EVODD

Enterprise Voice over Data Design

Version 3.3

Student Guide

Text Part Number:

Cisco Systems has more than 200 offices in the following countries and regions. Addresses, phone numbers, and fax numbers are listed on the Cisco Web site at www.cisco.com/go/offices.

Argentina • Australia • Austria • Belgium • Brazil • Bulgaria • Canada • Chile • China PRC • Colombia • Costa Rica • Croatia • Czech Republic • Denmark • Dubai, UAE • Finland • France • Germany • Greece • Hong Kong SAR • Hungary India • Indonesia • Ireland • Israel • Italy • Japan • Korea • Luxembourg • Malaysia • Mexico • The Netherlands • New Zealand • Norway • Peru • Philippines • Poland • Portugal • Puerto Rico • Romania • Russia • Saudi Arabia • Scotland • Singapore • Slovakia • Slovenia • South Africa • Spain • Sweden • Switzerland • Taiwan • Thailand • Turkey Ukraine • United Kingdom • United States • Venezuela • Vietnam • Zimbabwe

Course Introduction

Overview

The Enterprise Voice over Data Design (EVODD) course is designed to capture the breadth of technical issues surrounding the design of Voice over Data networks and provides a methodology that brings organization to the design issues faced by the systems engineer.

This three-day course covers both Cisco IOS and Cisco CallManager (CCM) voice network designs in depth. The course presents the technical issues of designing Voice over Data networks. Both sales and technical personnel will benefit from learning the Cisco methodology for implementing Voice over Data networks.

Major topics covered include: Voice over IP (VoIP), CCM, quality of service (QoS), gateways, legacy PBX integration, Cisco Unity, standards, dial plans, and migration strategies.

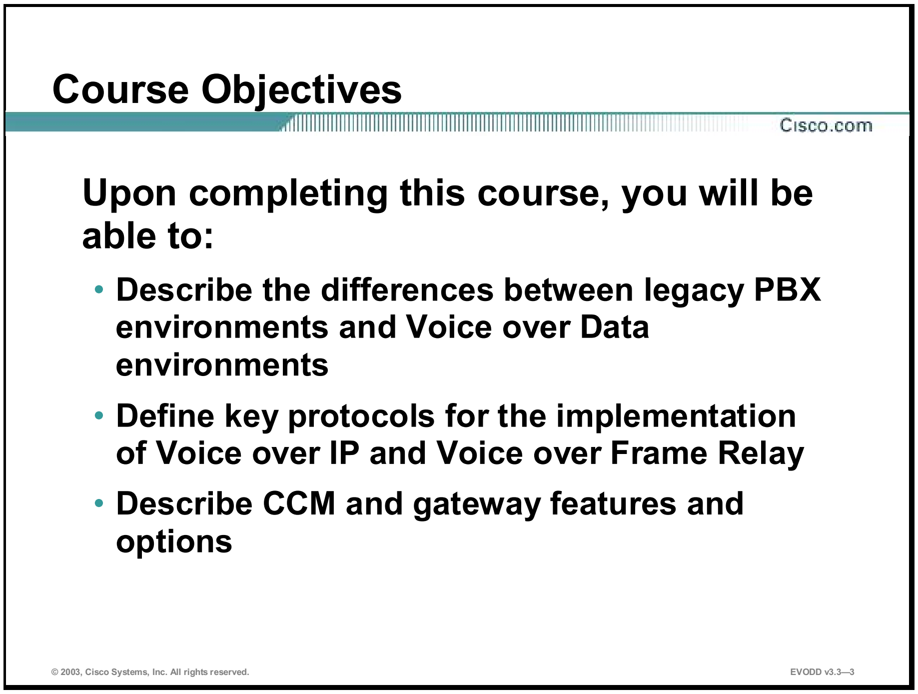
Outline

The Course Introduction includes these topics:

- Course Objectives
- Cisco Certification Track
- Learner Skills and Knowledge
- Learner Responsibilities
- General Administration
- Course Flow Diagram
- Icons and Symbols
- Sources of Information
- Learner Introductions

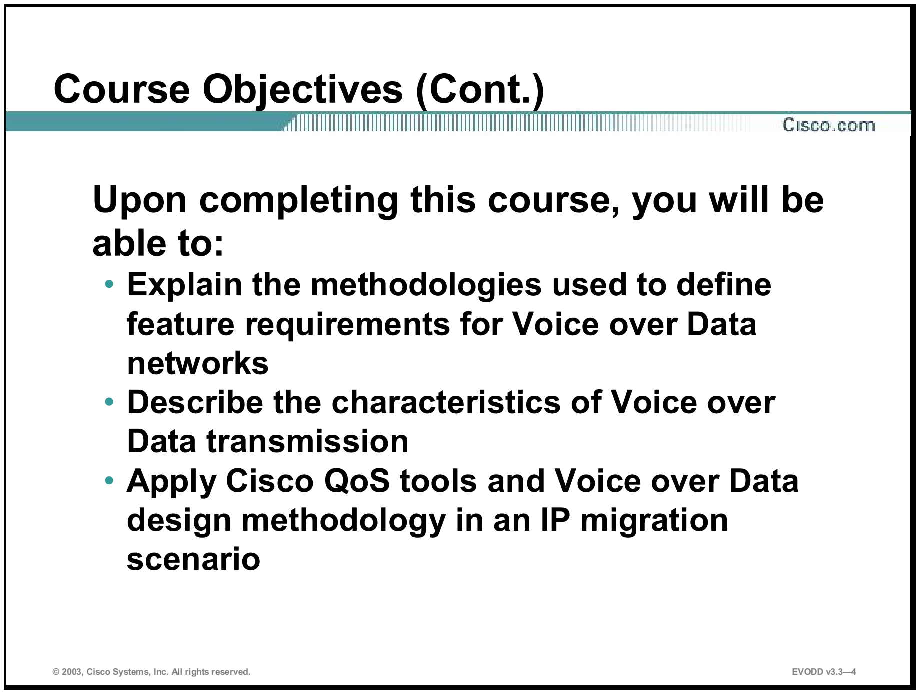
Course Objectives

This topic lists the course objectives.



Upon completing this course, you will be able to:

- Describe the differences between legacy PBX environments and Voice over Data environments
- Define key protocols for the implementation of Voice over IP and Voice over Frame Relay
- Describe CCM and gateway features and options

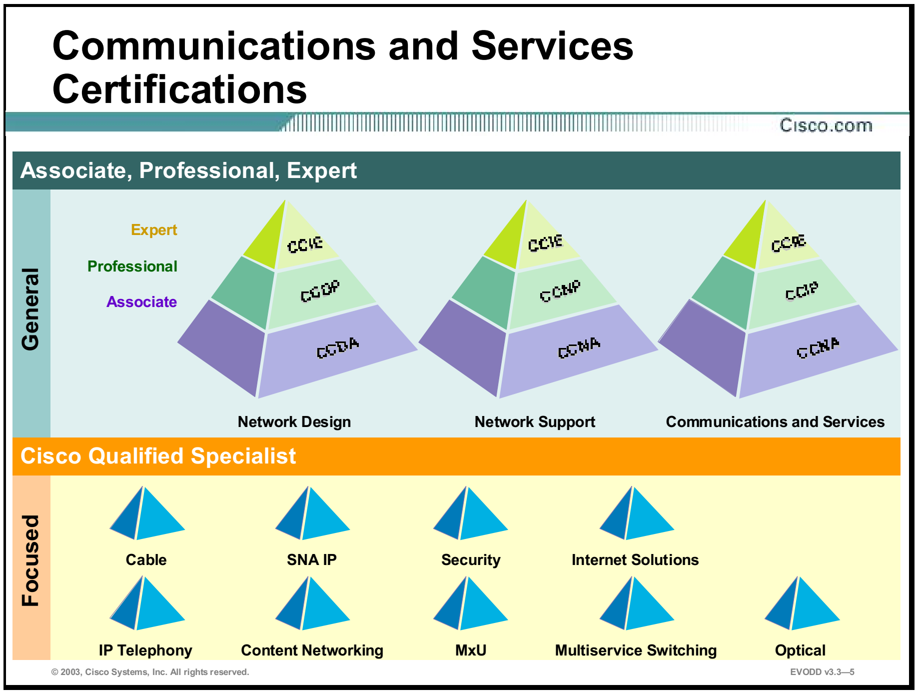


Upon completing this course, you will be able to:

- Explain the methodologies used to define feature requirements for Voice over Data networks
- Describe the characteristics of Voice over Data transmission
- Apply Cisco QoS tools and Voice over Data design methodology in an IP migration scenario

Cisco Certification Track

This topic lists the certification requirements of this course.



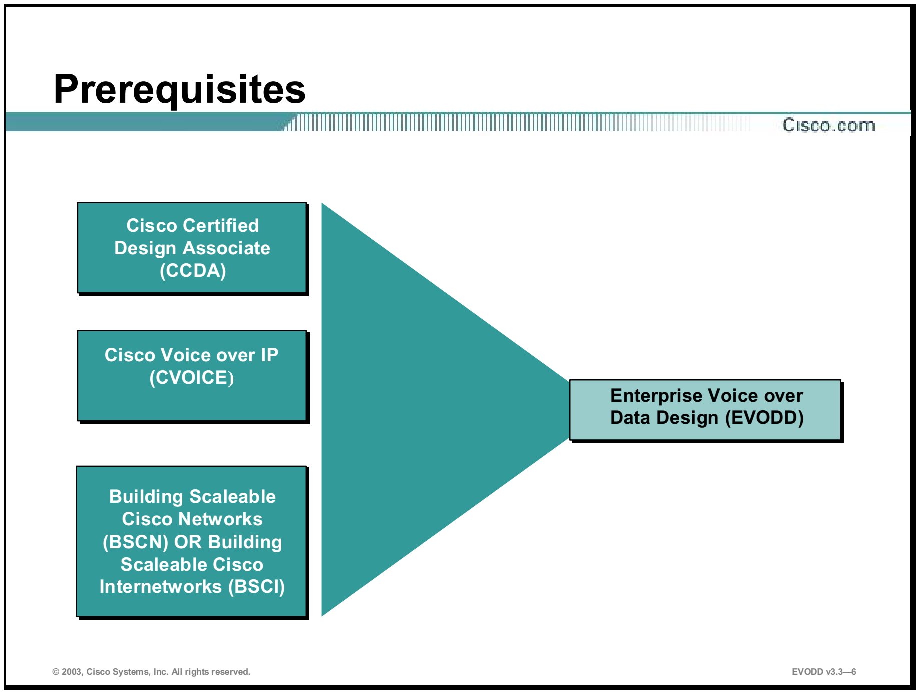
This education offering is a course focused on IP telephony, targeted towards the Cisco IP Telephony Design Specialist certification.

The *Cisco IP Telephony Design Specialist* certification is a Cisco Qualified Specialist (CQS) certification.

The requirements for the *Cisco IP Telephony Design Specialist* certification include the successful completion of the CCDA, DQoS, and EVoDD exams.

Learner Skills and Knowledge

This topic lists the course prerequisites.

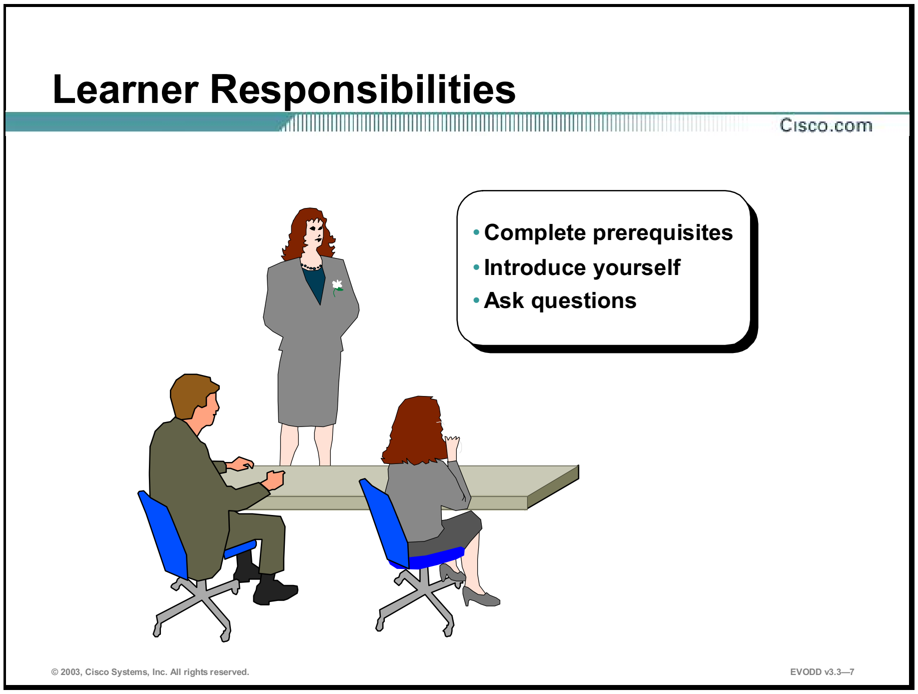


To fully benefit from this course, you must have these prerequisite skills and knowledge:

- Cisco Certified Design Associate (CCDA)
- *Cisco Voice over IP* (CVOICE)
- *Building Scalable Cisco Networks* (BSCN) or *Building Scalable Cisco Internetworks* (BSCI)

Learner Responsibilities

This topic discusses the responsibilities of the learners.



To take full advantage of the information presented in this course, you must have completed the prerequisite requirements.

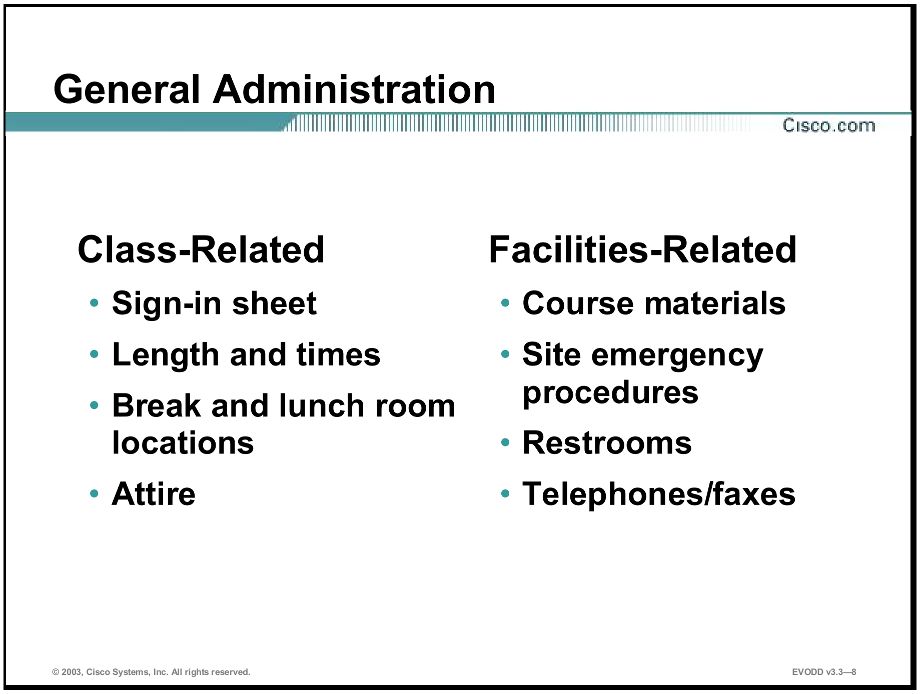
In class, you are expected to participate in all lesson exercises and assessments.

In addition, you are encouraged to ask any questions relevant to the course materials.

If you have pertinent information or questions concerning future Cisco product releases and product features, please discuss these topics during breaks or after class. The instructor will answer your questions or direct you to an appropriate information source.

General Administration

This topic lists the administrative issues for the course.



The instructor will discuss the administrative issues noted here so that you know exactly what to expect from the class:

- Sign-in process
- Starting and anticipated ending times of each class day
- Class breaks and lunch facilities
- Appropriate attire during class
- Materials you can expect to receive during class
- What to do in the event of an emergency
- Location of the rest rooms
- How to send and receive telephone and fax messages

Course Flow Diagram

This topic covers the suggested flow of the course materials.

Course Flow Diagram

Cisco.com

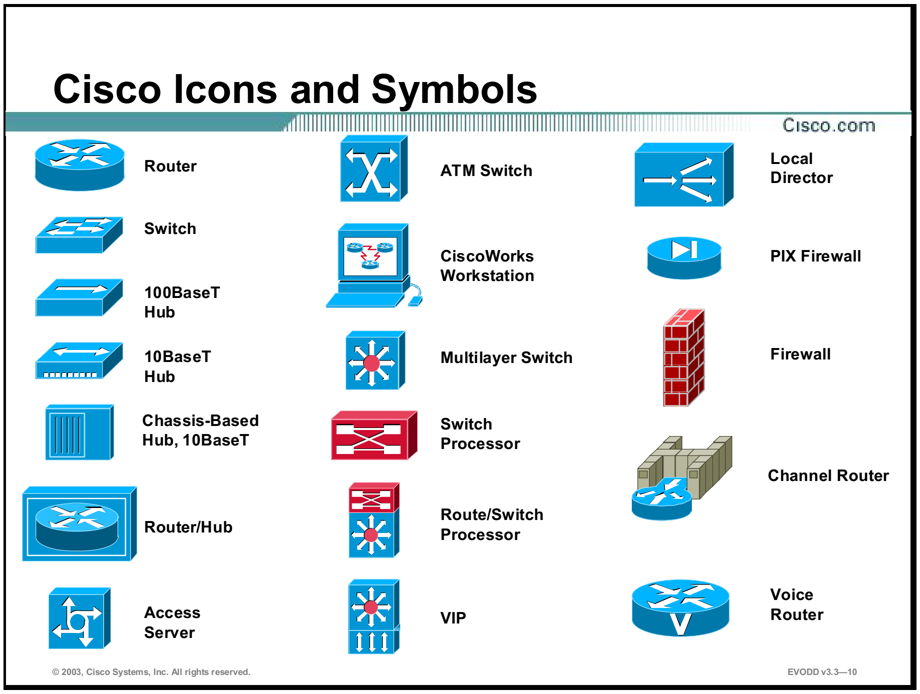
	Day 1	Day 2	Day 3
AM	Course Introduction	CallManager Overview and Gateway Selection	Voice Over Data Migration
AM	Voice Over Data Overview	Dial Plans and Voice Mail Considerations	
	Lunch		
PM	Voice Over Data Overview (Cont.)	Dial Plans and Voice Mail Considerations (Cont.)	Voice Over Data Migration (Cont.)
PM	Standards for Voice over IP Signaling	Voice Over Data Characteristics	

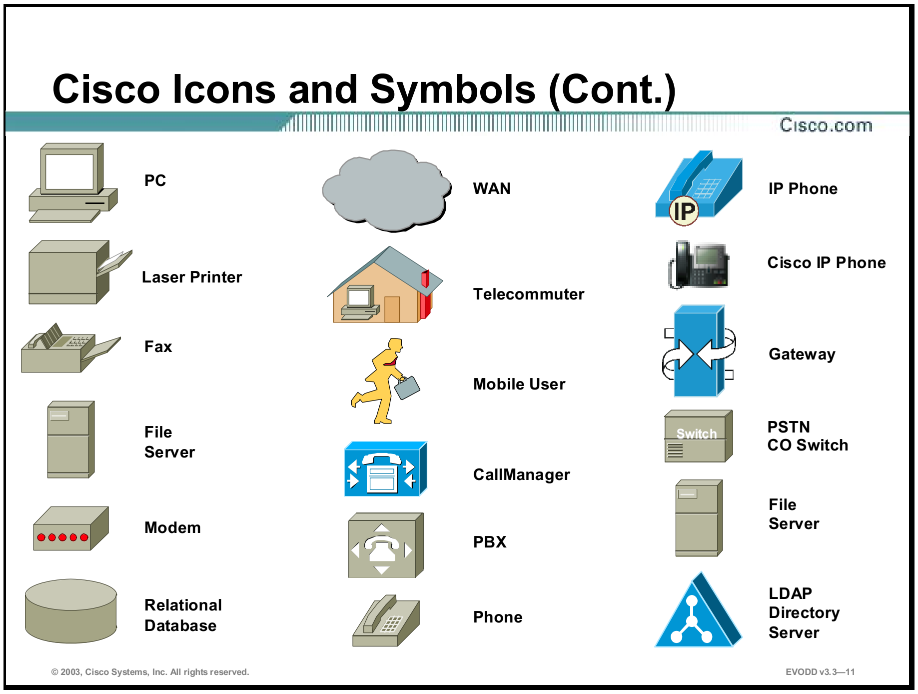
EVODD v3.3—9

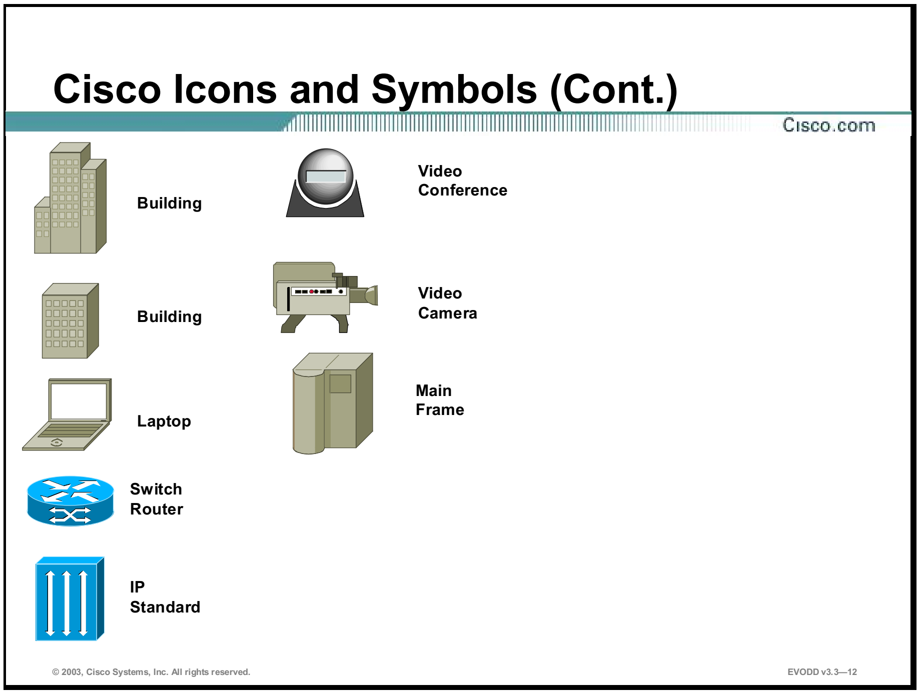
The schedule reflects the recommended structure for this course. This structure allows enough time for the instructor to present the course information and for you to work through the laboratory exercises. The exact timing of the subject materials and labs depends on the pace of your specific class.

