server can still be part of a workgroup in a remote branch.
- **Intranets** allow organization-wide client/server communications based on a Web technology. However, these new components allow more data to flow off of local subnets and onto a routed network, where a router's performance can become the bottleneck.

Layer 2 switches have the same limitations as bridge networks. They cannot break up broadcast domains, which can cause performance issues and limits the size of the network. Thus, broadcast and multicasts, along with the slow convergence of spanning tree, can cause major problems as the network grows. Table 1.3 briefly summarizes the differences between Layer 2 switching and bridges

*TABLE 1.3: Differences between Switches and Bridges*

<b>Operation / Occurrences</b>	<b>Switches</b>	<b>Bridges</b>
Ports	Numerous	Maximum of 16
Filters	Hardware based	Software based
Spanning Tree numbers	Many occurrences	One occurrence

Because of these problems, layer 2 switches cannot completely replace routers in the internetwork. They can however be used for workgroup connectivity and network segmentation. When used for workgroup connectivity and network segmentation, layer 2 switches allows you to create a flatter network design and one with more network segments than traditional 10BaseT shared networks.

**Address learning** occurs when Layer 2 switches and bridges learn the hardware addresses of all devices on an internetwork and enters it into a MAC database. A switch is in essence a multiport transparent bridge. Frame forwarding is based on the MAC addresses that each frame has. A switch forwards a frame when it knows the destination device's location.

The MAC filtering table has nothing in it when a switch is powered. Once a frame is received from a device, the switch retains information on which interface the device is located on. It inserts the source address into the MAC filter table. Since the device's location is unknown at this stage, the network is flooded with the frame.

When a device replies and returns a frame, the switch gets that frame's source address and inserts the MAC address in the MAC database. This source address is connected with the interface on which the frame was initially received. At this point, the switch has two MAC addresses in the MAC filtering table and the devices can create a point-to-point connection. Frames are transmitted just between the two devices.

**Forwarding and Filtering Decisions** is the procedure that a switch uses to establish which ports to forward a frame out of. In addition, the Layer 2 switch uses the MAC filter table to filter received frames. When a switch port receives a frame, it places the frame into one of its ingress queues. The switch then has to decide on the forwarding policies as well find the egress switch port. These decision processes are outlined below.

- **L2 Forwarding Table:** The destination hardware address is utilized as an input key and placed into the Content Addressable Memory (CAM). The egress switch port and its fitting VLAN ID are obtained from the address table if it is listed there. The frame is transmitted out on the correct exit interface.
- **Security Access Control Lists (ACLs):** The Ternary Content Addressable Memory (TCAM) holds ACLs that can be used to single out frames. Frames are identified on their MAC addresses, IP addresses, Layer 4 port numbers and protocol types when the frame is not an IP frame.
- **QoS ACLs:** These ACLs can be utilized to categorize received frames in relation to quality of service (QoS) parameters. In this manner, the extent of traffic flows can be controlled and QoS parameters in outbound frames can be noted.

Another function that Layer 2 switching is responsible for is **Loop Avoidance**. Network loops takes place when there are multiple links between switches that were established for redundancy. Although this can help to prevent network failures, redundant links can cause severe problems. These are noted below

- **Broadcast Storms** occur when switches continuously flood broadcasts all through the network. Loop avoidance help to avoid this situation
- Multiple frames can turn up from different links concurrently and cause **Multiple Frame Copies**. The switch would not know the location of device. **Thrashing the MAC table** happens when a switch cannot send a frame because it is continuously updating the MAC table

### 1.6.1.3 Layer 3 Switching

The difference between a layer 3 (Network) switch and a router is the way the administrator creates the physical implementation. In addition, traditional routers use microprocessors to make forwarding decisions, whereas the layer 3 switch performs only hardware-based packet switching. Layer 3 switches can be placed anywhere in the network because they handle high-performance LAN traffic and can cost-effectively replace routers. Layer 3 switching is all hardware-based packet forwarding, and all packet forwarding is handled by hardware ASICs. Furthermore, Layer 3 switches provide the same functionally as the traditional router. These are:

- Determine paths based on logical addressing;
- Run layer 3 checksums on header only;
- Use Time to Live (TTL);
- Process and responds to any option information;
- Can update Simple Network Management Protocol (SNMP) managers with Management Information Base (MIB) information; and
- Provide Security.

### Routers

Routers and layer 3 switches are similar in concept but not design. Like bridges, routers break up collision domains but they also break up broadcast/multicast domains.

The benefits of routing include:

- Break up of broadcast domains;
- Multicast control;
- Optimal path determination;
- Traffic management;
- Logical (layer 3) addressing; and
- Security.

Routers provide optimal path determination because the router examines every packet that enters an interface and improves network segmentation by forwarding data packets to only a known destination network. If a router does not know about a remote network to which a packet is destined, it will drop the packet. Because of this packet examination, traffic management is obtained. Security can be obtained by a router reading the packet header information and reading filters defined by the network administrator.

The benefits of Layer 3 switching include:

- Hardware-based packet forwarding;
- High-performance packet switching;
- High-speed scalability;
- Low latency;
- Lower per-port cost;
- Flow accounting;
- Security; and
- Quality of service (QoS).

#### **1.6.1.4 Layer 4 Switching**

Layer 4 (Transport) switching is considered a hardware-based layer 3 switching technology. It provides additional routing above layer 3 by using the port numbers found in the Transport layer header to make routing decisions. These port numbers are found in Request for Comments (RFC) 1700 and reference the upper-layer protocol, program, or application.

The largest benefit of layer 4 switching is that the network administrator can configure a layer 4 switch to prioritize data traffic by application, which means a QoS can be defined for each user. However, because users can be part of many groups and run many applications, the layer 4 switches must be able to provide a huge filter table or response time would suffer. This filter table must be much larger than any layer 2 or 3 switch. A layer 2 switch might have a filter table only as large as the number of users connected to the network while a layer 4 switch might have five or six entries for each and every device connected to the network. If the layer 4 switch does not have a filter table that includes all the information, the switch will not be able to produce wire-speed results.