Icons and Symbols

This topic shows the Cisco icons and symbols used in this course.

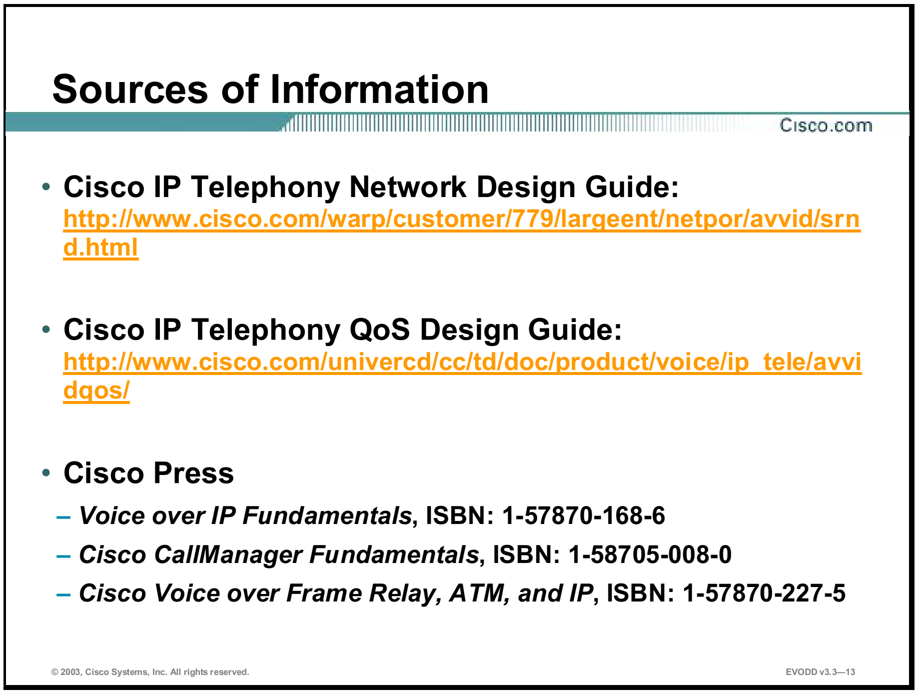



Cisco Icons and Symbols (Cont.)
Cisco.com
PC
Laser Printer
Fax
File
Server
Modem
Relational
Database
WAN
Telecommuter
Mobile User
CallManager
PBX
Phone
IP
IP Phone
Cisco IP Phone
Gateway
Switch
PSTN
CO Switch
File
Server
LDAP
Directory
Server
© 2003, Cisco Systems, Inc. All rights reserved.
EVODD v3.3—11


Cisco Icons and Symbols (Cont.)
Cisco.com
Building
Building
Laptop
Switch
Router
IP
Standard
Video
Conference
Video
Camera
Main
Frame
© 2003, Cisco Systems, Inc. All rights reserved.
EVODD v3.3—12

Sources of Information

This topic shows supplemental resources.

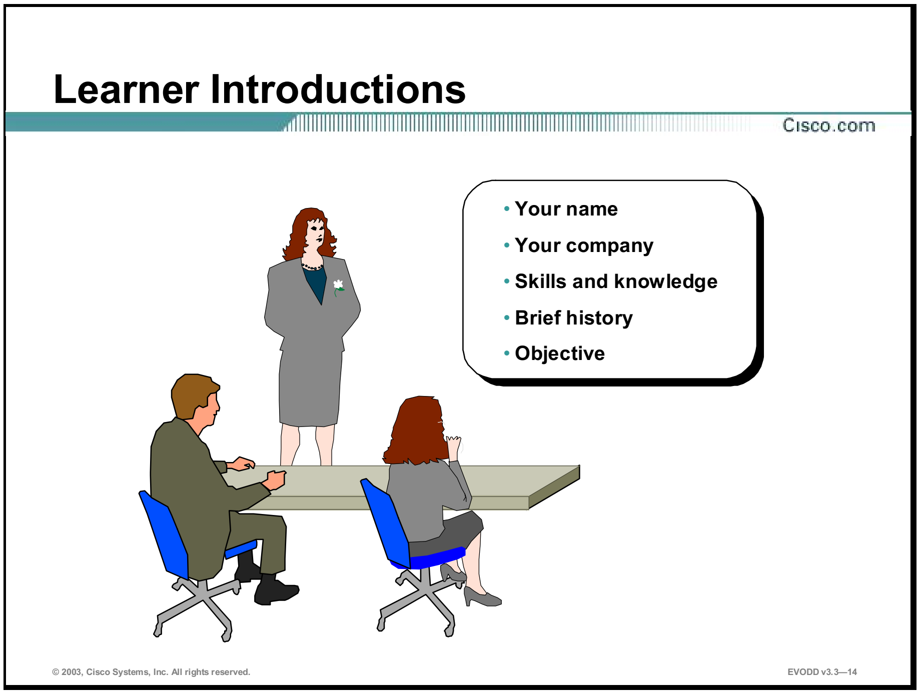


In addition to the course material, the following resources provide supplemental information regarding Voice over Data design:

- Cisco IP Telephony Network Design Guide: http://www.cisco.com/warp/customer/779/largeent/netpro/avvid/srnd.html
- Cisco IP Telephony QoS Design Guide: http://www.cisco.com/univercd/cc/td/doc/product/voice/ip_tele/network/
- Cisco Press
  - *Voice over IP Fundamentals*, ISBN: 1-57870-168-6
  - *Cisco CallManager Fundamentals*, ISBN: 1-58705-008-0
  - *Cisco Voice over Frame Relay, ATM, and IP*, ISBN: 1-57870227-5

Learner Introductions

This is the point in the course where you introduce yourself.



Prepare to share the following information:

- Your name
- Your company
- If you have most or all of the prerequisite skills
- A profile of your experience
- What you would like to learn from this course

Voice Over Data Overview

Overview

Voice over Data, also referred to as IP telephony, allows data, voice, and video to be transmitted over a single network infrastructure. To create a reliable Voice over Data design, you must have a sound understanding of the existing network. This module presents an overview of a legacy telephony system, a high-level view of IP telephony migration phases, and the various Voice over Data options. This module also presents a systematic approach for conducting a financial analysis for the migration of a project.

Upon completing this module, you will be able to:

- Describe an enterprise legacy PBX environment
- List the steps for planning an end-to-end IP telephony installation
- Describe the steps for router replacement
- Describe the process of delivering a complete Cisco IP telephony solution to a client
- Conduct a financial analysis of a proposed migration

Outline

The module contains these lessons:

- Enterprise and Public Switched Telephone Networks
- Voice Over Data Networks

Enterprise and Public Switched Telephone Networks

Overview

Enterprises are made up of dual infrastructures, one for data and another for voice. This lesson teaches you about basic PBX replacement and end-to-end IP telephony. You will learn how to implement an end-to-end IP telephony solution using a systematic approach. Note that while Cisco Architecture for Voice, Video and Integrated Data (AVVID) also encompasses the transmission of video across a data network, the *Enterprise Voice over Data Design* (EVODD) course focuses on the transmission of voice over an IP network.

Importance

To create a reliable Voice over Data design, you must have a sound understanding of the existing network.

Objectives

Upon completing this lesson, you will be able to:

- List characteristics of a legacy telephony network
- Identify the phases of IP telephony migration
- Explain route replacement in Voice over Data networking
- Describe a sample migration strategy
- List the options for Voice over Data deployments

Learner Skills and Knowledge

To fully benefit from this lesson, you must have these prerequisite skills and knowledge:

- A fundamental understanding of the services provided by legacy telephony systems
- A fundamental understanding of Cisco routing hardware features and functions

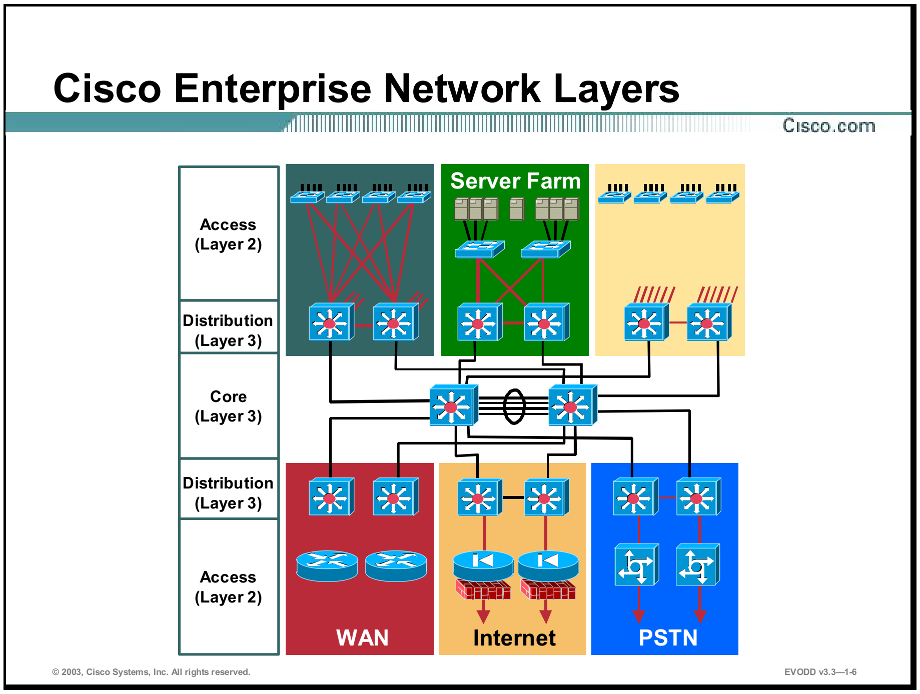
Outline

This lesson includes these topics:

- Overview
- Enterprise Network
- End Goal
- Route Replacement
- IP Telephony
- Voice Over Data Options
- Summary
- Lesson Review

Enterprise Network

This topic discusses enterprise network layers and correlates them to layers in the Open System Interconnection (OSI) protocol stack.

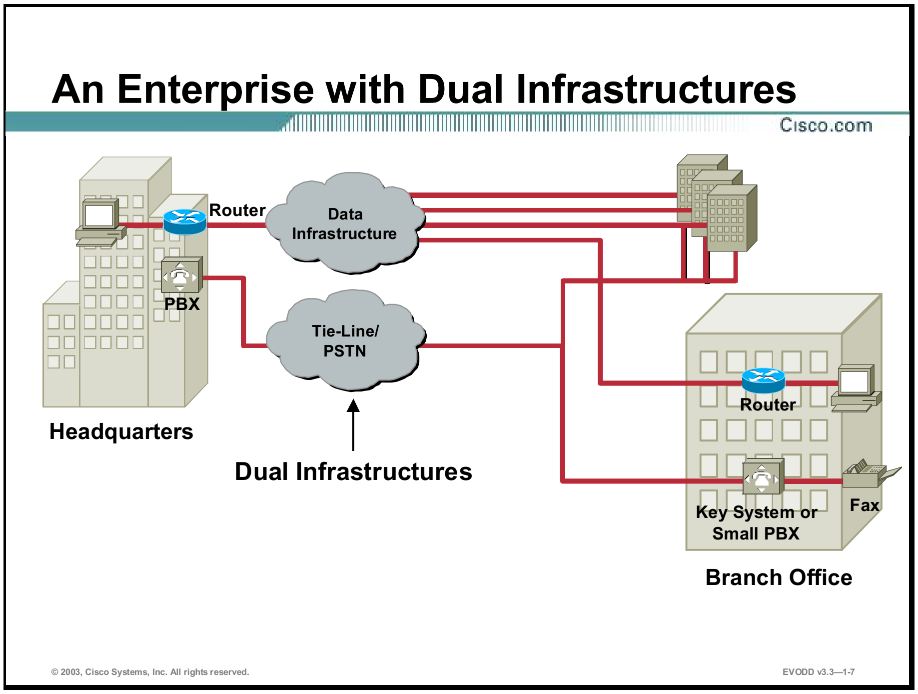


There are three enterprise network layers: access, distribution, and core. Although they are different than the protocol stack layers in the OSI network model, the two systems can be correlated. The figure shown here demonstrates the correlation.

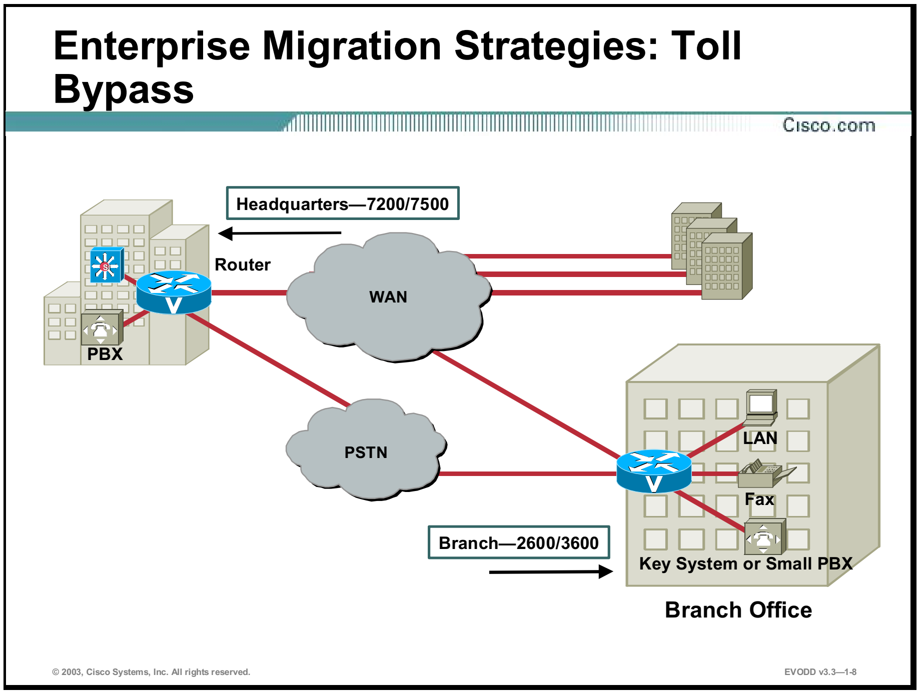
The Cisco access layer uses equipment that provides the OSI Layer 2 (data link) functionality. This equipment includes local routers, firewalls, gateways to the Public Switched Telephone Network (PSTN), and other access devices that are the closest endpoint. The access layer equipment also includes application servers and end-user desktop devices.

The Cisco distribution layer comprises equipment that supplies services corresponding to the OSI Layer 3 (network). This layer, the network edge, uses the following equipment: edge concentrators, campus routers, and smaller switches.

The Cisco core layer is generally populated with WAN backbone equipment that performs services corresponding to the OSI Layer 3. The main purpose of the core layer is to move information as quickly as possible; speed is the primary concern in the core layer. Equipment typically used at this layer includes such scalable solutions as the Cisco Catalyst 6500 switches and the Cisco 7200 and 7500 routers.



This figure illustrates a typical enterprise: a corporate headquarters and a branch office. The headquarters has a PBX voice network and a Cisco 7500 router. The branch office has a key system, or small PBX voice network, and a Cisco 2600 router. Enterprises support the expense of dual infrastructures, one for data and another for voice.



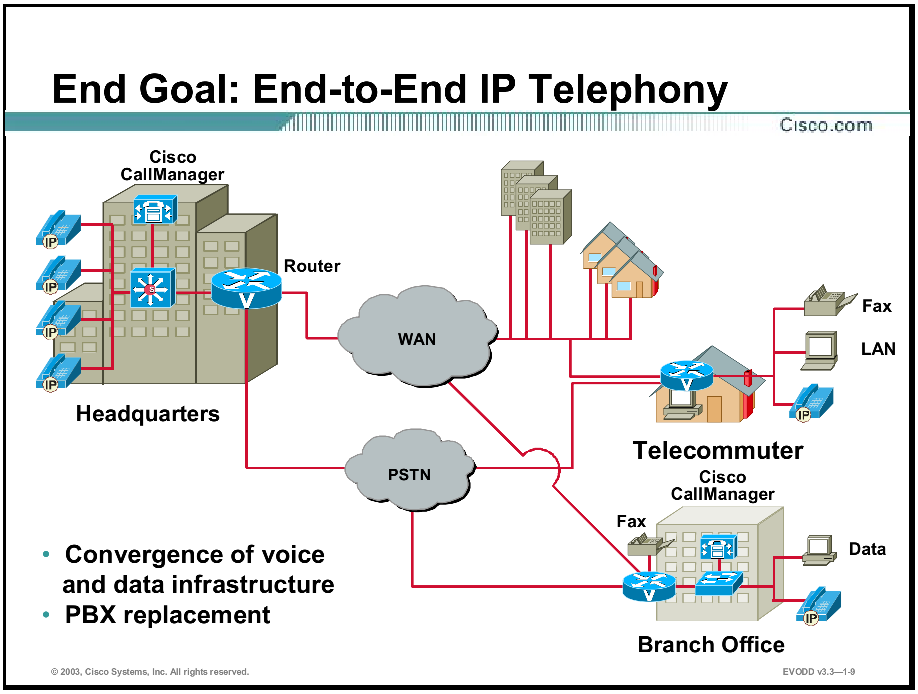
Although the Voice over Data design includes PBX replacement and end-to-end IP telephony, enterprises can take advantage of many interim strategies available for route replacement.

The figure shown here illustrates a typical enterprise that is replacing some voice lines by moving traffic to voice-enabled routers. This process, called toll bypass, allows customers to bypass the PSTN and use the packet network for long-distance (or toll) voice calls.

The enterprise maintains the PSTN for overflow traffic, as well as the PBX and key system for the branch office. This coexistence is common when IP telephony is installed in new branch offices, or when branch offices are replacing key systems.

End Goal

This topic discusses new telecommunications applications.



Voice over IP (VoIP), Voice over ATM (VoATM), and Voice over Frame Relay (VoFR) are all recent telecommunications applications. The challenge of integrating voice and data networks is fast becoming a priority for many network managers. Organizations want solutions that help take advantage of the excess capacity for voice and data transmissions on broadband networks, and use the Internet and company intranets as alternatives to more costly mediums.

Quality services and products require a voice access gateway to link the data and telephony networks. When sending voice over data networks, you must use this technology to ensure the quality of voice over streaming data.

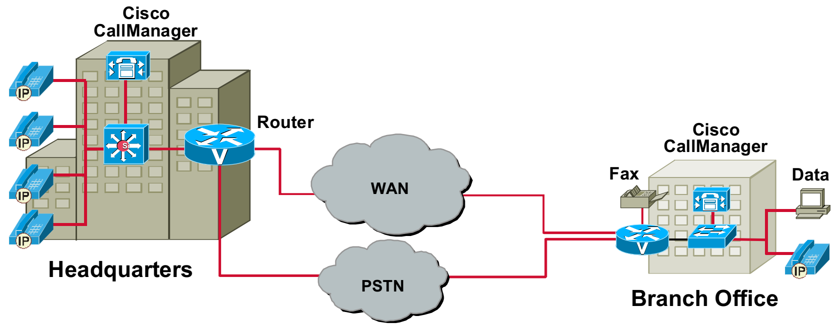
The voice-processing technology must handle greater and variable delays, cancel echoes from the telephony side so that the voice does not sound mechanical, and mask gaps caused by dropped packets during congestion.

The data network must adapt to variable networks and conditions, ensure the correct end-to-end connections, and handle the call setup translation for different types of networks, connections, and internetworking.

The goal of Voice over Data networking is the convergence of voice and data infrastructures into a single multiservice network. In a typical enterprise, this infrastructure will support voice, video, and integrated data, giving telecommuters the same access to corporate information services as onsite employees. Because of this technology, the enterprise is able to replace the PBX with an end-to-end IP telephony solution.

Enterprise Voice Over Data Building Blocks

Cisco.com



- PBX
- PSTN
- WAN
- Routers/Gateways
- Switches
- MCS
- QoS

 EVODD v3.3—1-10

A large organization must implement an end-to-end IP telephony solution in phases. Because each phase of the project builds upon the previous phase, the organization must understand the building blocks used throughout deployment and stay focused on the goal. The figure illustrates the location of each building block in the client solution. These building blocks include:

- **PBX:** Unless the organization is implementing a new installation, it must verify that its current PBX system is part of the new solution. Most organizations will implement the new solution slowly and leave the current PBX in place as they perform toll bypasses.

- **PSTN:** Because the organization is implementing a total Cisco IP telephony solution, it must have access to the PSTN. The organization uses the PSTN as a failover path, and to call locations that are not connected to its private network.

- **WAN:** The WAN is the current wide area data network for the client. In the figure shown, the client will also use it for calls between geographically dispersed locations.

- **Routers:** An organization must have routers in place as part of its data network. Some of these routers (such as the Cisco 2600 and 3600 routers) can support voice interfaces.

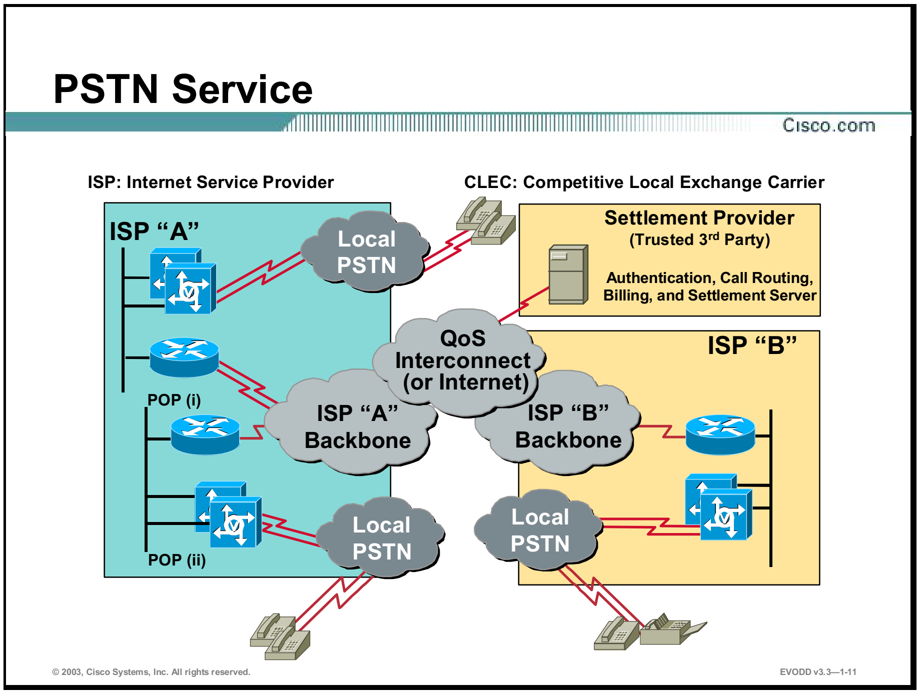
- **Switches:** When the data network transmits voice, the amount of bandwidth use increases. The network infrastructure must be able to handle the added traffic. Therefore, shared media hubs should be replaced with switches, due to collision issues. Also, some Catalyst switches support in-line power for Cisco IP Phones.

- **Media convergence server (MCS):** Cisco CallManager (CCM) software runs on the MCS. The MCS replaces the PBX and allows the administrator to implement a complete IP telephony solution.

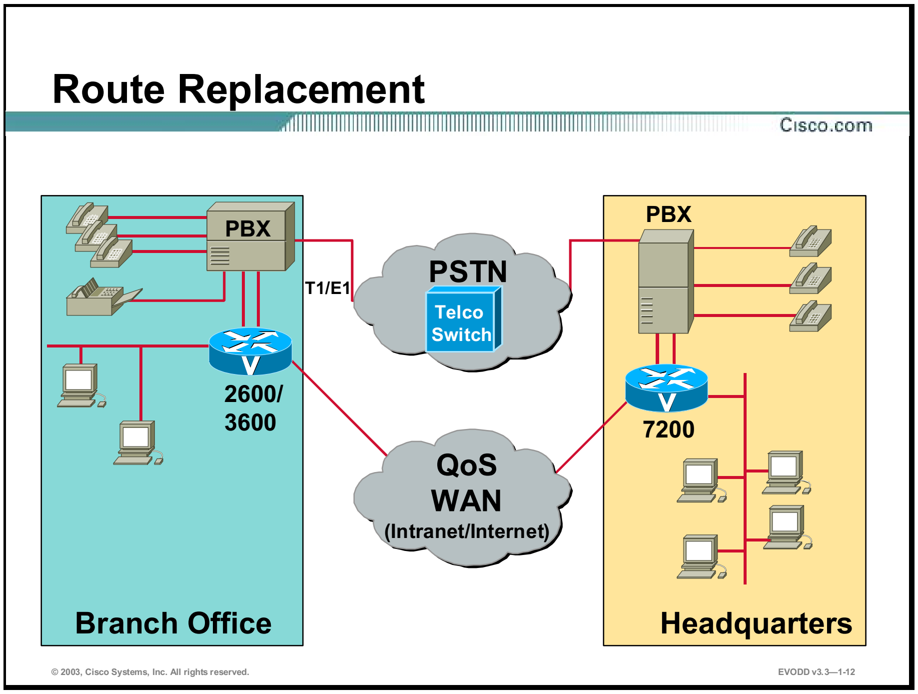
- **Quality of service (QoS):** As data and voice packets are moved to the same network, the organization needs to implement QoS to ensure that voice traffic has priority over data traffic, and that it moves in a timely fashion.

Route Replacement

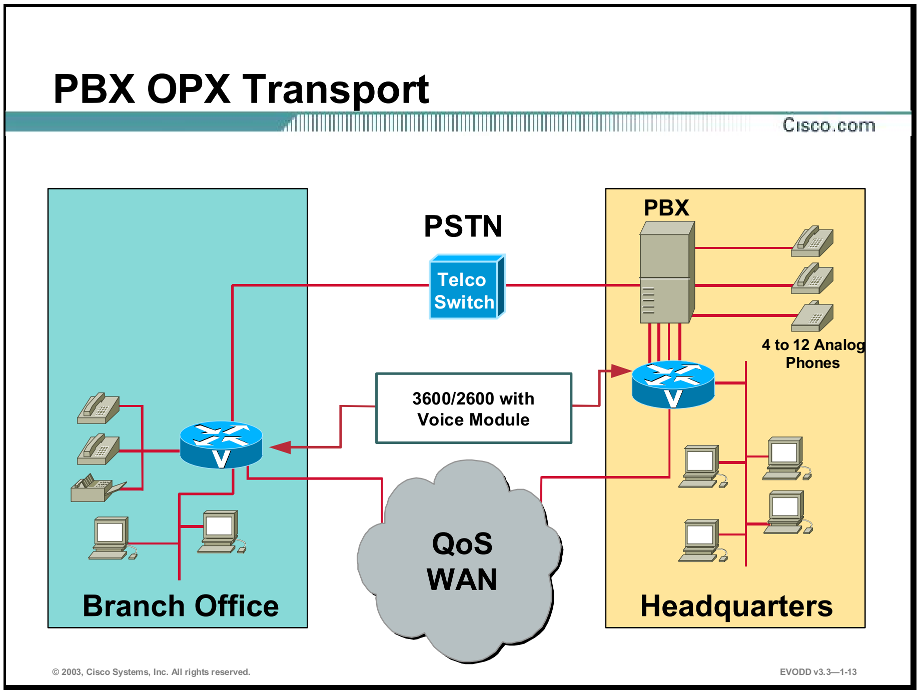
This topic focuses on IP WAN route replacement, an early step in IP telephony migration.



Internet telephony service providers and competitive local exchange carriers (CLECs) take advantage of data infrastructures and implement one of the IP telephony applications— they carry voice over their data networks. These new competitors (to traditional local exchange carriers [LECs] and long-distance companies) also use gateways to connect to the PSTN, relying on a trusted, third-party settlement provider to handle billing. Users set up voice rates as part of their Internet service provider (ISP) packages, creating a public alternative to the traditional PSTN.



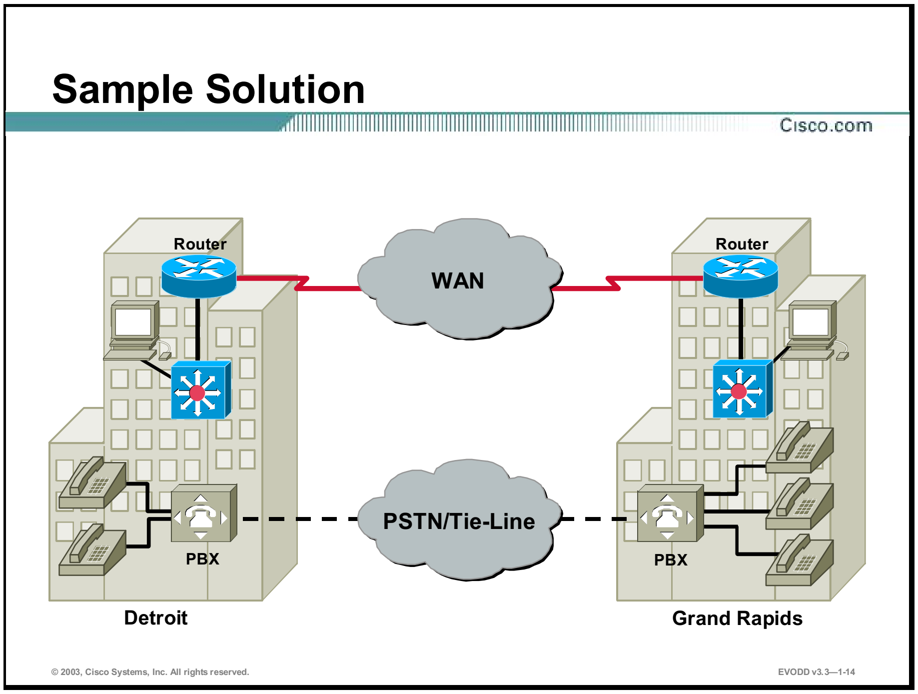
Route replacement is a common application for Voice over Data networking today. Traffic moves from either leased voice lines or the PSTN, and runs over the data infrastructure, where additional bandwidth handles the traffic. With data and voice converged on the same network, organizations use bandwidth more effectively and eliminate waste. To implement this strategy, an organization must have voice-enabled routers and updated software. The cost of this equipment is much less, however, than upgrading the legacy PBX hardware.



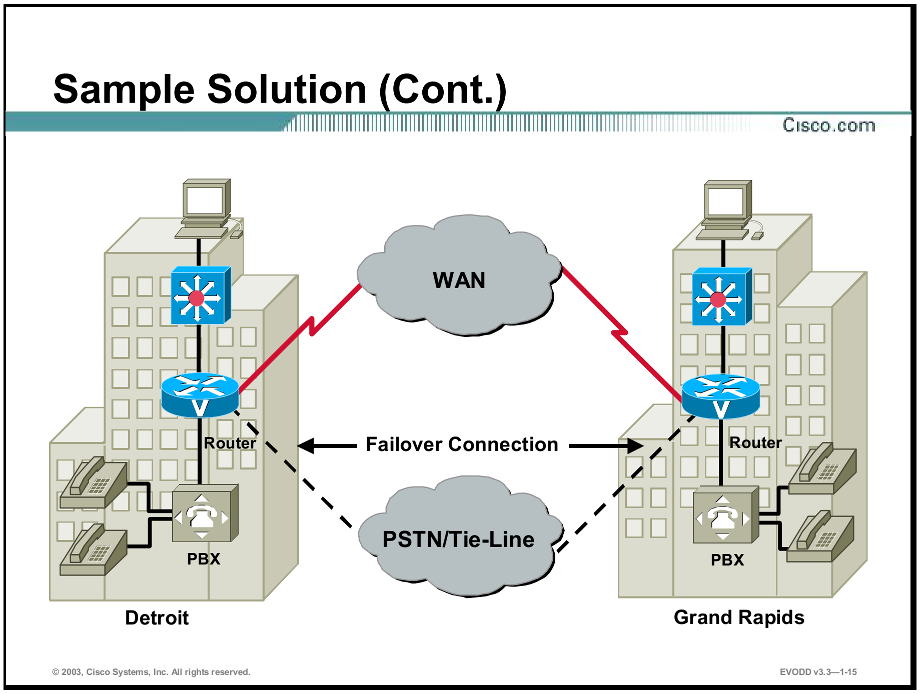
Another common application of Voice over Data is an Off-Premises eXtension (OPX) for a branch office. In the figure shown here, the branch office receives its dial tone from the PBX at the main office. The numbers and extensions appear as if they are located with the PBX, which is one way to avoid toll charges back to the main office. Voice modules are voice feature cards that go into a router or a switch to support voice services.

IP Telephony

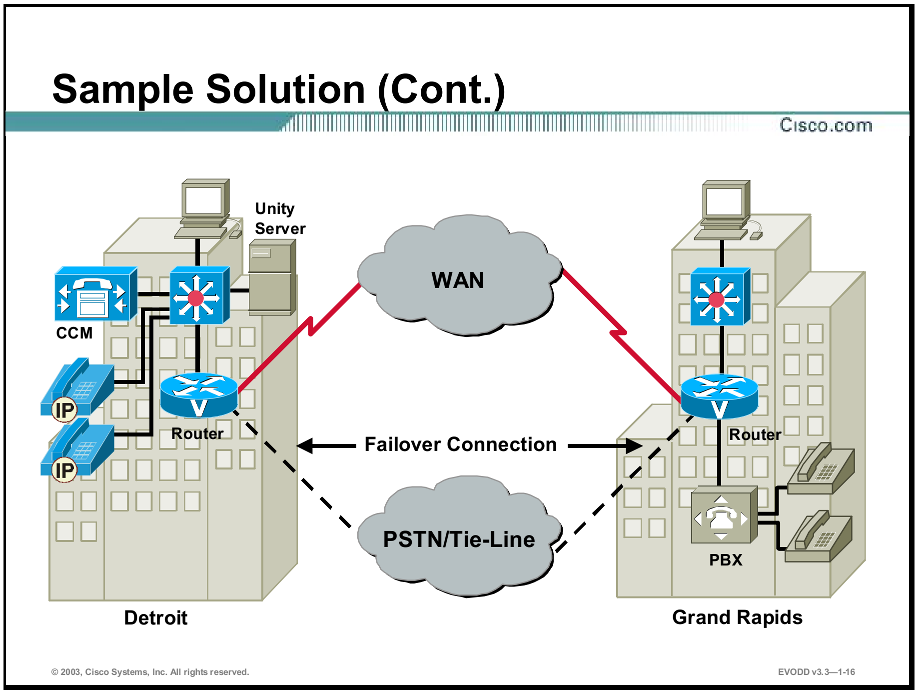
This topic illustrates the process of designing an IP telephony solution for a client.



A client asks you to design a Cisco IP telephony solution. The client is interested in implementing route replacement to test the reliability of this solution before moving to a complete Cisco IP telephony solution. The company has two locations with a PBX at each end. Each location has a WAN in place to access the main server located in Detroit, Michigan. Your goal is to deliver a complete Cisco IP telephony solution.



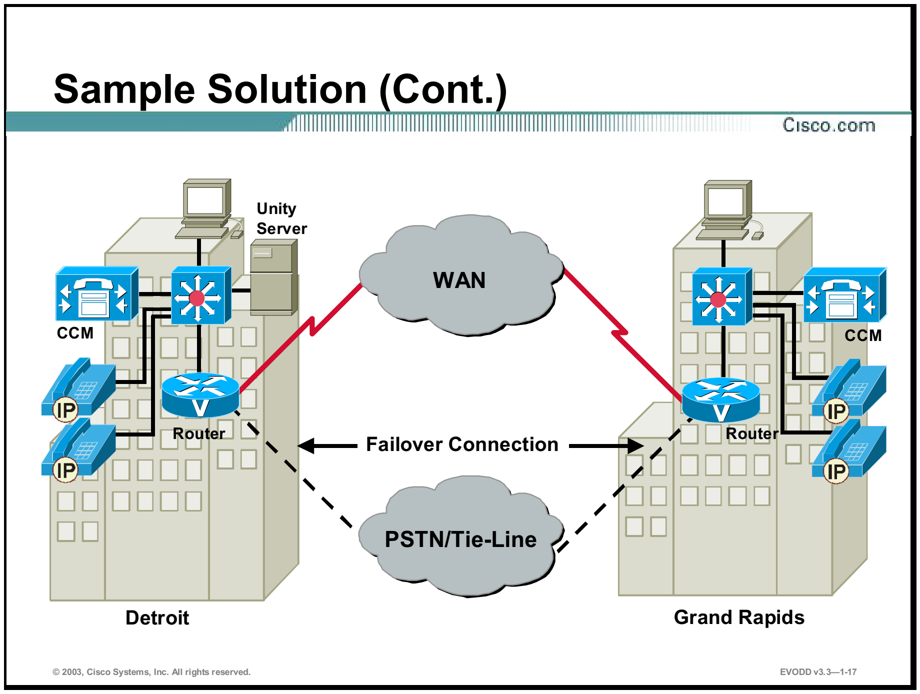
The client has three choices to implement route replacement: use the existing routers; upgrade the existing routers; or purchase new, voice-enabled routers. The PSTN connection still exists for failover purposes.



Router replacement is complete, and the satisfied client wants to convert one office to a complete Cisco IP telephony environment.

To create a Cisco IP telephony environment, you must install CCM. The CCM installation process includes IP Phones, because digital PBX telephones do not work in a Cisco IP telephony environment. Changes include the addition of the PSTN failover path on the Detroit side and the addition of a voice gateway. Because CCM is IP-based, it requires a gateway for converting voice that travels over the PSTN. This gateway is not always a stand-alone device. When the client is satisfied with the results, installing CCM in Grand Rapids, Michigan, is the next step.

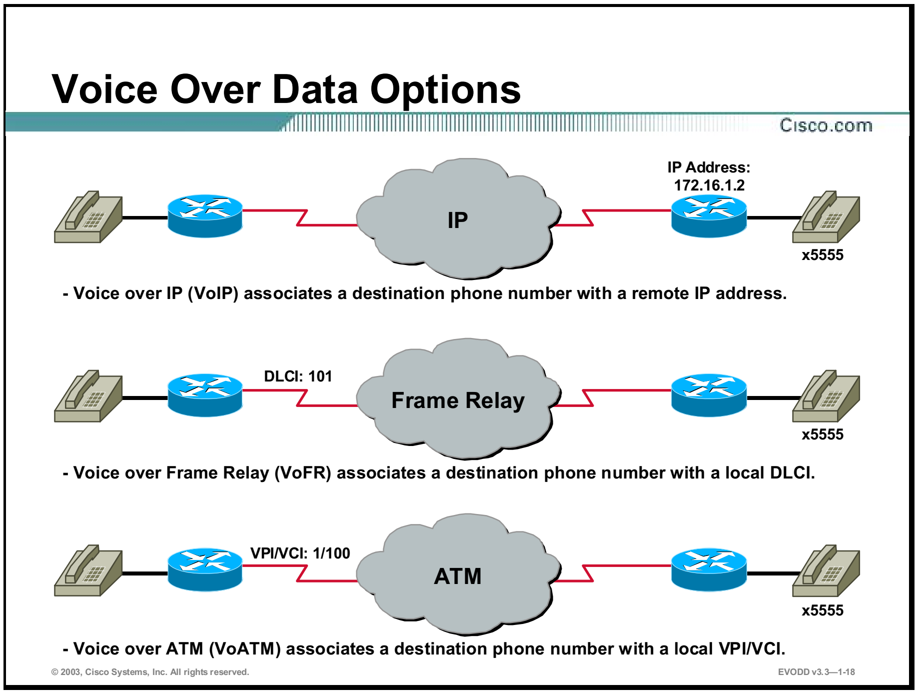
Note	Analog telephones can coexist in an IP telephony environment. In addition, Cisco supports shrink-and-grow migrations, in which legacy PBX systems and IP Phones coexist with (CCM).



After installing CCM and IP Phones in Grand Rapids, you can continue building the solution by installing a Cisco Unity server (unified messaging). Cisco Unity is just one example of an add-on application within the Cisco IP telephony environment. Cisco is developing more applications for Cisco IP telephony environments that will allow customers to continue building solutions to meet their specific needs.

Voice Over Data Options

This topic introduces other available Voice over Data transport options.



CCM uses VoIP, which sends voice traffic across a WAN by associating a destination telephone number with an IP address; for example, the IP address of the remote router that is directly attached to the telephone or to the PBX servicing the telephone. VoIP provides a homogeneous approach to voice transport, even if a company has multiple WAN technologies, such as VoFR, ATM, and PPP.

VoFR sends voice traffic across a Frame Relay WAN by associating a destination telephone number with a local data-link connection identifier (DLCI). A DLCI is a permanent virtual circuit (PVC) number that is recognized by the local router and the Frame Relay switch of the service provider. Because Frame Relay design did not originally support latency-sensitive voice transmission, you will need several QoS tools to preserve voice quality. However, VoFR transports voice using OSI Layer 2 (data link), so it may have less overhead than VoIP, which transports voice using Layer 3 (network).

VoATM sends voice traffic across an ATM WAN by associating a destination telephone number with a local ATM PVC identifier, such as a virtual path identifier (VPI) or virtual channel identifier (VCI). Unlike Frame Relay, the ATM design supports the latency-sensitive demands of voice and video; therefore, VoATM is easier to configure for voice traffic than for data traffic. Because VoATM uses Layer 2 for voice transport, it has less overhead than VoIP.

Summary

This topic summarizes the key points discussed in this lesson.

Summary

Cisco.com

- Legacy telephony networks operate separately from data networks.
- The migration from a legacy telephony network to a Voice over Data network occurs in phases: Cisco CallManager gradually replaces PBX.
- Although VoIP is the most popular Voice over Data technology, other options include VoFR and VoATM.
- To create a Cisco IP telephony environment, Cisco CallManager must be installed.
- Cisco CallManager uses VoIP, which sends voice traffic across a WAN by associating a destination telephone number with an IP address.