#### **1.6.1.5 Multi-Layer Switching (MLS)**

Multi-layer switching combines layer 2 switching, layer 3 switching, and layer 4 switching technologies and provides high-speed scalability with low latency. It accomplishes this by using huge filter tables based on the criteria designed by the network administrator. Multi-layer switching can move traffic at wire speed while also providing layer 3 routing. This can remove the bottleneck from the network routers. Multi-layer switching can make routing/switching decisions based on:

- The MAC source/destination address in a Data Link frame;
- The IP source/destination address in the Network layer header;
- The Protocol field in the Network layer header; and
- The Port source/destination numbers in the Transport layer header.

Two types of MLS are supported by catalyst switches:

- **Route caching** needs a switch engine (SE) and a route processor (RP). The RP processes a traffic flow's first packet in order to establish the destination, while the SE inserts an entry in its MLS cache to store

the relevant destination. These SE uses these entries when sending the next packets in the same traffic flow. Route caching is also known as demand-based switching, NetFlow LAN switching and flow-based switching.

- With **Topology-based** switching, also known as **Cisco Express Forwarding (CEF)**, Layer 3 routing data creates and preloads a database of the whole network topology. This database is checked when forwarding packets.

Packets entering the switch port are located in the ingress queue, and are then extracted and examined for Layer 2 and Layer 3 destination. A decision process is performed to determine the destination for the packet and the forwarding policies:

- **L2 Forwarding Table:** The destination MAC address is utilized like an input key to the CAM table. When the frame has a Layer 3 packet, the only action taken is to process the packet at that layer.
- **L3 Forwarding Table:** The destination IP address is utilized as an input key and checked against the FIB table. The FIB table also holds the egress switch port with its fitting VLAN ID, and each entry's Layer 2 MAC address.
- **Security Access Control Lists (ACLs):** The **Ternary Content Addressable Memory (TCAM)** holds ACLs that can be used to single out frames. A decision on whether to forward a packet is done as a single table lookup.
- **QoS ACLs:** These ACLs can be utilized to categorize received frames in relation to quality of service (QoS) parameters. Packet categorization and marking can be done as a single table lookups in the QoS TCAM.

The following are excluded from MLS because they cannot be directly forwarded by CEF:

- Cisco Discovery Protocol packets
- ARP requests and replies
- IP packets that needs a reply from a router
- Routing protocol updates
- IPX routing protocol
- Packets that are not IP and IPX protocol packets
- IP broadcasts to be passed on as unicast
- Packets setting off Network Address Translation (NAT)
- Packets that require encryption

### **1.6.2 The Cisco Hierarchical Model**

When used properly in network design, a hierarchical model makes networks more predictable. It helps to define and expect at which levels of the hierarchy we should perform certain functions. The hierarchy requires that you use tools like access lists at certain levels in hierarchical networks and must avoid them at others.

## CCNP/CCDP 642-891 (Composite)

The Cisco hierarchical model is used to design a scalable, reliable, cost-effective hierarchical internetwork. Cisco defines three layers of hierarchy: the core layer; the distribution layer; and the access layer. These three layers are logical and not necessarily physical. They are thus not necessarily represented by three separate devices.

The Cisco hierarchical model is used to design a scalable, reliable, cost-effective hierarchical internetwork. Cisco defines three layers of hierarchy: **the core layer**; **the distribution layer**; and **the access layer**. These three layers are logical and not necessarily physical. They are thus not necessarily represented by three separate devices. Each layer has specific responsibilities, allowing only certain traffic to be forwarded through to the upper levels. A filtering operation restricts unnecessary traffic from traversing the network. Thus, the network is more adaptable, scalable, and more reliable.

### 1.6.2.1 Core Layer

At the top of the hierarchy is the core layer. It is literally the core of the network and is responsible for switching traffic as quickly as possible. The traffic transported across the core is common to a majority of users. However, user data is processed at the distribution layer, and the distribution layer forwards the requests to the core, if needed. If there is a failure in the core, every all user can be affected; therefore, fault tolerance at this layer is critical.

As the core transports large amounts of traffic, you should design the core for high reliability and speed. You should thus consider using data-link technologies that facilitate both speed and redundancy, such as FDDI, FastEthernet (with redundant links), or even ATM. You should use routing protocols with low convergence times. You should avoid using access lists, routing between virtual LANs (VLANs), and packet filtering. You should also not use the core layer to support workgroup access and upgrade rather than expand the core layer if performance becomes an issue in the core.

The following Cisco switches are recommended for use in the core:

- **The 5000/5500 Series.** The 5000 is a great distribution layer switch, and the 5500 is a great core layer switch. The Catalyst 5000 series of switches includes the 5000, 5002, 5500, 5505, and 5509. All of the 5000 series switches use the same cards and modules, which makes them cost effective and provides protection for your investment.
- **The Catalyst 6500 Series**, which are designed to address the need for gigabit port density, high availability, and multi-layer switching for the core layer backbone and server-aggregation environments. These switches use the Cisco IOS to utilize the high speeds of the ASICs, which allows the delivery of wire-speed traffic management services end to end.
- **The Catalyst 8500**, which provides high performance switching. It uses Application-Specific Integrated Circuits (ASICs) to provide multiple-layer protocol support including Internet Protocol (IP), IP multicast, bridging, Asynchronous Transfer Mode (ATM) switching, and Cisco Assure policy-enabled Quality of Service (QoS). All of these switches provide wire-speed multicast forwarding, routing, and Protocol Independent Multicast (PIM) for scalable multicast routing. These switches are perfect for providing the high bandwidth and performance needed for a core router. The 6500 and 8500 switches can aggregate multiprotocol traffic from multiple remote wiring closets and workgroup switches.