 EVODD v3.3—1-19

Next Steps

After completing this lesson, go to:

- Voice over Data Networks lesson

References

For additional information, refer to these resources:

- *Technical Considerations for Converging Data, Voice, and Video Networks:* http://www.cisco.com/warp/public/cc/so/neso/vvda/avvid/tecon_wp.htm
- *Voice-Enabled WAN:* http://www.cisco.com/warp/public/779/largeent/design/voice_wan.html

Lesson Review

This practice exercise reviews what you have learned in this lesson.

Q1) What is the role of toll bypass?

- A) to allow telephone calls over the IP WAN if the PSTN becomes congested
- B) to allow telephone calls over the IP WAN to avoid the PSTN charges
- C) to allow telephone calls over the PSTN if the IP WAN becomes congested
- D) to allow telephone calls over the PSTN to avoid the IP WAN charges

Q2) Which of the following is NOT considered a Voice over Data building block?

- A) PBX
- B) PSTN
- C) Unity
- D) QoS

Q3) Which of the following best describes an OPX?

- A) An OPX is an on-premises extension located at the main office.
- B) An OPX is an off-premises extension located at a branch or home office.
- C) An OPX is an outgoing PBX extension located at the main office.
- D) An OPX is an outgoing PBX extension located at the branch or home office.

Q4) After the route replacement phase, which of the following would be a logical next step in a migration from a legacy PBX system to an IP telephony solution?

- A) removing the legacy PBXs
- B) connecting the digital telephones of the PBX to a Cisco Catalyst switch
- C) installing CCM for toll bypass
- D) connecting IP Phones to the PBX

Q5) Which of the following Voice over Data technologies associates a destination telephone number with a local DLCI?

A) VoIP

B) VoATM

C) VoHDLC

D) VoFR

Voice Over Data Networks

Overview

Deployment of Voice over Data solutions is divided into four stages: plan, design, implement, and operate. The plan and design stages are the most applicable to this lesson. Completing all of the steps within each stage is essential for success. This lesson teaches you how to plan a design methodology and how to apply this design methodology to customer problems.

Importance

This lesson benefits learners who want to increase their understanding of the key customer motivators for a Voice over Data network. In addition, this lesson provides information about clarifying the operation and performance goals for voice applications to meet customer expectations.

Objectives

Upon completing this lesson, you will be able to:

- Identify the deployment stages for Voice over Data solutions
- Identify the key considerations for planning a Voice over Data solution
- Identify the data-related and voice-related critical factors in an existing customer network
- Recognize pertinent voice traffic parameters and quantify capacity requirements
- Identify and justify costs of the Voice over Data solution
- Identify a poorly performing Voice over Data network
- Apply the Voice over Data network design methodology to real-life customer problems
- Interview customers to characterize customer costs
- Explain how to use the cost-savings model tool

- Describe the output of the cost-savings model tool, keeping model limitations in context

Learner Skills and Knowledge

To fully benefit from this lesson, you must have these prerequisite skills and knowledge:

- A fundamental understanding of the services provided by legacy telephony systems
- A fundamental understanding of Cisco routing hardware features and functions

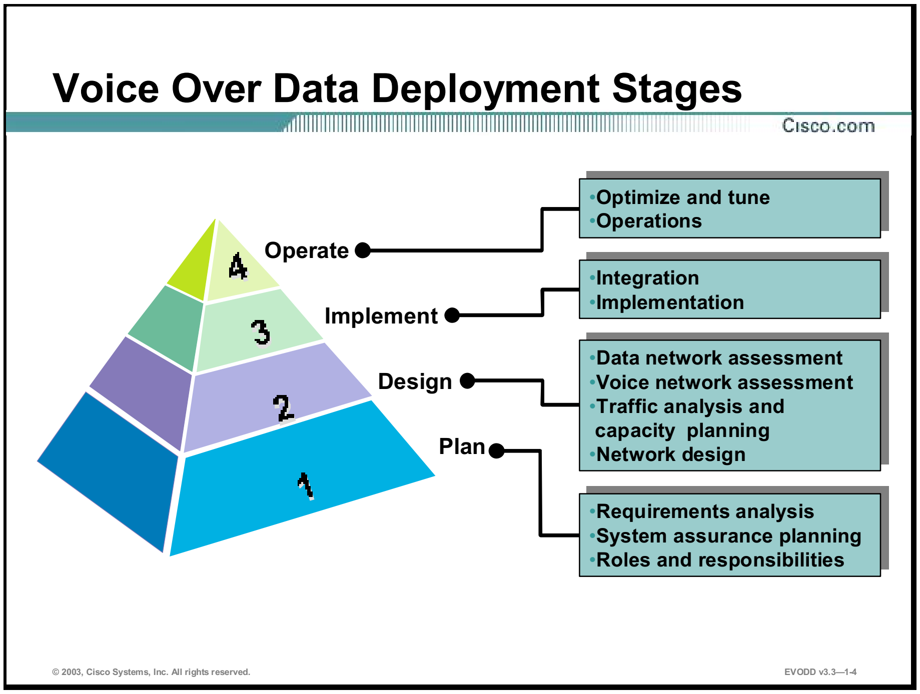
Outline

This lesson includes these topics:

- Overview
- Deployment Stages
- Plan a Design
- Network Design
- Traffic Analysis and Capacity Planning
- Cost Documentation
- Voice and Data Costs
- Other Cost Considerations
- Input Information
- Standard Model Output
- Summary
- Laboratory Exercise: Calculating Trunk Capacity
- Lesson Review

Deployment Stages

This topic introduces the four stages of Voice over Data deployment: plan, design, implement, and operate. You must complete specific steps in each of theses stages for success.



The planning stage focuses on gathering customer requirements, ensuring that their current network infrastructures can support efficient voice transport, and assigning roles and responsibilities between the customer and vendor.

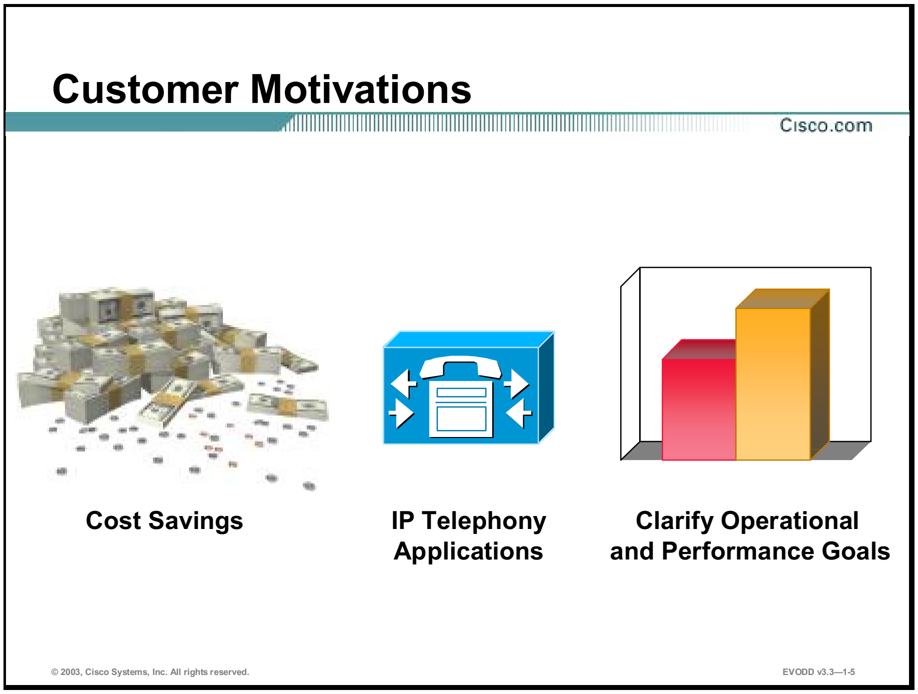
The design stage consists of all of the steps necessary to successfully complete a rigorous design of the new converged network. This stage requires recording information about the existing data and voice networks and planning the necessary bandwidth increases to carry additional voice traffic. Final network design includes infrastructure component changes, such as the deployment of switches or redundancy designs for Cisco CallManager (CCM).

The implementation stage is the domain of engineers, who focus on configuring the selected products and services.

The last stage, operate, requires optimizing the solution to ensure that all components work smoothly; for example, this stage may include the introduction of echo cancellation.

Plan a Design

This topic provides an in-depth discussion of the planning stage.



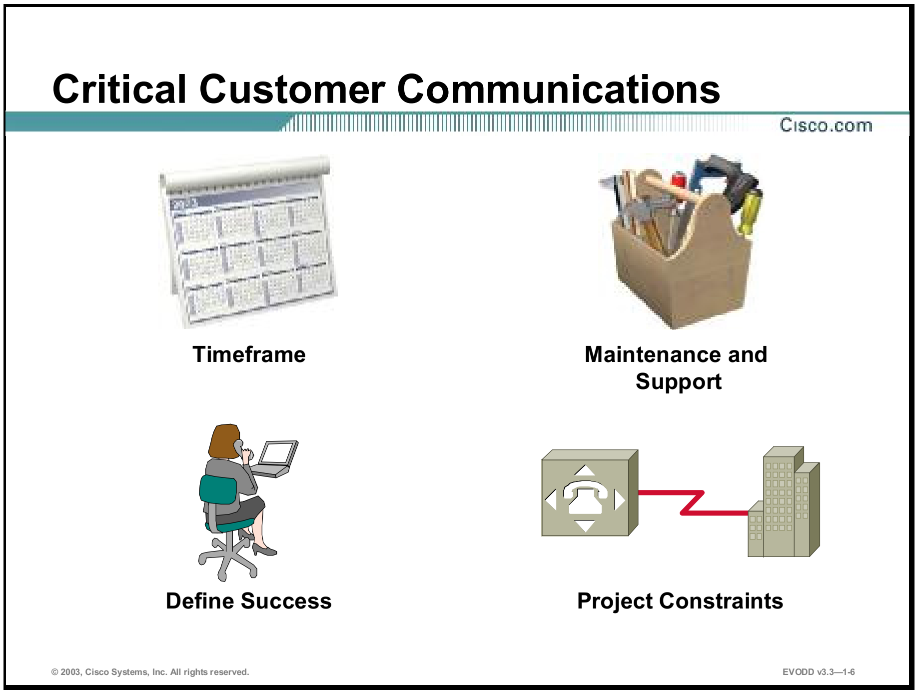
Before the requirements process begins, you should clarify the key motivators from the customer perspective. Motivators for IP Telephony are: a single network to own and administer, simpler adds/moves/ and changes, (less staff required – could be a touchy subject), and open standards allowing rapid development and implementation of applications. There are many key motivators for deploying Voice over Data. The following list represents important key motivators:

- **New applications:** Organizations want access to new applications, such as unified messaging. These applications usually have a data and data networking orientation.

- **Telephone mobility:** Organizations also need flexibility. Telephone mobility makes it easy for organizations to move their telephones to new offices or to connect from home using a software-based telephone (such as Cisco SoftPhone) on personal computers.

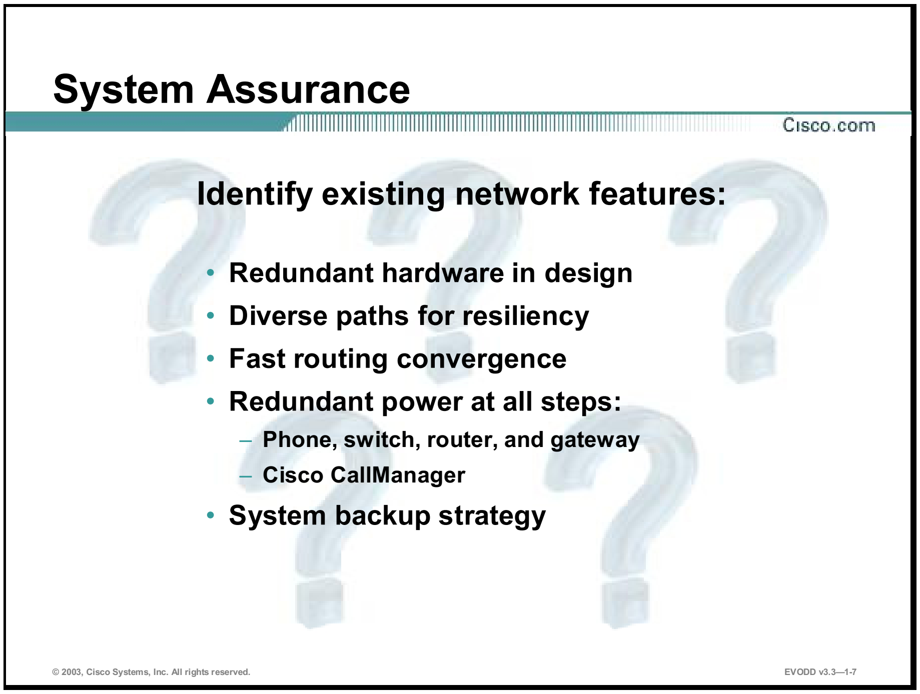
- **Remote office receptionist:** An organization can have all incoming calls answered by one receptionist, whose office may be at any of the company locations or a remote location in another city. The company is able to route these calls, via its intranet, while providing quality customer service at a reduced cost.

An equally important part of planning is clarifying customer expectations around the operation and performance of voice applications. You should ask questions about frequently used features and those considered most valuable, because customers often base their service expectations on existing PBX vendors. Managing customer expectations by advising the customer about the possibility of losing a valued feature is also essential. You will need to continue this practice throughout the deployment of other Cisco IP telephony solutions



Customers typically have an expectation of shorter response and repair times for voice problems than for data problems. With proper redundancy designed into the network, there should never be an outage that disrupts the entire voice network. However, the possibility exists, especially with 2:1 redundancy and oversubscribed SRT that an outage can occur and affect many people. The customer should have a plan in place for those rare possibilities.

Customers will define the success of their projects differently. Success for one customer can be the ability to scale the solution quickly to a large number of users in a short timeframe. Interoperability with existing PBX systems can be a key success criterion for another customer. Because customers typically use a combination of factors to define success, their feedback will provide guidelines for the design process. You should also discuss project constraints as early as possible. If there are no standards-based interfaces on the PBX for connecting Cisco products, project success will be more difficult, and maybe even impossible, to accomplish.



System assurance describes the minimum system features required to ensure a functional, converged network. The existing customer network must implement these features to support the reliable deployment of Voice over Data.

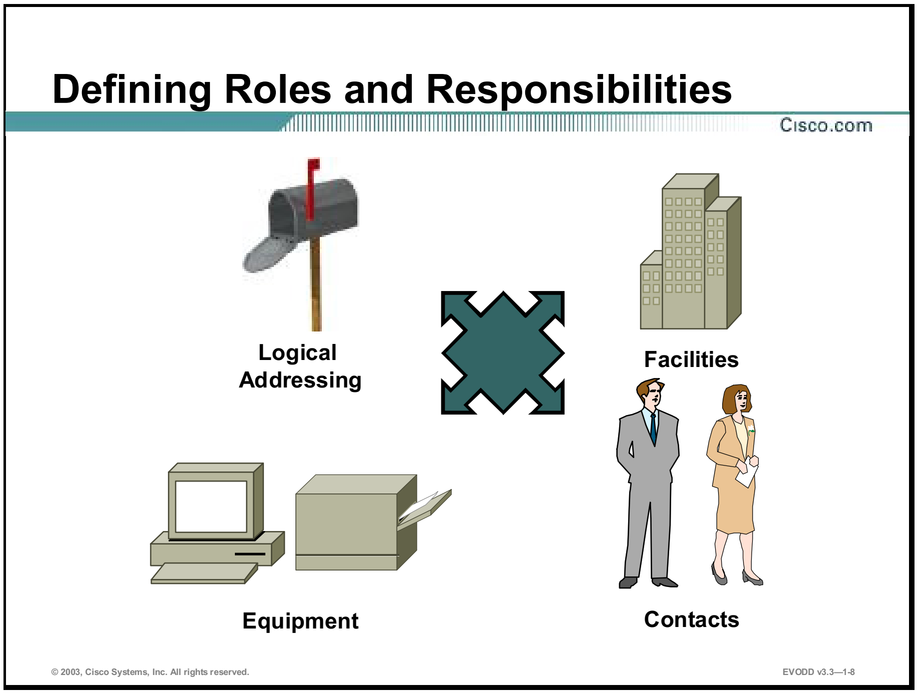
Redundancy is key to deploying successful, converged networks.

Redundancy begins at the hardware level, with redundant power supplies for telephones, switches, routers, and gateways. Platforms running CCM should have redundant power options, and routers and switches should have a redundant configuration design. Running protocols, such as the Multigroup Hot Standby Router Protocol (MHSRP), provides router redundancy, enabling recovery from total router failure.

Reliable Voice over Data solutions require diverse paths to ensure that the network can find an alternative route if there is a path failure. Dynamic routing algorithms, such as the Enhanced Interior Gateway Router Protocol (EIGRP) and Open Shortest Path First (OSPF), automatically route around path failures.

You should use rapid routing protocol convergence to avoid delays when the network is recovering from path failures. You should also identify the current routing protocols and the typical convergence times.

You must identify the system backup strategy for current servers and network devices. Make certain that the customer understands that individuals cannot make voice calls during the reconfiguration of a router. CiscoWorks provides backup tools for routers that allow reconfiguration in seconds.



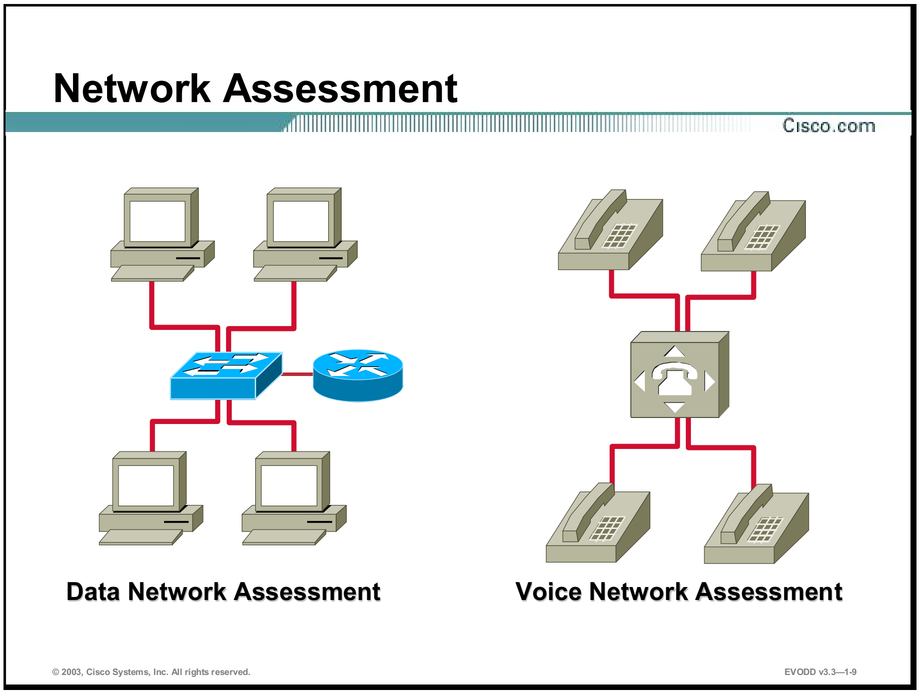
Because the customer and vendor expectations of roles and responsibilities may differ, you should document customer responses to the following questions during the planning stage:

- Who controls each part of the network?
- Who prepares the facilities?
- Who handles the equipment staging?
- Who develops the address and numbering plans?
- Who upgrades the existing systems?
- Who are the site and department contacts? If the customer expects the vendor to stage the equipment, a price increase is likely. You should clarify this point with the customer during the planning process to avoid hidden or unexpected costs.

IP addresses, telephone numbers, and dial-plan designs are critical for the successful deployment of a Voice over Data solution. The customer should understand all of the issues related to this topic and be prepared to perform the appropriate tasks.

Network Design

This topic describes all of the steps necessary to complete the rigorous design of a new, converged network.

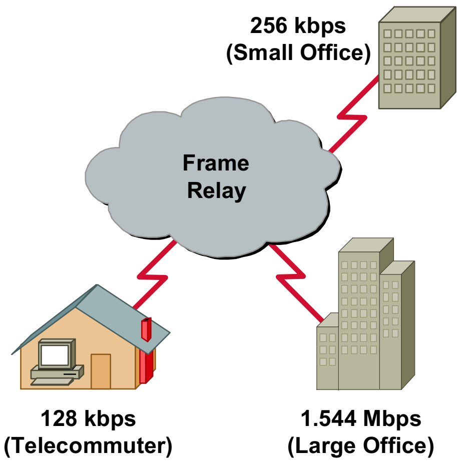


Network assessment is divided into two areas: data network assessment and voice network assessment. Data network assessment includes identifying WAN types, and understanding traffic characteristics and data network costs. The data network also determines any necessary IP telephony considerations. For example, shared media hubs do not provide the collision protection required for IP telephony. The voice network assessment is very similar for the parallel voice network.