### 1.6.2.2 Distribution Layer

## CCNP/CCDP 642-891 (Composite)

The distribution layer is the communication point between the access layer and the core. The primary function of the distribution layer is to provide routing, filtering, and WAN access and to determine how packets can access the core, if needed. The distribution layer must determine the fastest way that user requests are serviced. After the distribution layer determines the best path, it forwards the request to the core layer. The core layer is then responsible for quickly transporting the request to the correct service. You can implement policies for the network at the distribution layer. You can exercise considerable flexibility in defining network operation at this level.

Generally, you should:

- Implement tools such as access lists, packet filtering, and queuing;
- Implement security and network policies, including address translation and firewalls;
- Redistribute between routing protocols, including static routing;
- Route between VLANs and other workgroup support functions; and
- Define broadcast and multicast domains.

The distribution layer switches must also be able to participate in multi-layer switching (MLS) and be able to handle a route processor. The Cisco switches that provide these functions are:

- **The 2926G**, which is a robust switch that uses an external router processor like a 4000 or 7000 series router.
- **The 5000/5500 Series**, which is the most effective distribution layer switch, it can support a large amount of connections and also an internal route processor module called a Route Switch Module (RSM). It can switch process up to 176KBps.
- **The Catalyst 6000**, which can provide up to 384 10/100 Ethernet connections, 192 100FX FastEthernet connections, and 130 Gigabit Ethernet ports.

### 1.6.2.3 Access Layer

The access layer controls user and workgroup access to internetwork resources. The network resources that most users need will be available locally. Any traffic for remote services is handled by the distribution layer. At this layer access control and policies from distribution layer should be continued and network segmentation should be implemented. Technologies such as dial-on-demand routing (DDR) and Ethernet switching are frequently used in the access layer.

The switches deployed at this layer must be able to handle connecting individual desktop devices to the internetwork. The Cisco solutions that meet these requirements include:

- **The 1900/2800 Series**, which provides switched 10 Mbps to the desktop or to 10BaseT hubs in small to medium campus networks.
- **The 2900 Series**, which provides 10/100 Mbps switched access for up to 50 users and gigabit speeds for servers and uplinks.
- **The 4000 Series**, which provides a 10/100/1000 Mbps advanced high-performance enterprise solution for up to 96 users and up to 36 Gigabit Ethernet ports for servers.

- **The 5000/5500 Series**, which provides 10/100/1000 Mbps Ethernet switched access for more than 250 users.

### **1.6.3 Modular Network Design**

Cisco promotes a campus network design based on a modular approach. In this design approach, each layer of the hierarchical network model can be broken down into basic functional modules or blocks. These modules can then be sized appropriately and connected together, while allowing for future scalability and expansion. A building block approach to network design. Campus networks based on the modular approach can be divided into basic elements. These are:

- **Switch blocks**, which are access layer switches connected to the distribution layer devices; and
- **Core blocks**, which are multiple switch blocks connected together with possibly 5500, 6500, or 8500 switches.

Within these fundamental campus elements, there are other contributing variables that can be added to the network. These are:

- **Server Farm blocks**, which are groups of network servers on a single subnet
- **Enterprise Edge blocks** are centralized services to which the enterprise network is responsible for providing complete access, together with their related access and distribution switches.
- **Network Management blocks** are a set of network management resources with their accompanying access and distribution switches
- **Service Provider Edge blocks** are multiple connections to an ISP or multiple ISPs

#### **1.6.3.1 The Switch Block**

The switch block is a combination of **layer 2 switches** and **layer 3 routers**. The layer 2 switches connect users in the wiring closet into the access layer and provide 10/100 Mbps dedicated connections. 1900/2820 and 2900 Catalyst switches can be used in the switch block. From here, the access layer switches will connect into one or more distribution layer switches, which will be the central connection point for all switches coming from the wiring closets. The distribution layer device is either a switch with an external router or a multi-layer switch. The distribution layer switch will then provide layer 3 routing functions, if needed.

The distribution layer router will prevent broadcast storms that could happen on an access layer switch from propagating throughout the entire internetwork. Thus, the broadcast storm would be isolated to only the access layer switch in which the problem exists.

Switch block sizing at the access layer is based on the quantity of users or the port density. Distribution layer sizing is based on the quantity of access layer switches that are passed into a distribution mechanism. When sizing the distribution layer, the following should be considered:

- Traffic types and behaviors
- Quantity of users connected to access layer switches
- Layer 3 switching abilities on the distribution layer

- The size of Spanning Tree domains
- The physical confines of VLANs

Designing a switch block should be based essentially on traffic types and behaviors, and the quantity and extent of workgroups. Because a switch block can be too large or too small, the ability to break up or downsize a switch block should be catered for. A switch block is too large when multicast or broadcast traffic reduces speed of the switch block switches, or the distribution layer multilayer switches turn into traffic blockages.

Access switches are able to contain one or many redundant links to distribution layer mechanisms. This enables traffic to be load balanced across redundant links using redundant gateways.

### **1.6.3.2 The Core Block**

If you have two or more switch blocks, you need a core block which will be responsible for transferring data to and from the switch blocks as quickly as possible. You can build a fast core with a frame, packet, or cell (ATM) network technology. Typically, have two or more subnets configured on the core network for redundancy and load balancing.

Switches can trunk on a certain port or ports. This means that a port on a switch can be a member of more than one VLAN at the same time. However, the distribution layer will handle the routing and trunking for VLANs, and the core is only a pass-through once the routing has been performed. Because of this, core links will not carry multiple subnets per link. A Cisco 6500 or 8500 switch is recommended at the core. Even though one switch might be sufficient to handle the traffic, Cisco recommends two switches for redundancy and load balancing purposes.

#### **1.6.3.2.1 The Collapsed Core**

A **collapsed core** is defined as one switch device performing both core and distribution layer functions. The collapsed core is typically found in smaller campus networks where a separate core layer is not warranted. Although the distribution and core layer functions are performed in the same device, keeping these functions distinct and properly designed remain of importance. In the collapsed core design, each access layer switch has a redundant link to each distribution/core layer switch and each access layer switch may support more than one VLAN. The distribution layer routing is the termination for all ports. In a collapsed core network, Spanning-Tree Protocol (STP) blocks the redundant links to prevent loops. Hot Standby Routing Protocol (HSRP) can provide redundancy in the distribution layer routing. It can keep core connectivity if the primary routing process fails.

#### **1.6.3.2.2 Dual Core**

A dual core connects **two or more switch blocks** in a **redundant** fashion. Each connection would be a separate subnet. Redundant links connect the distribution layer portion of each switch block to each of the dual core switches. In the dual core, each distribution switch has two equal-cost paths to the core, providing twice the available bandwidth. The distribution layer routers would have links to each subnet in the routing tables, provided by the layer 3 routing protocols. If a failure on a core switch takes place, convergence time will not be an issue. HSRP can be used to provide quick cutover between the cores.

### **1.6.3.2.3 Core Size**

The dual core is made up of redundant switches, and is bounded and **isolated** by Layer 3 devices. Routing protocols determine paths and maintain the operation of the core. You must pay attention to the overall design of the routers and routing protocols in the network. As routing protocols propagate updates throughout the network, network topologies might be undergoing change. The size of the network, i.e., the number of routers, then affects routing protocol performance, as updates are exchanged and network convergence takes place. Large campus networks can have many switch blocks connected into the core block. Layer 2 devices are used in the core with usually only a single VLAN or subnet across the core. Therefore, all route processors connect into a single broadcast domain at the core.

Each route processor must communicate with and keep information about each of its directly connected peers. Thus, most routing protocols have practical limits on the number of peer routers that can be connected. Because two equal-cost paths from each distribution switch into the core, each router forms two peer relationships with every other router. Therefore, the actual maximum number of switch blocks that can be supported is half the number of distribution layer routers. In the case of dual core design, the equalcost paths must lead to isolated VLANs or subnets if a routing protocol supports two equal-cost paths. Thus, two equal-cost paths are used in a dual core design with two Layer 2 switches. Likewise, a routing protocol that supports six equal-cost paths requires that the six distribution switch links be connected to exactly six Layer 2 devices in the core. This gives six times the redundancy and six times the available bandwidth into the core.

### **1.6.3.2.4 Core Scalability**

As the number of switch blocks increases, the core block must also be capable of scaling without needing to be redesigned. Traditionally, hierarchical network designs have used Layer 2 switches at the access layer, Layer 3 devices at the distribution layer, and Layer 2 switches at the core. This design is called a Layer 2 Core has been very cost effective and has provided high-performance connectivity between switch blocks in the campus. As the network grows, more switch blocks must be added to the network, which in turn requires more distribution switches with redundant paths into the core. The core must then be scaled to support the redundancy and the additional campus traffic load.

Providing redundant paths from the distribution switches into the core block allows the Layer 3 distribution switches to identify several equal-cost paths across the core. If the number of core switches must be increased for scalability, the number of equal-cost paths can become too much for the routing protocols to handle. Because the core block is formed with Layer 2 switches, the Spanning-Tree Protocol (STP) is used to prevent bridging loops. If the core is running STP, then it can compromise the high-performance connectivity between switch blocks. The best design on the core is to have two switches without STP running. You can do this only by having a core without links between the core switches.

### **1.6.3.2.5 Layer 3 Core**

Layer 3 switching can also be used in the core to fully scale the core block for large campus networks. This approach overcomes the problems of slow convergence, load balancing limitations, and router peering limitations. In a Layer 3 core, the core switches can have direct links to each other. Because of Layer 3 functionality, the direct links do not impose any bridging loops.

With a Layer 3 core, the path determination intelligence occurs in both the distribution and core layers, allowing the number of core devices to be increased for scalability. Redundant paths also can be used to

interconnect the core switches without concern for Layer 2 bridging loops, eliminating the need for STP. If you have only Layer 2 devices at the core layer, the STP will be used to stop network loops if there is more than one connection between core devices. The STP has a **convergence time** of over 50 seconds, and if the network is large, this can cause an enormous amount of problems if it has just one link failure. However, STP would not be implemented in the core if the core has Layer 3 devices. Instead, routing protocols, which have a much faster convergence time than STP, could be implemented. In addition, the routing protocols can **load balance** with multiple equal-cost links. STP is discussed in more detail in [Section 12.2](#).

Router peering problems are also overcome as the number of routers connected to individual subnets is reduced. Distribution devices are no longer considered peers with all other distribution devices. Instead, a distribution device peers only with a core switch on each link into the core. This advantage becomes especially important in very large campus networks involving more than 100 switch blocks. However, Layer 3 devices are more expensive than Layer 2 devices. The Layer 3 devices also need to have switching latencies comparable to their Layer 2 counterparts. Using a Layer 3 core also adds additional routing hops to cross-campus traffic.

### **1.6.3.3 Additional Building Blocks**

Additional resources can be assembled into building blocks, and can be located and arranged in the same manner as common switch block modules.

- **Server Farm Blocks:** - Enterprise servers comprising of company e-mail, intranet services, mainframe systems and Enterprise Resource Planning (ERP) applications normally belong to a server farm. These enterprise resources are accessed by most of the connected users. The whole server farm can be structured into a switch block with its own layer of access switches. These access switches are then uplinked to dual distribution switches that are connected into the core layer by means of redundant high-speed links. **Dual-homing** the servers occurs when each server has dual network connections - one to each distribution switch.
- **Enterprise Edge Blocks:** - Campus networks connect to service providers at the **edge of the campus network** to gain access to these service providers' external resources or services. The resources are utilized by the whole network and can be grouped into a single switch block that is connected to the core network. These resources can comprise of internet access, WAN access, e-commerce and remote access and VPNs.
- **Network Management Blocks:** - Network management resources and policy management applications such as system logging servers and authentication, authorization, and accounting (AAA) servers, can be grouped into a single network management switch block. These resources access application servers, network devices and user connectivity and actions. This single network management switch block has a distribution layer that links into the core switches. Redundant links and redundant switches are usually utilized to ensure that the resources are always available.
- **Service Provider Edge Blocks:** - A service provider has its own hierarchical network design and can be viewed as an enterprise or campus network. A campus network contains an edge block and connects to each service provider's network edge from there.