Data Network Assessment: WAN Media

Cisco.com

- **Existing WAN media can influence:**
  - Selection of voice/data transport
  - Voice/data bandwidth
  - Selection of QoS features
  - Need for virtual networking services (Frame Relay and ATM)

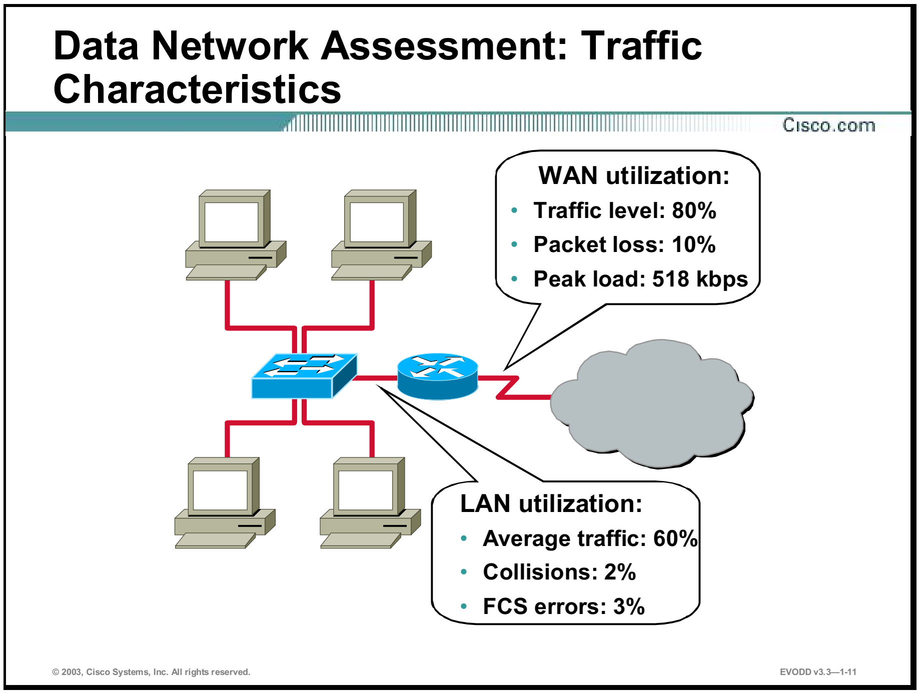


EVODD v3.3—1-10

WAN media can influence several areas of the design process. If a customer uses Frame Relay and does not require an IP telephony solution, then you may suggest Voice over Frame Relay (VoFR) as an immediate solution.

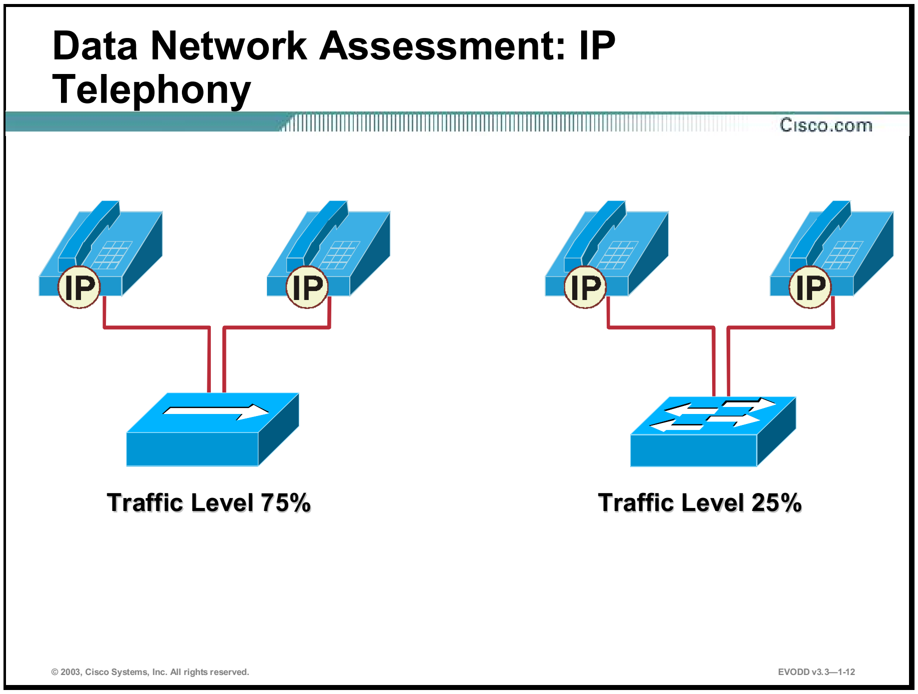
WAN media also influences the bandwidth and quality of service (QoS) considerations required to support the voice call. A Voice over IP over Frame Relay (VoIPovFR) call has different bandwidth requirements than a Voice over IP over ATM (VoIPovATM) or Voice over ATM (VoATM) call. WAN media also influences the QoS techniques applied at the WAN interface level.

If current PBXs are providing tandem-switching functions, you should consider using voice network switching with Frame Relay and ATM networks. However, the future of Voice over Data is Voice over IP (VoIP) because of the universal adoption of TCP/IP by other data applications and its compatibility with them.



Sufficient bandwidth for the existing data traffic is necessary to prevent overloading the network. Adding voice to an existing overloaded network produces poor voice quality. You should sample the network at key points, such as access, distribution, and core, as well as at peak load times.

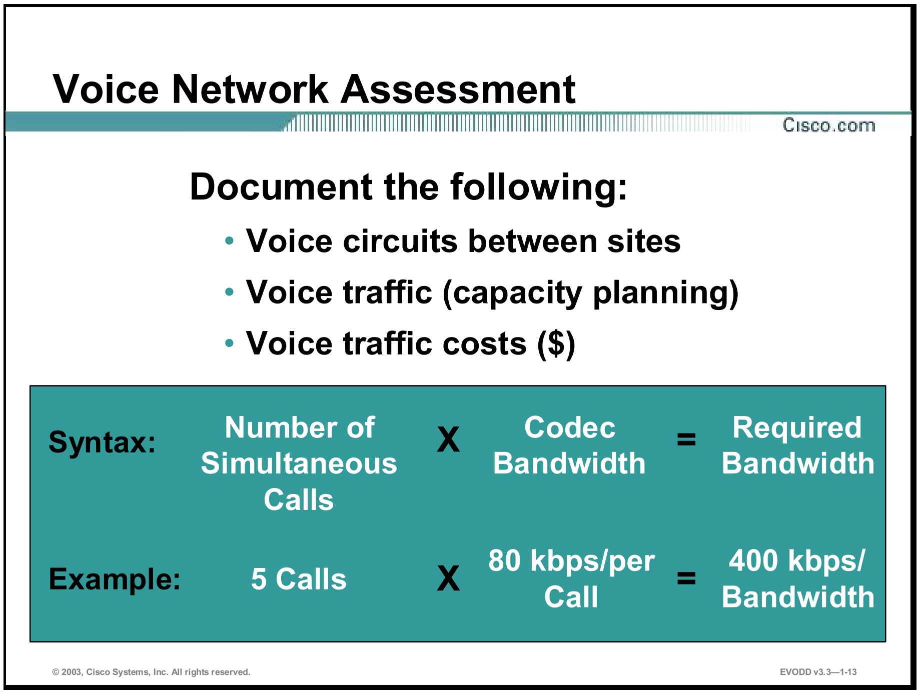
You need to verify the application mix because bursty application traffic may affect voice quality. When gathering traffic statistics, measure packet loss on the WAN links, and collision rates on the LAN. Ideally, you should use switches in the LAN to eliminate collision problems.



Provide LAN connectivity via switched ports to minimize collisions (delay and jitter) in an IP telephony environment, where users have IP Phones on their desktops. It is marginally acceptable to use hubs, but only if you reduce the number of devices connecting into a single hub.

Note Remember that a shared LAN reverts to the speed of the slowest device.

Firewalls present problems for IP telephony traffic. Their primary use is to protect access to and from the Internet. Because enterprise Voice over Data solutions are targeted at the corporate intranet, most firewalls are not properly equipped to handle this traffic.



Voice network assessment involves the detailed recording of voice circuits between all offices, particularly international sites. You should record the type (analog or digital) and number of voice circuits. You should also assess voice traffic characteristics by using capacity-planning techniques.

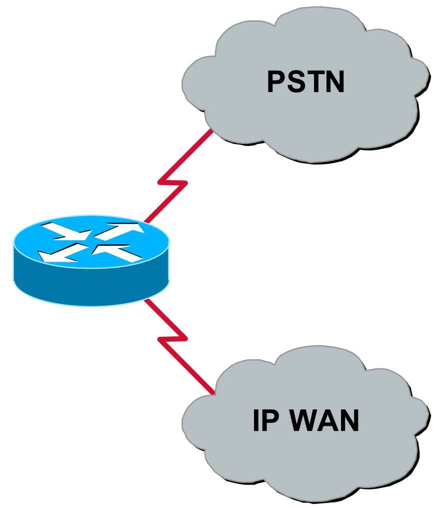
Note You should record current voice traffic costs so you can use this data in a cost-justification model.

Verify that you have assessed and recorded all voice components, PBX models, and supported interfaces, as well as any installed voice-mail systems and their interfaces to the PBX. You need to identify the PBX protocols that are used to connect to the central office (CO) switch and other PBXs. If the inter-PBX connectivity does not use Q Signaling (QSIG), identify whether Cisco cross-connect methods, or frame forwarding, is necessary to transport proprietary signals between PBXs.

Voice Network Assessment: Voice Circuit and Costs

Cisco.com

- List types of voice circuits and service from telco.
- Assess costs of the services.
- Consider how service can be delivered more effectively using Cisco Voice over Data and IP telephony solutions.



 EVODD v3.3—1-14

You should consider several topics when assessing voice circuits. You should identify tie-lines when considering route replacement. You should also consider other deployed services, such as Direct Inward Dialing (DID), Off-Premises eXtension (OPX), and Centrex services. The solution that you choose should support the effect of these services.

Use the cost of these services to build a cost justification model for Voice over Data. Consider how to offer the identified services more effectively using Cisco Voice over Data and IP telephony solutions. You should provide all of this information to the customer to demonstrate that you have considered all existing requirements in the solution.

Voice Network Assessment: Inventory PBX(s)

Cisco.com

- **PBX vendor and equipment model**
- **Locations and current PBX signaling protocols:**
  - **Proprietary PBX protocols will be transparently transported through Cisco router.**
  - **Tandem PBX(s) can be replaced for cost savings.**
- **PBX services that must be preserved:**
  - **Industry-standard PBX protocols (QSIG) do not always support critical PBX services.**
- **Existing voice-mail assessment:**
  - **Bandwidth considerations for centralized or distributed voice-mail access.**
  - **Triggering MWI is often a critical requirement.**

 EVODD v3.3—1-15

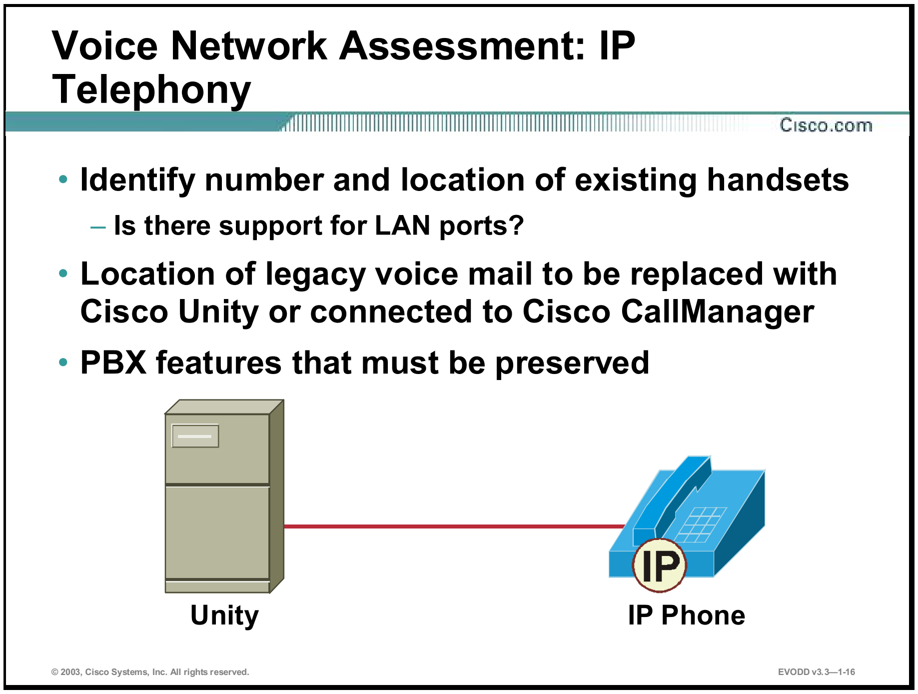
The installed PBX is the heart of the customer voice network and will probably coexist with CCM for the short term. It is critical that you connect the PBX to the data network.

Determine the PBX vendor and equipment model to help identify any proprietary PBX protocols. Although QSIG is the industry standard for interconnecting PBXs, the customer environment does not often support it because of cost or lack of implemented features. Cisco frame forwarding or cross-connect can transport proprietary protocols, but you must clearly identify the proprietary protocol.

Consider the importance of the PBX location from a design perspective. You should identify whether the data network extends to remote PBX locations. Consider PBXs with the tandem switching function and how the data network can provide this function.

Most enterprise telephone networks make use of a voice-mail system. Bandwidth must be available for retrieving voice-mail messages from a remote voice mailbox on the other side of the WAN. Because most systems do not save voice mail in a compressed format, voice-mail traffic, and compression of the voice signal as the recording plays out, must be part of the traffic analysis and design stage.

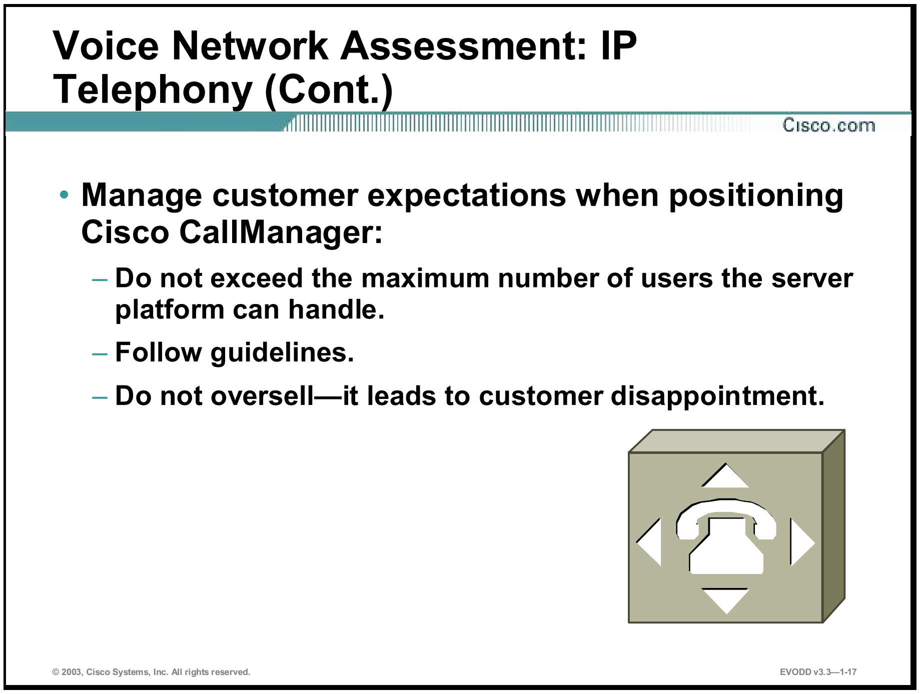
You need to investigate how important an illuminated, message-waiting light is to the customer. Although this Message Waiting Indicator (MWI) appears to be a simple request, it does not always have a simple solution. For example, during the migration phase on Lucent and Octel systems, you must set the dialing sequence that is used to enable and disable the MWI on telephones connected to the legacy PBX.



When designing an IP telephony solution, consider these additional questions:

- Where are the existing telephone handsets?
- Are LAN data ports extended to the locations of the existing telephone handsets?
- Where is the voice-mail system that Cisco Unity is replacing?
- Is there sufficient bandwidth if individuals access voice mail from Cisco Unity across the WAN?
- What PBX features must remain on CCM?

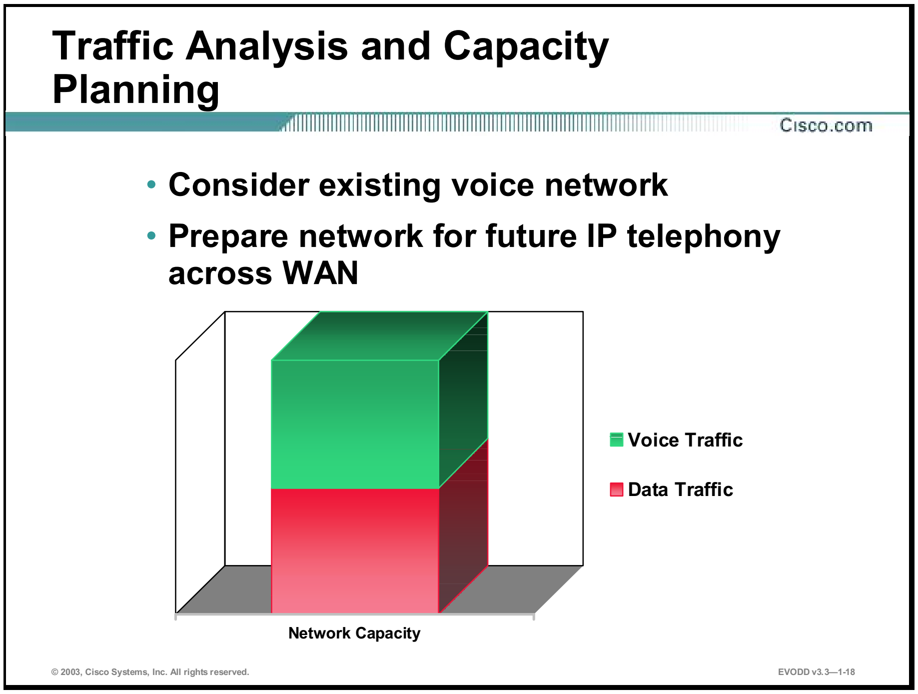
Note	Some locations will replace the PBX.



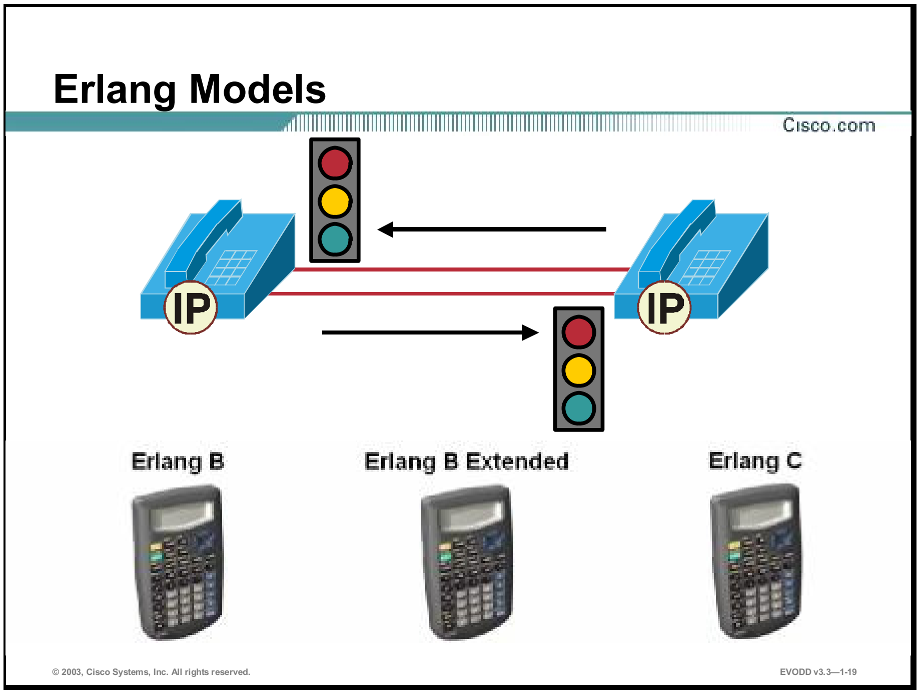
You should discuss CCM features with the customer, including features that are currently available and features that will be available in the future. This communication is an effective way to manage customer expectations. Cisco attempts to implement 98 percent of all PBX features that an enterprise typically uses. Many PBXs support a long list of features, but users typically understand and use very few.

Traffic Analysis and Capacity Planning

This topic identifies and describes the need for proper capacity planning.



You must consider proper capacity planning when building IP telephony voice networks. If you understand how much bandwidth each VoIP call uses, you can apply capacity-planning tools, such as Erlang models.



In traffic theory, you measure traffic load. When measuring traffic load, you compare the ratio of call arrivals in a specific period of time to the average amount of time taken to service each call during that time. These measurement units are based on average handle time (AHT). AHT is the total time of all calls in a specific period, divided by the number of calls in that period, as shown in the following example:

(3976 total call seconds)/(23 calls) = 172.87 sec per call = AHT of 172.87 seconds

One Erlang is 3600 seconds of calls on the same circuit, or enough traffic load to keep one circuit busy for 1 hour. Traffic in Erlangs is the product of the number of calls multiplied by AHT and divided by 3600, as shown in the following example:

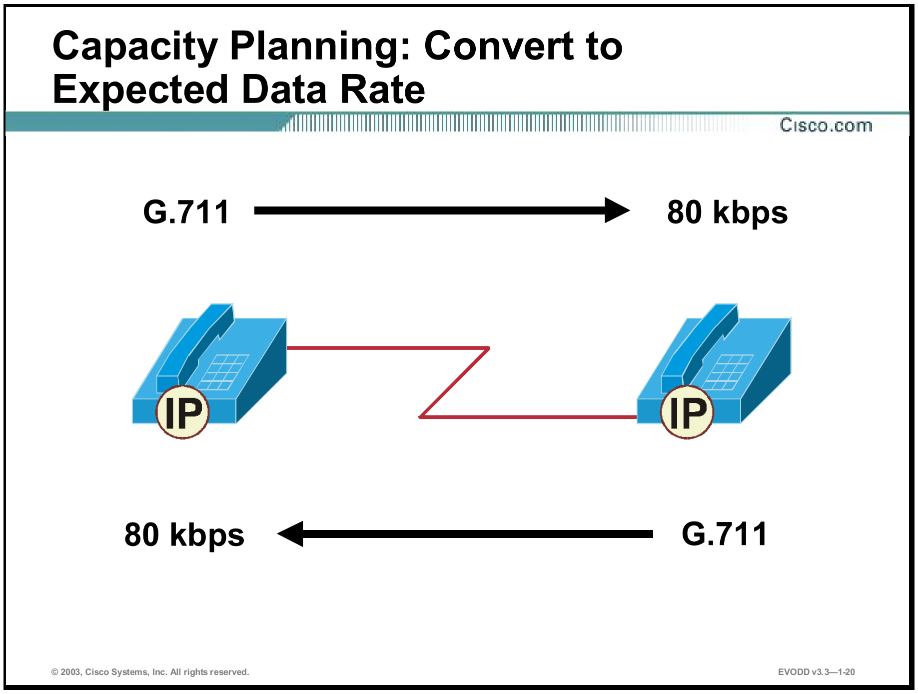
(23 calls * 172.87 AHT)/3600 = 1.104 Erlangs

Erlang models can help to determine the required bandwidth. The equations for solving these models are quite involved. Typically, designers use reference tables or Erlang calculators.

Use the Erlang B model for overflow paths when trunks are busy. Use the Erlang B Extended model when there is not an alternate route. In this case, a caller receives a busy signal and does not use the alternate route to the PSTN.

Use the Erlang C model in call center environments, where the system places a caller on hold if no bandwidth is available.

Note You can find more details on the Erlang models at www.erlang.com. Erlang calculators are also available at this site.

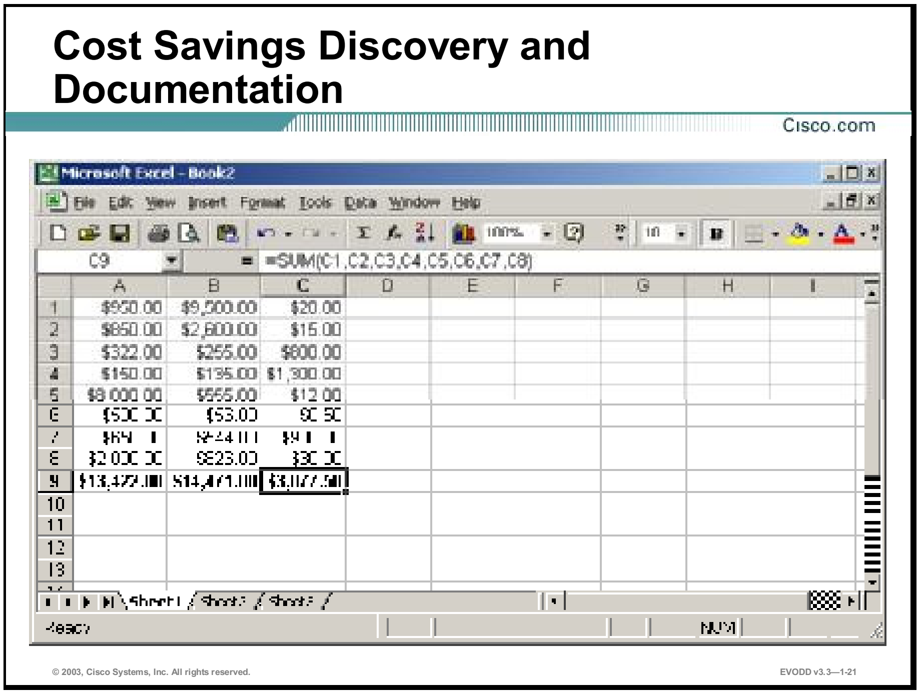


You should use the Erlang model and data rate conversion tables for VoIP, VoFR, and VoATM to determine the required bandwidth. Many additional factors affect bandwidth. These factors include encoding algorithm; voice and/or data transport protocol; the use of voice activity detection (VAD); and music on hold (MOH), which makes VAD less effective. Bandwidth calculations typically assume that there is one talker at a time. The bandwidth requirement doubles when both parties talk at the same time.

Network applications today require more bandwidth, so overbuilding bandwidth is a good practice if the budget permits the expense. Overbuilt bandwidth will support additional growth and unanticipated data demands.

Cost Documentation

This topic explains the importance of cost discovery in network planning.

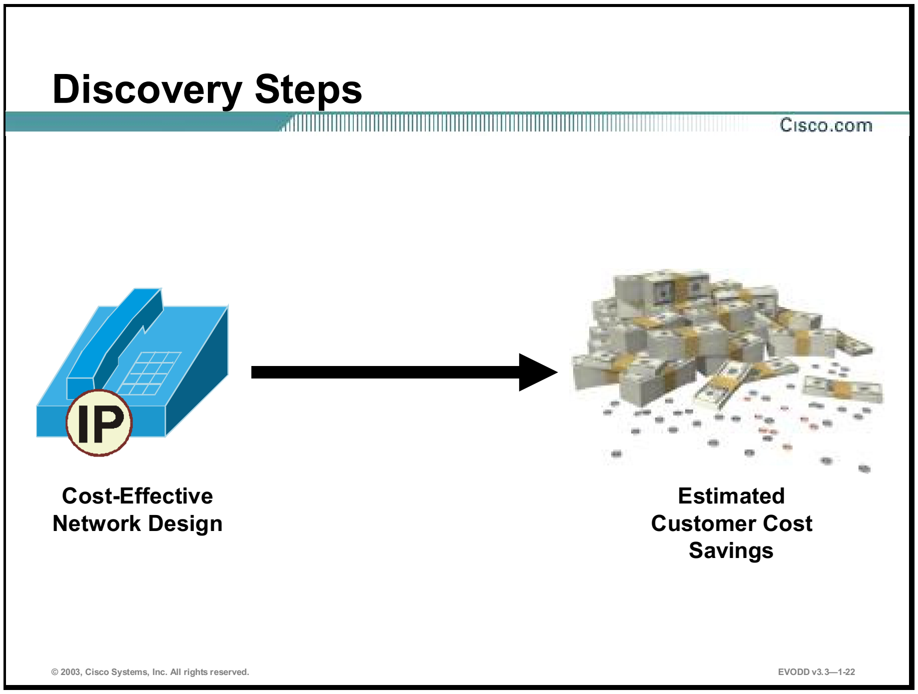


Discovering current and expected costs, under the current or converged design, is the most difficult step in the process.

Customers must use data from multiple sources to identify costs. Costs are arrived at by a combination of calculation and estimation. Spreadsheets like the one shown here capture these costs and create a template for each customer.

Insert your cost information into your Cisco spreadsheet model. This model computes and displays return on investment (ROI), using typical economic parameters as determined by the financial decision process of the customer.

Although the Cisco tool uses an Excel spreadsheet, the underlying details are not yet available to the user, which simplifies its use.



In the discovery process, you work with the customer to capture, estimate, or assume costs. When you participate to this degree, you can learn many important details about the existing systems and associated costs for the customer.

You can use this knowledge to create a cost-effective Voice over Data design that meets customer expectations and assists in the decision-making process. You can use this design to arrive at the estimated cost.

Spreadsheets are a convenient way to capture the results of joint discovery and provide input for the Cisco cost model. These results include capital costs associated with the new design, one-time cutover cost estimates, and recurring cost estimates.

Discovery of Current Costs

Jointly gather and document current system costs:

- **Voice**
- **Data**
- **Distribution/cable**
- **Battery/backup power**
- **Site**
- **LEC/circuit costs**
- **Toll costs**
- **Staffing**

EVODD v3.3—1-23

Consider the following cost areas to help you in the discovery process:

- Voice
- Data
- Distribution and cable
- Battery and backup power
- Site
- Local exchange carrier (LEC) and circuit costs
- Toll costs
- Staffing

You can create other cost areas that may be more appropriate for individual customers.

Existing Access

Cisco.com

Number of DID trunks?	PSTN services?
Number of CO trunks?	Ratio of analog or digital to users?
Fax or other individual lines?	How many users? How many locations?
Digital or analog?	Current provisioning (DS-1/T1)?
Monthly access cost?	Continuing costs irrespective of PBX technology?
Number of Off-Premises eXtensions (OPXs)?	Station lines at remote sites?
Monthly cost per OPX?	$20 to $80 per month per OPX
Number of remote sites?	Physical concentration/location?
OPX sites served by routers?	If routers, multiservice?
Tie trunks to other PBXs?	Leveraged to router backbone?
Monthly cost of tie trunks?	Potentially displaceable costs?
Drop and insert potential?	Savings for voice/data on same DS-1/T1?
Inter-exchange carrier (IXC)?	Information and potential partner (a network powered by Cisco)?

EVODD v3.3—1-24

Customer analysis plays a major role in identifying access items. The customer should analyze access bills and be prepared to share this information with the Cisco design team. If the customer cannot share this information, discovery may be a very lengthy process.

Customer Premises Equipment

Cisco.com

PBX manufacturer/model?	Lucent, Nortel, NEC/Definity, Meridian, NEAX?
Is PBX software generic?	G3, Option 11, etc.?
Remote shelves/EPNs?	Physical distribution of single PBX often in MAN?
Number of attendant positions?	Number of operators answering incoming calls?
Number of station lines, including OPX?	Extensions on PBX?
Number of digital handsets? What models?	Proprietary and often display-capable?
Speakerphone equipped?	Do digitals have speakerphones?
Number of analog handsets?	Industry-standard 2500 style?
CDR/SMDR-equipped PBX?	Call detail record/station message detail-capable?
Call costing/accounting software?	Vendor/model: internally developed or outsourced?
Customer charges back long distance?	Real charge-back to user departments?
Call center?	Typically in customer service organizations?
ACD via separate switch?	If yes, vendor/model?
Call centers?	Number and locations?
Average number of trunks at call centers?	More than number of agents - calls queued?
Average number of agents?	(More than 100 is considered large.)
Interactive voice response/manufacturer?	IRV or Voice Response Unit(s) (VRU)?

EVODD v3.3—1-25

System operations staff typically understands most of the customer premises equipment (CPE) items shown in this figure. The design team should discover any items not understood by the system operations staff.

Voice and Data Costs

This topic explains why the costs for voice and data systems are important inputs to the Cisco cost model.

Voice Mail/Unified Messaging

Cisco.com

Voice-mail manufacturer/model?	Voice mail provided by PBX vendor? Octel, Intuity-Audix, Meridian Mail, Active Voice Repartee, etc.
Number of voice-mail ports?	Number of concurrent messages left/retrieved?
Voice-mail ports?	Analog or digital?
Does voice mail have SMDI?	Voice mail with Standard Message Detail; interface potential for integrating mail controls.
Voice-mail users?	How many?
Minutes of storage per user?	How many?
Customer employs or wants Unified Messaging? How soon?	Access voice mail or text messages from voice mail or e-mail?
Current mail environment?	MS Exchange, Notes, Eudora-POP3, IMAP, or HTTP DAV-based?

 EVODD v3.3—1-26

Voice mail is a critical business capability for most customers today. A converged voice and data approach provides forward-looking capabilities. This discovery area, and the discovery process, can help customers understand the advanced technologies that are available.

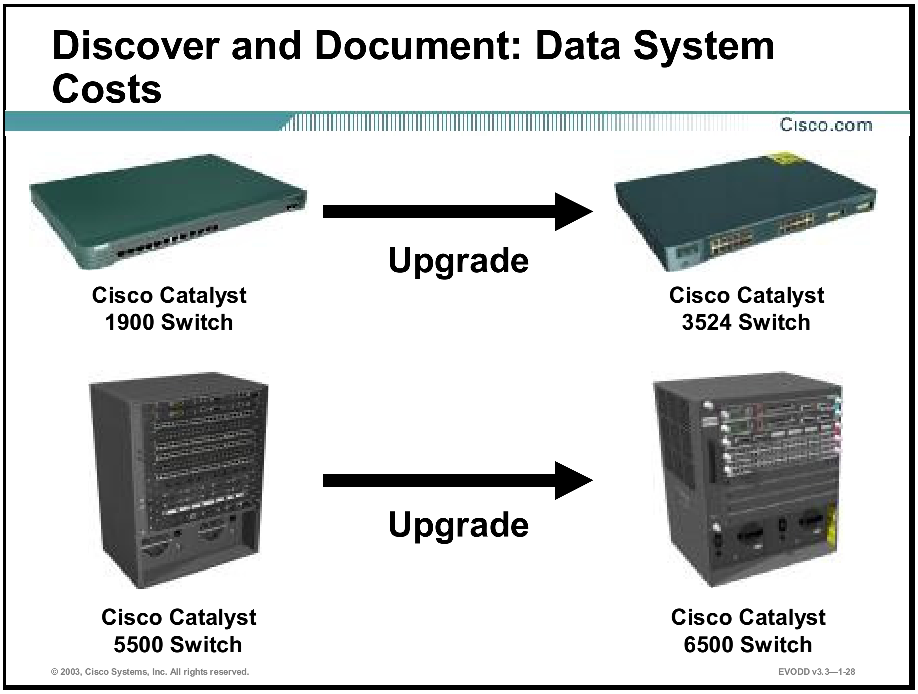
Other Information

Cisco.com

Customer TAPI applications? Names and types of TAPI applications?	Telephony programming interface applications in use?
Voice system UPS/backup power capabilities?	Operations for commercial power failure?
Annual number of moves, adds, and changes? Cost per MAC?	Cost typically ranges between $40 to $200 per MAC.
Audio teleconferencing? Vendor/provider?	Cisco has relationships, among others, with Polycom and Latitude
Video conferencing? Vendor/provider?	Cisco has OEM relationship with RadVision.
Windows NT/2000 customer strategy?	Positive or negative bias?
Customer uses SQL server databases?	Cisco CallManager uses SQL.
Customer uses NetMeeting or NT-Phone Dialer?	These can be integrated into Cisco CallManager as H.323 clients.

 EVODD v3.3—1-27

The voice-mail and/or unified messaging area can help the customer focus on the future while discovering current system parameters and costs.



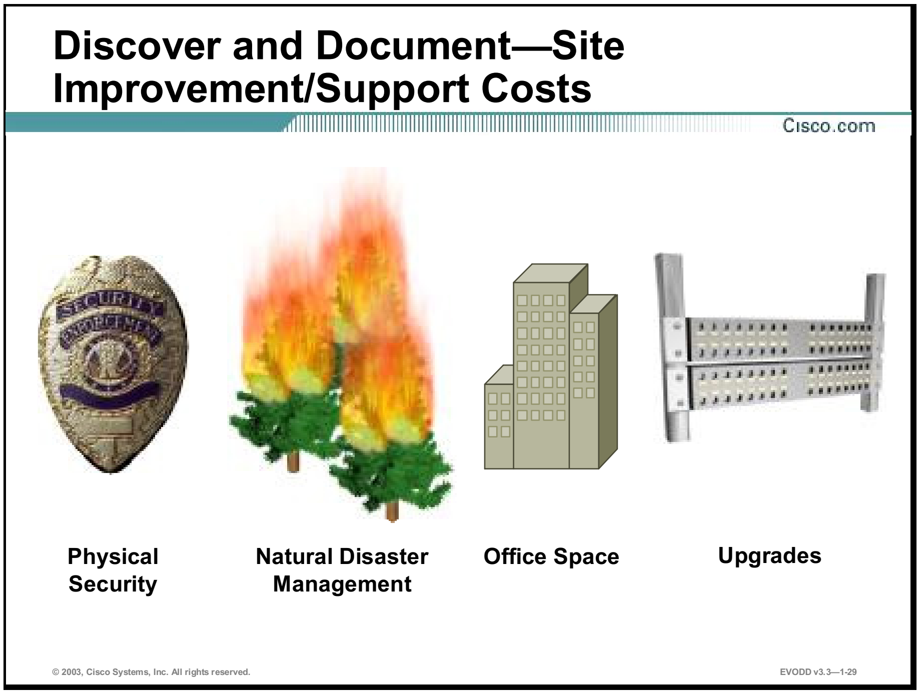
Enter the data system costs into the model. Data system costs do not include cabling and/or wire-fiber hardware and software components.

You should document data system costs including:

- LAN hubs, switches, ports, and management tools (upgrades, maintenance and recurring costs for hardware and software)
- Other needed data network upgrades (maintenance and recurring costs)
- Moves, additions, and changes
- Management, security, and privacy tools

Other Cost Considerations

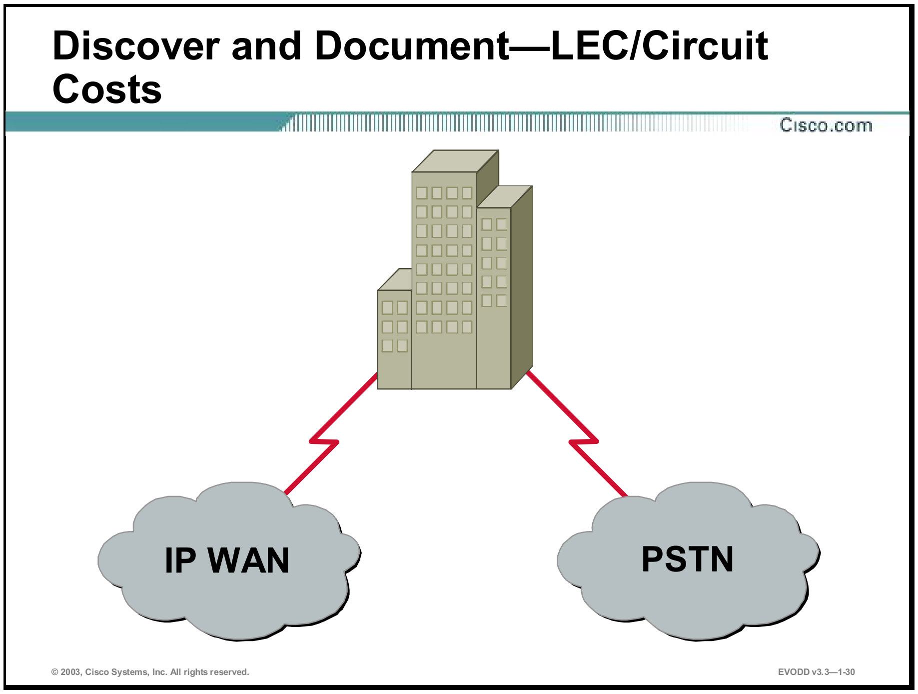
This topic describes several additional cost considerations that you must identify for the cost-savings model.



The cost-savings model should include site considerations from the discovery and modeling process. There should be recurring savings associated with a converged environment.

You should document improvement and/or support costs including:

- Physical security
- Fire and natural disaster management
- Space costs
- Necessary build-out for upgrades or changes

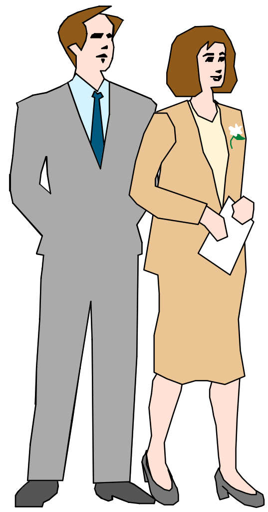


Important considerations include changes for leased circuits, or contract provisions associated with changes or decommissioning. You may have gathered this data when you considered the area of access.

Examples of LEC and/or circuit costs include:

- **Voice:** DIDs, central office (CO) trunks, OPX circuits, and tie-lines
- **Data:** WAN circuits, inter-building, or inter-company metropolitan-area network (MAN) links
- **Contract change costs:** Circuit installation, reconfiguration, and circuit decommissioning

Discover and Document—Associated Staff Costs



EVODD v3.3—1-31

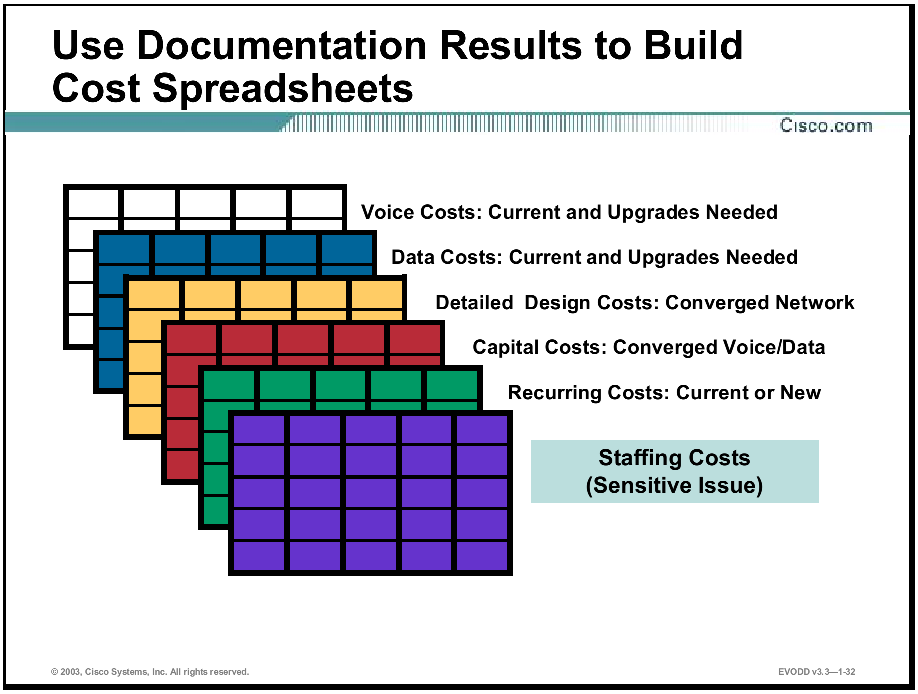
Staffing costs are one of the most sensitive areas to consider. If the budget allows, the customer may choose not to make changes in this area, except for duties. You should consider the following when documenting staff costs: voice and/or data analysts, network engineers, operations staff, voice systems operators and attendants, and voice storage management. Most customers will realize savings associated with converging data and voice environments.