## **1.7 Alleviating Congestion**

### **1.7.1 Access Lists**

You can alleviate congestion by controlling network traffic. Cisco routers have features, such as **access lists**, that you can use to control network traffic. Access lists are crucial to the programming of a Cisco router and allow for the control of traffic. Given that the router operates at Layer 3, the control that is offered is extensive. The router can also act at higher layers of the OSI model. This is useful when identifying particular traffic and protocol types for prioritization across slower WAN links.

Access lists can be used to either restrict or police traffic entering or leaving a specified interface. The access lists used for IP enable you to apply great subtlety in the router's configuration. Access lists are linked lists with a top-down logic, ending in an implicit **deny any** command, which will deny everything. Top-down logic means that the process will read from the top of the access list and stop as soon as it meets the first entry in the list that matches the packet's characteristics. Therefore, it is crucial that careful attention be given to their creation. Writing down the purpose of the proposed access list before attempting to program the system also proves helpful. Access lists block traffic traversing the router but does not prevent or block the traffic generated by the router.

The syntax for a standard **access-list** command and an **ip access-group** command are:

```
access-list access-list_number { permit | deny }  
    { source [ source_wildcard | any ] }
```

and

```
ip access-group access-list_number { in | out }
```

The **access-list\_number** must be 1-99 to create a standard access list. Standard access lists are implemented at Layer 3. In general, both the source and destination addresses are identified as criteria in the logic of the list.

IP access lists use the source address only. The placement of the access list is crucial because it may determine the effectiveness of the control imposed. Because the decision to forward can be made on the source address only, the access list is placed as close to the destination as possible to allow connectivity to intermediary devices. You can place an access list on either an inbound or an outbound interface. If this option is not configured, the default is for the access list to be placed on the outbound interface. The access list will examine traffic flowing only in the direction stated. In this way, traffic subject to an inbound access list will be examined before it is sent to the routing process. To ensure that all paths to the remote location have been covered, access lists should be implemented with reference to the network topology map.

### 1.7.2 Extended Access Lists

Although the same rules apply for all access lists, **extended access lists** allow for a far greater level of control because decisions are made at higher levels of the OSI model. The syntax of an extended **access-list** command is:

```
access-list access-list_number { deny | permit } protocol  
    source source_wildcard destination destination_wildcard  
ip access-group access-list_number { in | out }
```

The **access-list-number** must be 100-199 to create an extended access list.

TABLE 1.4: Parameters for the Extended access-list Command

Command	Description
<i>access-list</i> <i>number</i>	Specifies the number of an access list.
{ deny   permit }	Denies or permits access if the conditions are matched.
<i>source</i> <i>source_wildcard</i>	Gives the source address and the wildcard mask.
<i>destination</i> <i>destination_wildcard</i>	Gives the destination address and the wildcard mask.
[ precedence <i>precedence</i> ]	Packets can be filtered by precedence level, as specified by a number from 0 to 7, or by <i>name</i> .
[ tos <i>tos</i> ]	Packets can be filtered by type of service level, as specified by a number from 0 to 15, or by <i>name</i> .
[ established ] (For TCP only)	Indicates an established connection. A match occurs if the TCP datagram has the ACK or RST bits set. The nonmatching case is that of the initial TCP datagram to form a connection.
[ log ]	Gets access list logging messages, including violations.

Top-down logic is employed in extended access lists. It is therefore important to consider the sequence of conditions within the list.

You can use **show** commands to verify the filter configuration for either IP or IPX filters. The two commands to use are:

```
show access list
```

and

```
show ip interface
```

Although access lists are used mainly to manage traffic, they can be used to solve many other problems. An access list is a list that uses a simple logic to decide whether to forward traffic. As such, access lists are sometimes used as a security system. However, although access lists are complex, they are easily spoofed and defeated. Starting with IOS release 11.3, Cisco has implemented full security features that should be utilized in preference to access lists.

In addition, access lists can be used to control Telnet traffic. However, access lists filter traffic traversing the router but not the traffic generated by the router. To control Telnet traffic in which the router is the end station, an access list can be placed on the virtual terminal line (vty). Five terminal sessions are available: vty 0 through vty 4. Because anticipating which session will be assigned to which terminal is difficult,

control is generally placed uniformly on all virtual terminals. Although this is the default configuration, some platforms have different limitations on the number of vty interfaces that can be created.

The syntax for virtual terminal `line` commands are:

```
line { vty_number | vty_range }  
access-class access-list_number { in | out }
```

### 1.7.3 Distribution Lists

Traffic management is most easily accomplished at Layer 3 of the OSI model. However, limiting traffic at Layer 3 can also limit connectivity. Therefore, careful design is required. Routing updates convey information about the available networks.

In most routing protocols, these updates are sent out periodically to ensure that every router's perception of the network is accurate and current. Access lists applied to routing protocols restrict the information sent out in the update and are called **distribute lists**. They work by omitting certain networks based on the criteria in the access list. The result is that remote routers that are unaware of these networks are not capable of delivering traffic to them. Distribute lists are also used to prevent routing loops in networks that have redistribution between routing protocols. When connecting two separate domains, the connection point of the domains or the entry point to the Internet is an area through which only limited information needs to be sent. Otherwise, routing tables become unmanageably large and consume large amounts of bandwidth.

### 1.7.4 Other Solutions to Traffic Control

It is popular to tune the update timers between routers, trading currency of the information for optimization of the bandwidth. All routers running the same routing protocol expect to hear these updates with the same frequency that they send out their own. If any of the parameters defining how the routing protocol works are changed, these alterations should be applied consistently throughout the network; otherwise, routers will time out and the routing tables will become unsynchronized.