Staffing costs include:

- Internal voice staffing costs
- Internal data staffing costs
- Contract staff or outsource staffing costs

Input Information

This topic explains how to use cost information to identify potential cost savings for the customer.



The inputs to the Cisco standard model focus on the costs associated with voice in the new environment. In most cases, there is not a clear dividing line between voice duties and other duties. You must make assumptions about the portions of costs that are voice related and use them in the cost calculations.

Note	Spreadsheet example templates are available as files.

Model Constraints in Economic Decision-Making

Cisco.com

- Cost-savings tool may not indicate sufficient payback for network integration.
- Do not use cost-savings tool in isolation.

 EVODD v3.3—1-33

Determining the economic benefits of any technology transition, such as Voice over Data, is more complex than calculating net present value (NPV) and payback.

The economic benefits typically include:

- Reduced costs from the use of integrated management tools and staff
- Reduced forward costs for scaling, integration, and adding new technology associated with a standards-based network
- Leveraged budget from the integrated voice and data network

Economic decision-making that involves technology transitions creates a multidimensional problem, often more complex than simple NPV and payback calculations indicate. In the case of multiservice solutions, like VoIP, VoFR, and VoATM, consider the following:

- Network integration reduces overhead, including management tools and the staff to operate the network.
- Network integration reduces risk factors from technology obsolescence, compatibility, and scaling.
- Network integration improves budget use. Each dollar spent improves both data and voice services.

Note	Disclaimer for the Cisco MultiService Network Feasibility Tool: This tool may not indicate sufficient payback for multiservice network integration, even though it is the best business decision. Therefore, you should not use this tool in isolation, but as support for higher-level decisions, which include factors listed in this topic.

Information Input

Cisco.com

Cisco Systems Multiservice Networking Feasibility Tool

BASED ON EIGHT ASSUMPTIONS:	
1. Capital requirements for VoIP:	$250,000
2. Projected one-time cutover expenses:	$25,000
3. Current monthly voice expense:	$15,000
Annual growth rate for voice	10%
4. Projected monthly voice expense (Portion of network used for voice)	$5,000
Annual growth rate for network expense	5%
5. Number of voice ports required	150
6. Annual tax rate (38% assumed; can change)	38%
7. Project life in years(3, 4, or 5 years)	3
8. Discount rate/cost of capital (12% assumed; can change)	12%

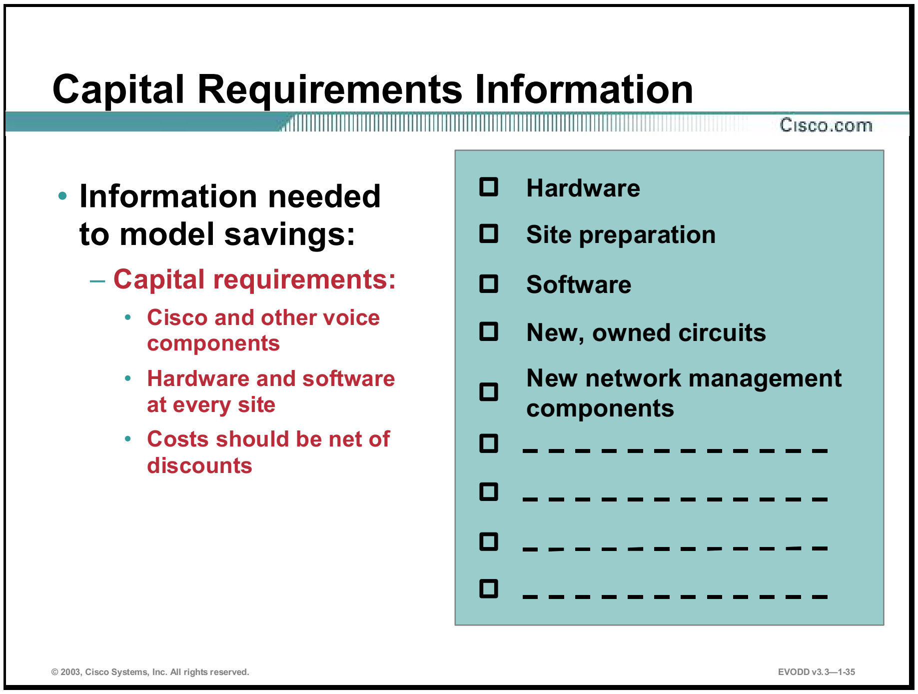
 EVODD v3.3—1-34

The model shown in this figure includes five tabs in a single Excel file (.xls). The first tab, labeled "Introduction," provides an overview of the model and states the disclaimer. The second tab, labeled "Instructions," provides information on assuming or estimating input to the model. It also briefly describes three results from the model:

- Approximate payback, in months
- NPV
- Internal rate of return (IRR)

You enter data into the third tab. The fourth and fifth tabs display the results.

The default model values—$250,000 for capital requirements and 10 percent of $250,000 for a one-time cutover—are illustrative. Users should not assume that one-time cutover expenses are 10 percent. In less complex solutions, where Cisco components replace carrier voice circuits in a link-to-link configuration, cutover costs may be lower than 10 percent. In other cases, the cost of cutover (as a percent of capital) can be much higher than 10 percent. For example, implementing PBX interfaces requires site preparation, extensive customer support, and user involvement. To model savings, you must use a like comparison. Most projects add new capacity or features. Additions are extras; they should not add to the calculated values used in the model (although they are a real overall project cost).

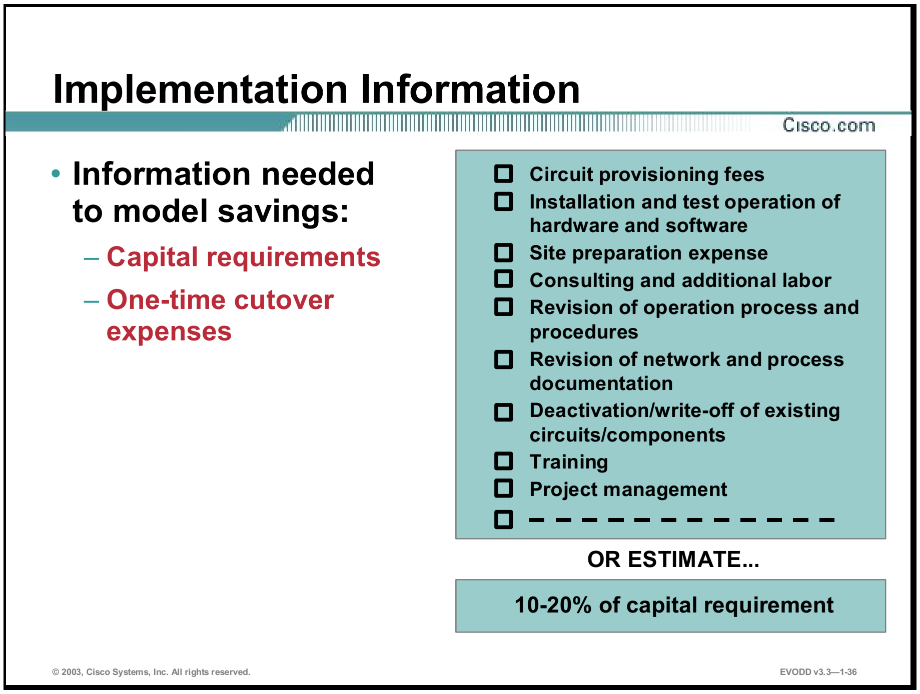


The first entry to the model shown in this figure is capital requirements. This entry includes all Cisco components, as well as all other voice components and capitalized items necessary for the project. It should include all hardware and software needed for implementation at every customer site in the project.

Note	Use costs that are the net of discounts for calculations.

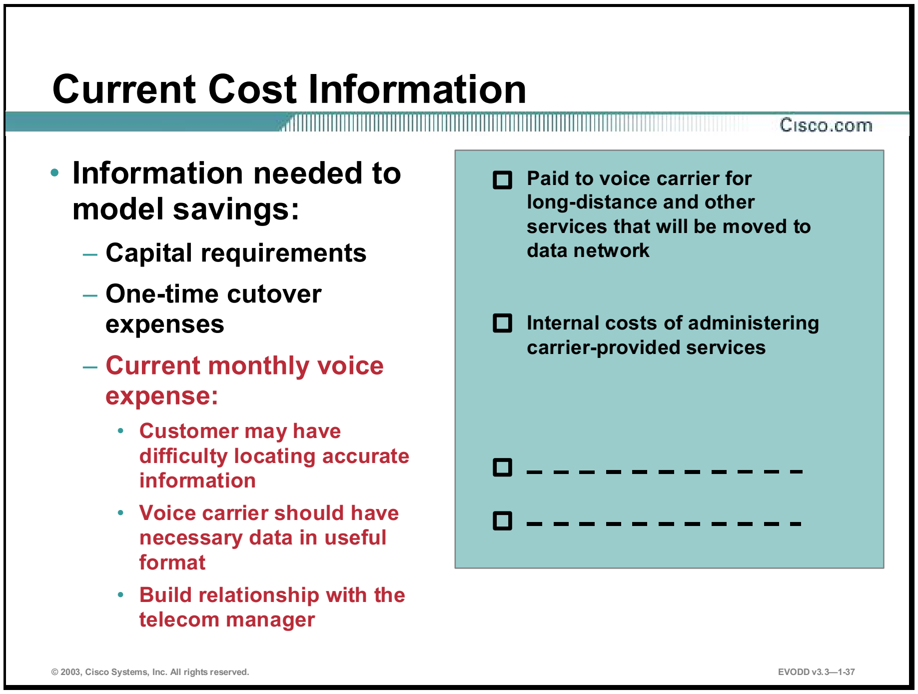
You should include all of the items that significantly contribute to the overall cost, especially if you estimate the cutover costs as a percent of the capital requirements.

Create a project checklist with hardware, software, site preparation, new network management items, and the costs for each.



The second value as shown in this figure, is the one-time expense for the cutover. These implementation costs typically include Cisco or partner fees, third-party fees, and extra (paid or costed) labor provided by the customer staff. There may be fees associated with decommissioning existing circuits and other telephone company services. Revising customer documentation also has associated costs. You will need to develop customer procedures and train the staff to use the new technology.

You can estimate the cutover costs at 10 to 15 percent of the capital costs if a more detailed analysis of the cutover costs is not available. To be more conservative, you can use 20 percent. For a successful project, you should ensure that you have sufficient project management. You can estimate that project management costs will add 20 percent to the cutover costs.

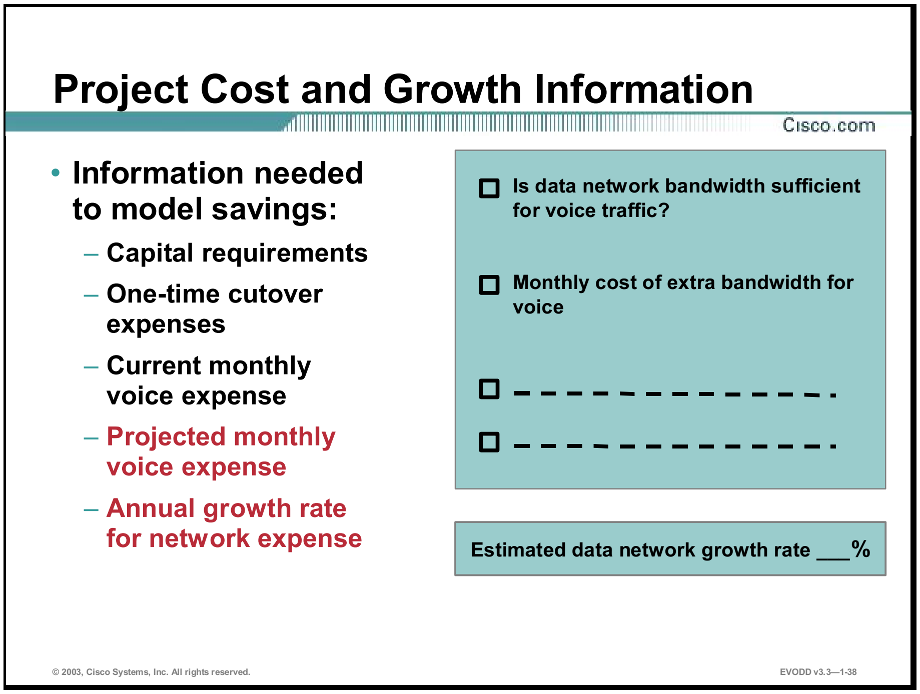


When you enter the current monthly expense for voice—as the value for cutover—you should include all of the costs: fees paid to carriers and in-house costs.

The customer may have difficulty locating accurate records, especially the plain old telephone service (POTS) costs. Typically, customers do not have accurate or current records; they will only know their overall monthly costs.

The voice carrier should have this information and the ability to format it in a useful way. If the project involves replacement of TDM components and leased or measured circuits, POTS bills provide most of the information for this value.

The telecom manager may have much of the required information, or may be able to access it if necessary. You should establish a good relationship with the telecom manager and with key staff members who can help locate this information.



You should next enter the projected monthly voice expenses for the project into the model, and an annual growth rate for the data network, if additional capacity is necessary to support voice.

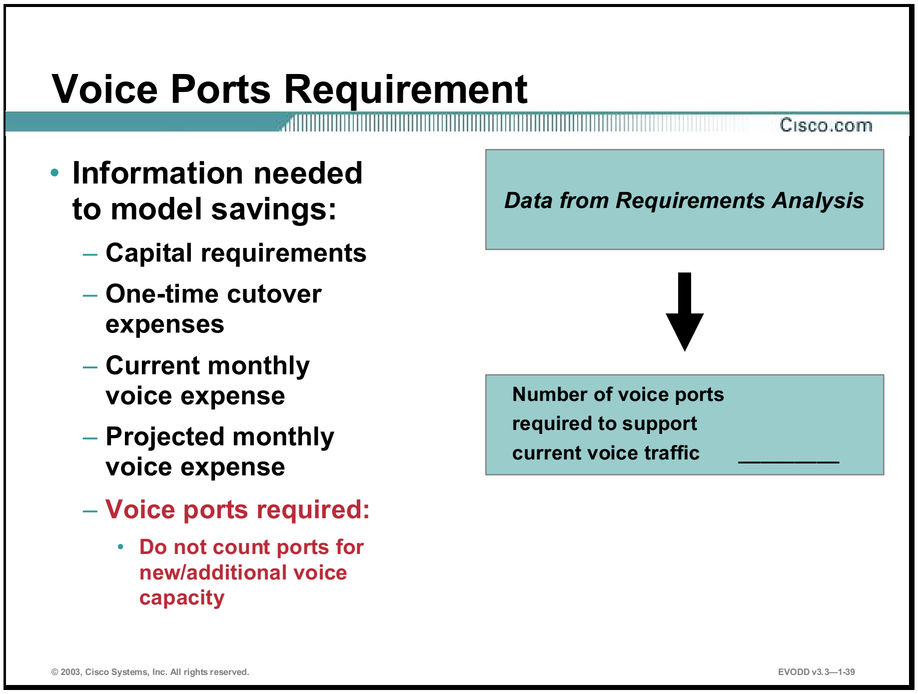
In some cases, the customer has sufficient data network capacity to support voice traffic without adding to the network. In this case, enter 0.

When upgrading a data network, include the monthly costs associated with that upgrade in the model.

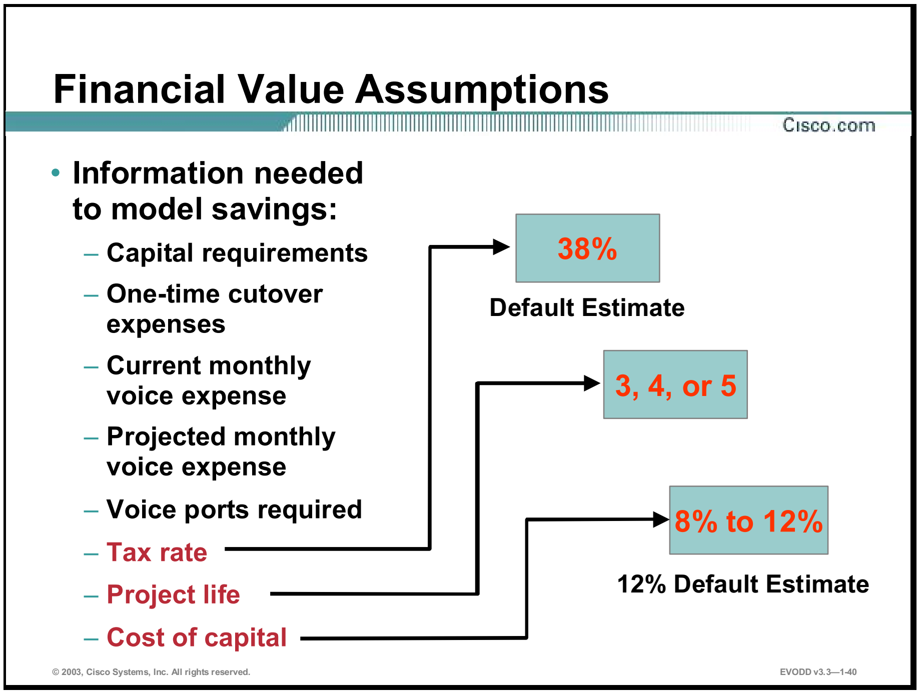
Note	When calculating savings associated with a migrating voice to the data network, it is important to assure that you are comparing similar values.

You should not charge the additional cost of new voice and video applications that were made possible by the migration against the new Voice over Data cost.

The model shown here uses the estimated growth rate to show the effect of growth on costs, with and without integration of voice and data.



The initial design of the project network determines the number of voice ports required to support the current voice traffic in a VoIP network. Do not charge ports for new or additional voice capacity as a new cost (so that a sound comparison to the old values can be made). Although these are costs of the overall project, you should associate these costs with the new application capability.



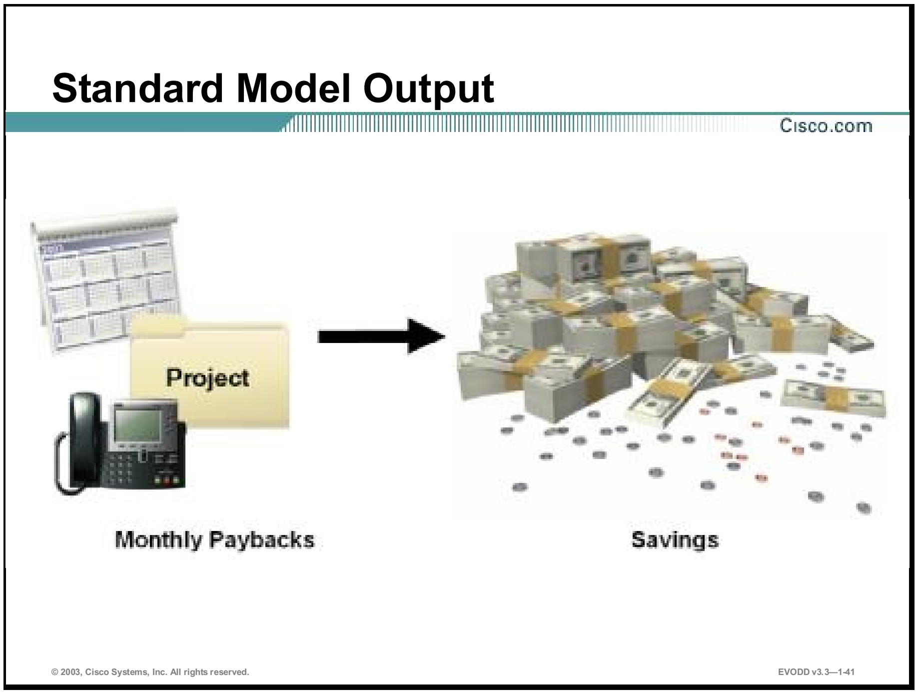
Customers should be able to provide their approximate annual corporate tax rates. You can use 38 percent in the United States if this information is not available.

The model accepts 3, 4, or 5 years as a value for the projected life. If the customer does not provide a suggestion, use 3 years. Using 3 years provides a worst-case analysis solution; the savings is greater if you select 4 or 5 years.

The cost of capital varies widely among countries and among corporations. The customer provides the best value for the model. In the United States, you can use a rate between 8 and 12 percent as a default. Use the cost of capital to discount the calculated savings associated with the project and to present the NPV. These values help to calculate the IRR or the hurdle rate (the rate that the customer must achieve for a cost-effective project) for the customer.

Standard Model Output

This topic shows you how the standard cost-savings model calculates and renders results associated with moving to VoIP.



The payback calculation for the project, its NPV, and annual IRR are on tab 1 of the model. Tabs 2 and 3 present more information about projected savings.