It may be advantageous to completely turn off routing updates across WAN networks and to manually or statically define the best path to be taken by the router. Routing protocols such as EIGRP or OSPF send out only incremental updates. However, these are correspondingly more complex to design and implement, although the configuration is easier. Another method of reducing routing updates is to implement the technology snapshot routing available on Cisco routers and designed for use across WAN links. This allows the routing tables to be frozen and updated either at defined times, such as every two days or whenever the dialup line is raised. For more information on this topic, refer to the Cisco web page.

### 1.7.5 Prioritization

Access lists are not used just to determine which packets will be forwarded to a destination. On a slow network connection where bandwidth is at a premium, access lists are used to determine the order in which traffic is scheduled to leave the interface. Unfortunately, some of the packets may time out. Therefore, it is important to plan the prioritization based on your understanding of the network. It is important to ensure that the traffic most likely to time out, such as IBM's Systems Network Architecture (SNA), is handled first.

There are many types of prioritization available. These types of prioritization are referred to as queuing techniques. They are implemented at the interface level and are applied to the interface queue. These include:

- **Weighted Fair Queuing (WFQ)**, which is replacing the **First-In, First-Out (FIFO)** queuing mechanism as the default. The queuing process analyzes the traffic patterns on the link, based on the size of the packets and the nature of the traffic, to distinguish interactive traffic from file transfers. The queue then transmits traffic based on its conclusions.
- **Priority Queuing**, which is a method of dividing the outgoing interface into four virtual queues. Importance or priority ranks these queues, so traffic is queued based on its importance and will be sent out of the interface accordingly. This method ensures that sensitive traffic, such as SNA traffic, on a slow or congested link is processed first.
- **Custom Queuing**, which is a method of dividing the outgoing interface into many subqueues. Each queue has a threshold stating the number of bytes that may be sent before the next queue must be processed. In this way, it is possible to determine the percentage of bandwidth that each protocol is given.
- **Class-based Weighted Fair Queuing (CBWFQ)**, which extends the standard WFQ functionality to provide support for user-defined traffic classes. For CBWFQ, you define traffic classes based on match criteria, including protocols, access lists, and input interfaces. Packets satisfying the match criteria for a class constitute the traffic for that class. A queue is reserved for each class, and traffic belonging to a class is directed to that class's queue.
- **Low-Latency Queuing (LLQ)**, which brings strict priority queuing to CBWFQ. Configured by the priority command, strict priority queuing gives delay-sensitive data, such as voice, preferential treatment over other traffic. With this feature, delay-sensitive data is sent before packets in other queues.

#### **1.7.5.1 First In, First Out (FIFO)**

FIFO is the most basic queuing strategy. It is the first-come, first-served approach to data forwarding. In FIFO, packets are transmitted in the order in which they are received. Until recently, FIFO was the default queuing strategy for all interfaces on a router. However, should it become necessary for the traffic to be reordered in any way, another strategy must be invoked because FIFO gives no regard to one type of traffic over another. It simply dispatches data as it receives it. This, however, is not really queuing; but is more along the lines of buffering. The packets are routed to the interface and stored in router memory until transmittal. The transmission order is based on the arrival order of the first bit of the packet.

#### **1.7.5.2 Weighted Fair Queuing (WFQ)**

Weighted Fair Queuing (WFQ) enables Telnet and other interactive traffic to have priority over FTP and other large transfers. This improves overall throughput. The FTP packets get through with relatively little delay, and Telnet users see improved response times. In WFQ, traffic is sorted by high-volume and low-volume communications (sessions). The traffic in a session is kept within one session, and the records are handled FIFO within a particular session. The lower volume interactive traffic is given a priority and flows first. The necessary bandwidth is allocated to the interactive traffic, and the high-volume traffic equally shares the remaining bandwidth. WFQ is the default queuing strategy on interfaces of less than 2 MB because at higher speeds, queuing is usually not necessary. In addition, WFQ is on by default for interfaces that support it. WFQ is not used by default on Link Access Procedure on the B channel (LAPB) for X.25, compressed Point-to-Point Protocol (PPP), or Synchronous Data Link Control (SDLC) interfaces.

Discrimination of traffic conversations is based on source and destination packet header addresses. Other factors such as source and destination MAC addresses, source and destination port or socket numbers, data-link connection identifier (DLCI) in Frame Relay deployments, Quality of Service (QoS) values, and Type of Service (ToS) values also provide discriminatory criteria to the WFQ process. With WFQ, the low-volume traffic is given the priority on the outbound interface.

Configuring WFQ involves adjusting the queue limits. To keep some sessions from overwhelming the circuit, you can configure the maximum number of records that any high-volume traffic allows into the queue. The default setting is 64 records, but the supported range is from 1–4096. If a session reaches the queue limit, no further records are queued for that session until the percentage of the entries in the queue for that conversation drops. All new packets for the over-queue-limit conversation are dropped and lost. TCP window sizes control the amount of data that can be transmitted between two hosts without an acknowledgement. A larger window size enables a higher transmission threshold. Sessions using TCP that suffer a packet drop retransmit automatically as part of the Layer 4 flow control process. As a consequence, the communicating parties in the TCP conversation are forced to reduce their window sizes. The **fair-queue** command is shown below:

```
RouterA(config-if)#fair-queue [ queue_limit ]
```

WFQ can be disabled using the **no fair-queue** command.

### **1.7.5.3 Priority Queuing (PQ)**

When absolute control over the throughput is necessary, priority queuing should be utilized. Priority queuing gives the network administrator granular control that reduces network delay for high-priority traffic. Variations of priority queuing have been in use for a number of years in differing vendor implementations. Cisco's implementation of priority queuing utilizes four queues: high, medium, normal, and low. For traffic placed in individual queues, the output strategy is FIFO. The traffic defined as high priority receives the benefit of all available resources on the output interface until the queue is empty. Once the high queue is complete, the medium queue traffic is dispatched in the same manner until empty. At this stage, the high queue is again checked for content and emptied if there is any new traffic. If there are no entries in the high queue after servicing the medium queue, the normal queue is emptied. Once normal traffic has been dispatched, the high queue is checked again, followed by the medium and normal queues. If all three are empty, the low queue is serviced. The result is that high-priority traffic always suffers the shortest delay in awaiting dispatch.

The low-priority traffic must wait until it can be serviced. The traffic can even age out and be purged from memory if the queue overflows. Once an overflow occurs, all new packets for that queue are dropped until space is freed up in the queue. Each queue has a fixed length, which is configurable. The defaults are 20 records for high, 40 records for medium, 60 records for normal, and 80 records for low. The lower priority queues are larger, by default, than the higher priority queues to accommodate the queuing algorithm and the fact that the lower priority queues might wait longer to be serviced.

The configuration of priority queuing, in the most basic of configurations, entails configuring each protocol that traverses a particular WAN link to enter a specific queue. In more advanced configurations, standard or extended access lists can be defined for specific traffic types and applied to a queue configuration. In priority queuing, the **priority-list** commands are read in the order of their appearance until a matching protocol or interface type is found. When a match is found, the packet is assigned to the appropriate queue

and the search ends. Therefore, some planning needs to go into the creation of the list. The configuration of priority queuing entails defining specific access lists if they are to be used; creating the priority list; applying the priority list to the interface; and verifying the queuing process.

- If it is necessary to queue traffic based on a specific network address, protocol, or application, access lists can be put in place to sort the traffic. Standard or extended access lists can be defined to specify the traffic type or types that should be placed into a specific queue.
- The command parameters for priority queue configuration for a specific protocol or traffic type is:

```
RouterA(config)#priority-list list_number protocol protocol  
{ high | medium | normal | low } queue_keyword keyword_value
```

The *list\_number* argument can be an arbitrarily selected number from 1–16; however, all lines for a particular priority list must have the same *list\_number* to function properly. The *queue\_keyword* and *keyword\_value* parameters are used to associate access lists with the priority list.

- Once the priority list is created, it must be associated with an interface. The priority list is activated on the interface by the **priority-group** command.
- Verifying the queuing configuration can be performed by using the **show queueing** command, which shows the detail of the priority lists configured on the router and the appropriate details of each list.

**Note:** The command used to verify the queuing configuration is **show queueing** and not **show queuing**. The latter command is not recognized by the Cisco IOS.

### 1.7.5.4 Custom Queuing

Custom queuing enables the sharing of available bandwidth across all types of traffic. This technique allocates a percentage of bandwidth to each of the various traffic types. The difference between this approach and priority queuing is that the queues are processed in round-robin sequence. Therefore, it is possible that high priority traffic would not be serviced quickly enough because although each type of traffic would get some bandwidth, no traffic would be designated with a higher priority than the rest. Custom queuing employs 17 queues with queue 0, which the system queue, being used for the system. The remaining 16 queues can be configured by the administrator.

By default, queues evenly balance traffic. There are two thresholds by which queues are measured: **queue limit** and **byte count**. The queue limit default is 20 records. The byte count limit default is 1500 bytes. Whichever limit is reached first signifies the end of a particular queue's time with the processor. If the byte count limit is reached during the transmission of a packet, the entire packet is dispatched.

As with priority queuing, the configuration of custom queuing involves the creation of a list and associating a group with an interface. Traffic in the queues can be configured based on a specific traffic type, protocol, or input interface. Access lists can be configured to place specific traffic types into a particular queue, and traffic not designated to a particular queue can be placed in a default queue.

**Note:** To implement custom queuing on a Frame Relay interface, Frame Relay traffic shaping must be disabled.

## CCNP/CCDP 642-891 (Composite)

- If it is necessary to queue traffic based on a specific network address, protocol, or application, access lists can be put in place to sort the traffic. Standard or extended access lists can be defined to specify the traffic type or types that should be placed into a specific queue.
- The command parameters for custom queue configuration is:

```
RouterA(config)#queue-list list_number protocol protocol queue_number  
queue_keyword keyword_value
```

The *list\_number* argument can be an arbitrarily selected number from 1–16; however, all lines for a particular queue list must have the same *list\_number* to function properly. The *queue\_keyword* and *keyword\_value* parameters are used to associate access lists with the queue list.

It is also possible to specify that any traffic that entered the router through a particular interface be placed into a particular queue, using the following command:

```
RouterA(config)#queue-list list_number interface interface_type  
interface_number queue_number
```

Any traffic that does not match any lines in a priority list is placed in the default queue. For custom queuing, the default queue is queue 1. The command for assigning a default queue is:

```
Router(config)#queue-list list_number default queue_number
```

The amount of data a queue can service before having to move on to the next queue is known as a **service threshold**. You can alter the service threshold of each individual queue. The command for resizing a queue's record limit service threshold is:

```
RouterA(config)#queue-list list_number queue queue_number  
limit limit_number
```

Valid entries for this command are 0–32, 767.

The command structure for altering the byte-count service threshold is as follows:

```
RouterA(config)#queue-list list_number queue queue_number  
byte-count byte_count_number
```

- Once the queue list is created, it must be associated with an interface. The queue list is activated on the interface by the `custom-queue-list` command.
- Verifying the queuing configuration can be performed by using the `show queueing` command, which shows the detail of the priority lists configured on the router and the appropriate details of each list.

**Note:** The command used to verify the queuing configuration is `show queueing` and not `show queuing`. The latter command is not recognized by the Cisco IOS.

### 1.7.5.5 Class-Based Weighted Fair Queuing (CBWFQ)

Class-based Weighted Fair Queuing (CBWFQ) extends the standard functionality of WFQ to provide support for user-defined traffic classes. It allows you to define traffic classes based on criteria such as protocols, access lists, and input interfaces. Once a class has been defined, you can assign it characteristics such as bandwidth, weight, and maximum packet limit. However, to characterize a class, you also specify the maximum number of packets allowed to accumulate in the queue, or the queue limit, for that class. Once a queue has reached its queue limit, the queuing of additional packets to the class causes tail drop or packet drop to take effect, depending on how class policy is configured.

If a default class is configured with the `bandwidth policy-map` class configuration command, all unclassified traffic is put into a single FIFO queue and treated according to the configured bandwidth. If a default class is configured with the `fair-queue` command, all unclassified traffic is flow classified and given best-effort treatment. If no default class is configured, then the traffic that does not match any of the configured classes is flow classified and given best-effort treatment. Once a packet is classified, all of the standard mechanisms that can be used to differentiate service among the classes apply.

For CBWFQ, the weight specified for the class becomes the weight of each packet that meets the match criteria of the class. Packets that arrive at the output interface are classified according to the match criteria filters you define, then each one is assigned the appropriate weight. The weight for a packet belonging to a specific class is derived from the bandwidth you assigned to the class. After the weight for a packet is assigned, the packet is queued in the appropriate class queue.

### 1.7.5.6 Low-Latency Queuing (LLQ)

Low-Latency Queuing (LLQ) allows you to implement strict priority queuing to CBWFQ. This gives delay-sensitive data preferential treatment over other traffic and ensures that delay-sensitive data is sent before packets in other queues. LLQ is configured by using the `priority` command. It enables the use of a single, strict priority queue within CBWFQ at the class level, allowing you to direct traffic belonging to a class to the CBWFQ strict priority queue. To queue class traffic to the strict priority queue, you must specify the named class within a policy map and then configure the `priority` command for the class.

### 1.7.6 Null Interface

Access lists are not always the most suitable solution to alleviate congestion. Access lists require CPU processing from the router. The more complex or the longer the access list, the greater the amount of CPU processing is required. The null interface, which is a virtual interface, is an alternative to access lists. Traffic may be sent to the null interface, but the traffic disappears because the interface has no physical layer. Thus, while access lists require CPU processing to determine which packets to forward, the null interface just routes the traffic to nowhere. The null interface is configured by using the following command syntax:

```
ip route ip_address subnet_mask null0
```

### 1.7.7 Fast, Autonomous, and Silicon Switching

Fast, autonomous and silicon switching techniques were created to improve the capability of the router to forward traffic at speed. After the routing process has made a routing decision, it sends the packet to the appropriate outbound interface. Meanwhile, the router holds a copy of the address details of the outbound frame in memory, along with a pointer to the appropriate outbound interface. This means that incoming

traffic can be examined as it comes into the router. The router looks in the cache to see whether a routing decision has already been made for that set of source and destination addresses. If an entry exists, the frame can be switched directly to the outbound interface, and the routing process is bypassed.

### **1.7.8 Cisco Express Forwarding (CEF)**

Another solution is Cisco Express Forwarding. This is very high-end solution and is available on 7500 routers with Versatile Interface Processors (VIPs) and the 8510 router. It is a distributed switching mechanism that keeps copies of route cache information in several different forms to be used for efficient switching. It is designed for high-performance, highly resilient Layer 3 IP backbone switching. It optimizes network performance and scalability for networks with large and dynamic traffic patterns, such as the Internet, on networks characterized by intensive Web-based applications, or interactive sessions. CEF offers three benefits: improved performance because it is less CPU-intensive than fast switching route caching; scalability; and resilience as CEF can switch traffic more efficiently than typical demand caching schemes.

Information conventionally stored in a route cache is stored in two data structures for CEF switching. The data structures are the **Forwarding Information Base (FIB)** and the **Adjacency Tables**, and provide optimized lookup for efficient packet forwarding.

- The **Forwarding Information Base (FIB)** is used to make IP destination prefix-based switching decisions. The FIB is similar to a routing table and maintains a mirror image of the forwarding information contained in the IP routing table. When routing or topology changes, the IP routing table is updated, and those changes are reflected in the FIB. The FIB also maintains next-hop address information based on the information in the IP routing table. Because there is a one-to-one correlation between FIB entries and routing table entries, the FIB contains all known routes and eliminates the need for route cache maintenance that is associated with switching paths.
- **Adjacency Tables** are used to maintain Layer 2 next-hop addresses for all FIB entries. The adjacency table is populated as adjacencies are discovered. Each time an adjacency entry is created; a link-layer header for that adjacent node is computed and stored in the adjacency table. Once a route is determined, it points to a next hop and corresponding adjacency entry. It is subsequently used for encapsulation during CEF packet switching. However, a route might have several paths, such as when a router is configured for load balancing and/or redundancy. For each resolved path, a pointer is added for the adjacency corresponding to the next-hop interface for that path. This mechanism is also used for load balancing across several paths.

CEF can be enabled in one of two modes: **Central CEF Mode**; and **Distributed CEF Mode (dCEF)**

- When in **Central CEF Mode** is enabled, the CEF FIB and adjacency tables reside on the route processor, and the route processor performs the express forwarding. This mode can be used when line cards, such as VIP line cards or GSR line cards, are not available for CEF switching or when you need to use features not compatible with distributed CEF switching.
- When **Distributed CEF Mode (dCEF)** is enabled, line cards maintain an identical copy of the FIB and adjacency tables. The line cards perform the express forwarding between port adapters, relieving the RSP of involvement in the switching operation. dCEF uses an Inter Process Communication (IPC) mechanism to ensure synchronization of FIBs and adjacency tables on the route processor and line cards.

### **1.7.9 Enhanced Interior Gateway Routing Protocol (EIGRP)**

## **CCNP/CCDP 642-891 (Composite)**

EIGRP a Cisco propriety routing protocol that is designed to make efficient use of the available network bandwidth. It can be used for IP, AppleTalk and IPX. EIGRP sends incremental updates, i.e., it sends updates only when a change in the network is experienced. EIGRP is particularly efficient in sending network and server information for client/server products such as NetWare for IPX and AppleTalk because it automatically redistributes routing updates into the local protocol updates. EIGRP is discussed in more detail in [Section 7](#).