Standard model outputs include:

- Approximate payback of project, in months
- NPV of project
- Annual rate of return for the project
- Voice savings sensitivity analysis
- Graphic presentation of daily call minutes required to yield various cost-per-minute savings
- Summary of project savings by year of life
- Graphic presentation of voice cost trends with and without the Cisco solution

Project Results

Cisco.com

ASSUMPTIONS:		
1.	Capital requirements for VoIP	$250,000.00
2.	Projected one-time cutover expenses	$25,000.00
3.	Current monthly voice expense	$15,000.00
	Annual growth rate for voice	10%
4.	Projected monthly voice expense (Portion of network used for voice)	$5000.00
	Annual growth rate for network expense	5%
5.	Number of voice ports required	150
6.	Annual tax rate (38% assumed; can change)	38%
7.	Project life in years(3, 4, or 5 years)	3
8.	Discount rate/cost of capital (12% assumed; can change)	12%
PROJECT RESULTS:		
1.	Approximate payback of project in months	28
2.	Net present value of project	$38,426.00
3.	Project annual internal rate of return	21.44%

 EVODD v3.3—1-42

You do not have to access the spreadsheet to change internals or add features.

Voice Savings Sensitivity Analysis

Cisco.com

- Based on current monthly voice expense and port requirements from tab 1

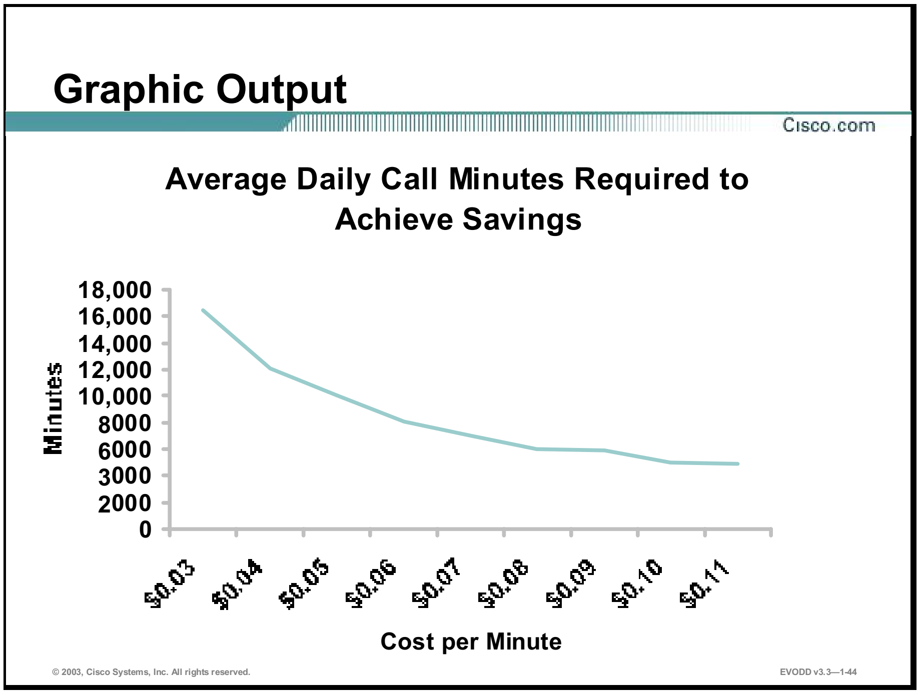
If your average cost per minute for voice is:	Then the average daily minutes required to generate current voice expenses is:	Average daily minutes required per voice port:	Average daily port utilization % (1):
$0.03	16,438	109.6	22.8%
$0.04	12,329	82.2	17.1%
$0.05	9863	65.8	13.7%
$0.06	8219	54.8	11.4%
$0.07	7045	47.0	9.8%
$0.08	6164	41.1	8.6%
$0.09	5479	36.5	7.6%
$0.10	4932	32.9	6.8%
$0.11	4483	29.9	6.2%

(1) Percentage of time each port is used, assuming 8 hours per day

EVODD v3.3—1-43

The table in this figure shows standard model output that demonstrates the relationships among:

- Cost per minute
- Average daily minutes for generating the current voice expense if the Call Detail Record (CDR) reporting tools do not easily display it
- Average daily minutes per voice port
- Associated voice port utilization



This figure shows the relationship between the number of call minutes and the cost per minute, which you can use to determine the number of call minutes required to reduce costs. In this example, the customer achieves savings after 7500 minutes if the current charge per minute is $.07.

Summary Output

Cisco.com

Voice Costs and Project Savings Summary

Life of Project

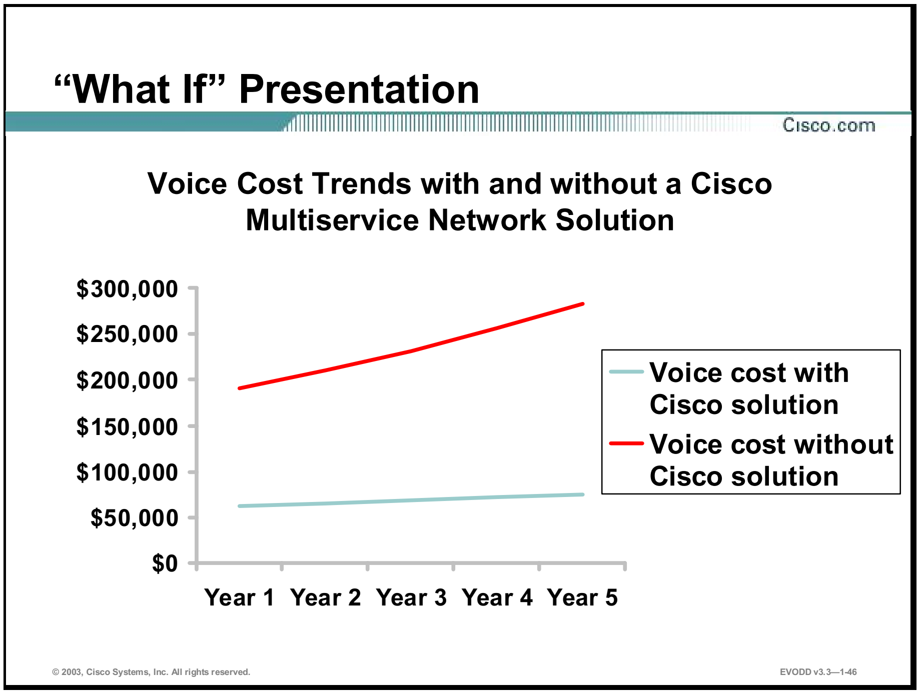
	Year 1	Year 2	Year 3	Year 4	Year 5
Voice Costs without Cisco Solution	$190,054	$209,955	$231,940	$256,228	$283,058
Voice Costs with Cisco Solution	$61,650	$64,804	$68,120	$71,605	$75,268
Estimated Savings from VoIP*	$128,404	$145,151	$163,821	$184,623	$207,790

*** Note: Model does not round to significant digits when presenting data to customer.**

 EVODD v3.3—1-45

The table in this figure illustrates an increase in savings when monthly voice costs increase. This chart is based on the percent growth you entered in the model.

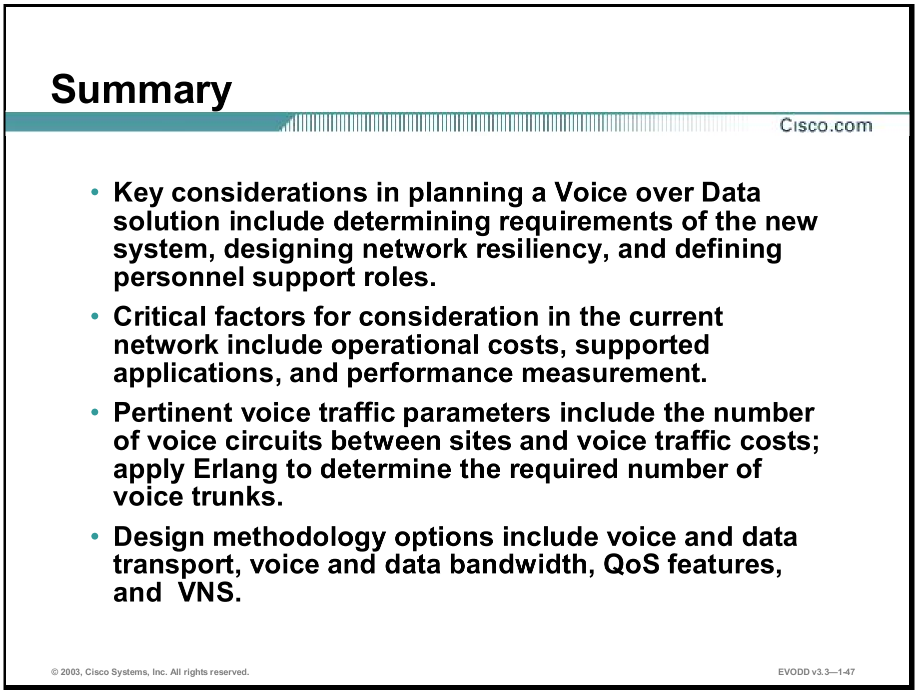
Note	The model does not round to significant digits when presenting data to the customer.



This figure shows the graphic representation of the previous table.

Summary

This topic summarizes the key points discussed in this lesson.



Summary (Cont.)

Cisco.com

- **Poor voice quality signals the need for tuning.**
- **Design methodology includes determining the features that currently exist, identifying the features that must be maintained, and meeting the design requirements of the proposed IP telephony solution.**
- **Customer costs are clarified by determining the costs of current access, existing CPE, and messaging services.**
- **The cost-savings model requires the completion of five worksheets in an Excel spreadsheet.**
- **The cost-savings model output includes project payback, NPV, IRR, voice savings sensitivity analysis, call savings, project savings summary, and voice cost trends with and without the Cisco solution.**

 EVODD v3.3—1-48

Next Steps

After completing this lesson, go to:

- Standards for Voice over IP Signaling module

References

For additional information, refer to these resources:

- Cisco AVVID website:

  http://www.cisco.com/warp/public/cc/so/cuso/epso/avpnpg/vvaid_wp.htm

- IP Communications website:

  http://www.cisco.com/warp/public/779/largeent/avvid/

Laboratory Exercise: Calculating Trunk Capacity

The laboratory exercises are designed to reinforce concepts discussed throughout the course. This laboratory exercise focuses on capacity planning.

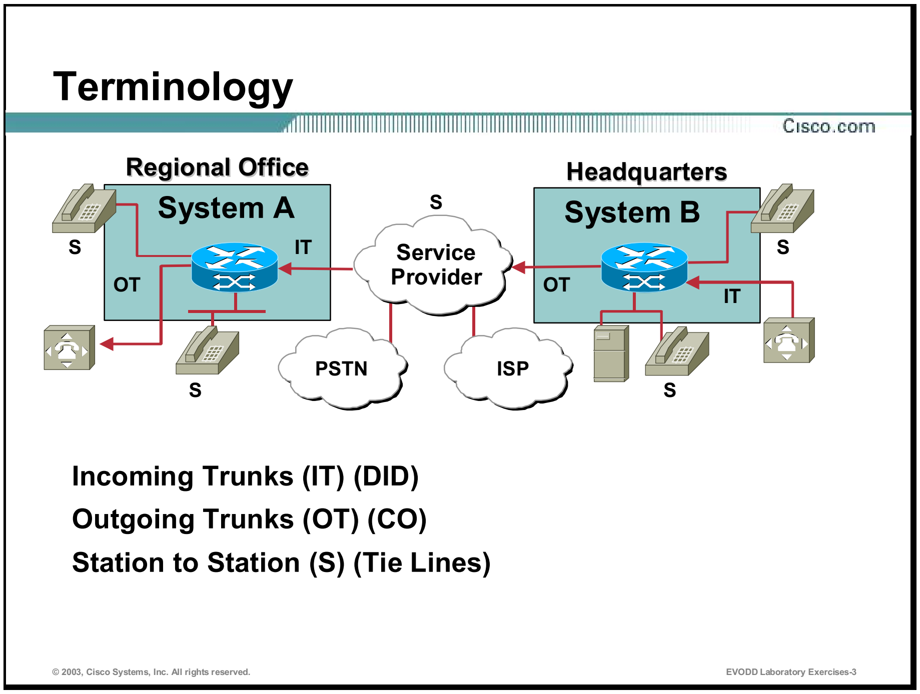
Exercise Objective

In this exercise, you will calculate trunk capacity.

After completing this exercise, you will be able to:

- Obtain information from customers on voice traffic patterns
- Classify traffic into specific types
- Calculate trunk capacity

Key Terms



Key terms for this lab are defined below:

- Incoming Trunk (IT): This trunk can be used as an incoming only Central Office (CO) trunk, a Bothway CO trunk, or a Direct Inward Dialing (DID) trunk.

- Outgoing Trunk (OT): This trunk can be used as an outgoing only CO trunk or a Bothway CO trunk.

- Station to Station (S): This trunk is a Tie Line (or Tie Trunk).

Job Aids

These job aids are available to help you complete the laboratory exercise:

- You will be using the following calculations when determining trunk needs:

  - Telephony Calculations= IT+OT+S

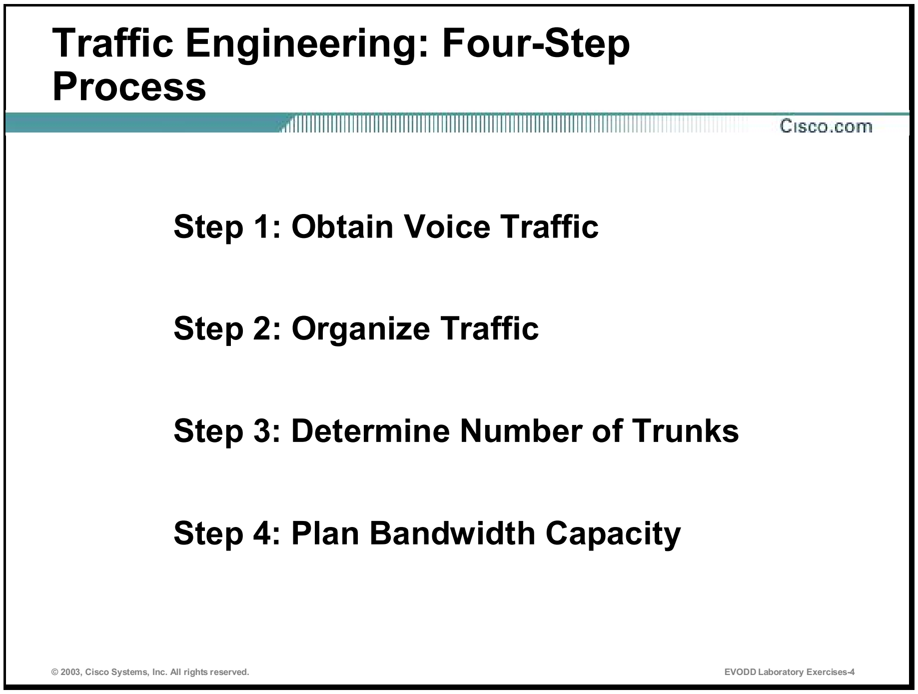
  - Switched Calculations= IT+OT+S+S

Table: ERLANG B Table

n	Grade of Service										n
	0.007	**0.008**	**0.009**	**0.01**	**0.02**	**0.03**	**0.05**	**0.1**	**0.2**	**0.4**	
1	.00705	.00806	.00908	.01010	.02041	.03093	.05263	.11111	.25000	.66667	**1**
2	.12600	.13532	.14416	.15259	.22347	.28155	.38132	.59543	1.0000	2.0000	**2**
3	.39664	.41757	.43711	.45549	.60221	.71513	.89940	1.2708	1.9299	3.4798	**3**
4	.77729	.81029	.84085	.86942	1.0923	1.2589	1.5246	2.0454	2.9452	5.0210	**4**
5	1.2362	1.2810	1.3223	1.3608	1.6571	1.8752	2.2185	2.8811	4.0104	6.5955	**5**
6	1.7531	1.8093	1.8610	1.9090	2.2759	2.5431	2.9603	3.7584	5.1086	8.1907	**6**
7	2.3149	2.3820	2.4437	2.5009	2.9354	3.2497	3.7378	4.6662	6.2302	9.7998	**7**
8	2.9125	2.9902	3.0615	3.1276	3.6271	3.9865	4.5430	5.5971	7.3692	11.419	**8**
9	3.5395	3.6274	3.7080	3.7825	4.3447	4.7479	5.3702	6.5464	8.5217	13.045	**9**
10	4.1911	4.2889	4.3784	4.4612	5.0840	5.5294	6.2157	7.5106	9.6850	14.677	**10**
11	4.8637	4.9709	5.0691	5.1599	5.8415	6.3280	7.0764	8.4871	10.857	16.314	**11**
12	5.5543	5.6708	5.7774	5.8760	6.6147	7.1410	7.9501	9.4740	12.036	17.954	**12**
13	6.2607	6.3863	6.5011	6.6072	7.4015	7.9667	8.8349	10.470	13.222	19.598	**13**
14	6.9811	7.1155	7.2382	7.3517	8.2003	8.8035	9.7295	11.473	14.413	21.243	**14**
15	7.7139	7.8568	7.9874	8.1080	9.0096	9.6500	10.633	12.484	15.608	22.891	**15**
16	8.4579	8.6092	8.7474	8.8750	9.8284	10.505	11.544	13.500	16.807	24.541	**16**
17	9.2119	9.3714	9.5171	9.6516	10.656	11.368	12.461	14.522	18.010	26.192	**17**
18	9.9751	10.143	10.296	10.437	11.491	12.238	13.385	15.548	19.216	27.844	**18**
19	10.747	10.922	11.082	11.230	12.333	13.115	14.315	16.579	20.424	29.498	**19**
20	11.526	11.709	11.876	12.031	13.182	13.997	15.249	17.613	21.635	31.152	**20**
21	12.312	12.503	12.677	12.838	14.036	14.885	16.189	18.651	22.848	32.808	**21**
22	13.105	13.303	13.484	13.651	14.896	15.778	17.132	19.692	24.064	34.464	**22**
23	13.904	14.110	14.297	14.470	15.761	16.675	18.080	20.737	25.281	36.121	**23**
24	14.709	14.922	15.116	15.295	16.631	17.577	19.031	21.784	26.499	37.779	**24**
25	15.519	15.739	15.939	16.125	17.505	18.483	19.985	22.833	27.720	39.437	**25**
26	16.334	16.561	16.768	16.959	18.383	19.392	20.943	23.885	28.941	41.096	**26**
27	17.153	17.387	17.601	17.797	19.265	20.305	21.904	24.939	30.164	42.755	**27**
28	17.977	18.218	18.438	18.640	20.150	21.221	22.867	25.995	31.388	44.414	**28**
29	18.805	19.053	19.279	19.487	21.039	22.140	23.833	27.053	32.614	46.074	**29**
30	19.637	19.891	20.123	20.337	21.932	23.062	24.802	28.113	33.840	47.735	**30**
31	20.473	20.734	20.972	21.191	22.827	23.987	25.773	29.174	35.067	49.395	**31**
32	21.312	21.580	21.823	22.048	23.725	24.914	26.746	30.237	36.295	51.056	**32**
33	22.155	22.429	22.678	22.909	24.626	25.844	27.721	31.301	37.524	52.718	**33**
34	23.001	23.281	23.536	23.772	25.529	26.776	28.698	32.367	38.754	54.379	**34**
35	23.849	24.136	24.397	24.638	26.435	27.711	29.677	33.434	39.985	56.041	**35**
36	24.701	24.994	25.261	25.507	27.343	28.647	30.657	34.503	41.216	57.703	**36**
37	25.556	25.854	26.127	26.378	28.254	29.585	31.640	35.572	42.448	59.365	**37**

38	26.413	26.718	26.996	27.252	29.166	30.526	32.624	36.643	43.680	61.028	**38**
39	27.272	27.583	27.867	28.129	30.081	31.468	33.609	37.715	44.913	62.690	**39**
40	28.134	28.451	28.741	29.007	30.997	32.412	34.596	38.787	46.147	64.353	**40**
41	28.999	29.322	29.616	29.888	31.916	33.357	35.584	39.861	47.381	66.016	**41**
42	29.866	30.194	30.494	30.771	32.836	34.305	36.574	40.936	48.616	67.679	**42**
43	30.734	31.069	31.374	31.656	33.758	35.253	37.565	42.011	49.851	69.342	**43**
44	31.605	31.946	32.256	32.543	34.682	36.203	38.557	43.088	51.086	71.006	**44**
45	32.478	32.824	33.140	33.432	35.607	37.155	39.550	44.165	52.322	72.669	**45**
46	33.353	33.705	34.026	34.322	36.534	38.108	40.545	45.243	53.559	74.333	**46**
47	34.230	34.587	34.913	35.215	37.462	39.062	41.540	46.322	54.796	75.997	**47**
48	35.108	35.471	35.803	36.109	38.392	40.018	42.537	47.401	56.033	77.660	**48**
49	35.988	36.357	36.694	37.004	39.323	40.975	43.534	48.481	57.270	79.324	**49**
50	36.870	37.245	37.586	37.901	40.255	41.933	44.533	49.562	58.508	80.988	**50**
51	37.754	38.134	38.480	38.800	41.189	42.892	45.533	50.644	59.746	82.652	**51**
	0.007	**0.008**	**0.009**	**0.01**	**0.02**	**0.03**	**0.05**	**0.1**	**0.2**	**0.4**	
n	**Loss probability (E)**										**n**

Exercise Procedure



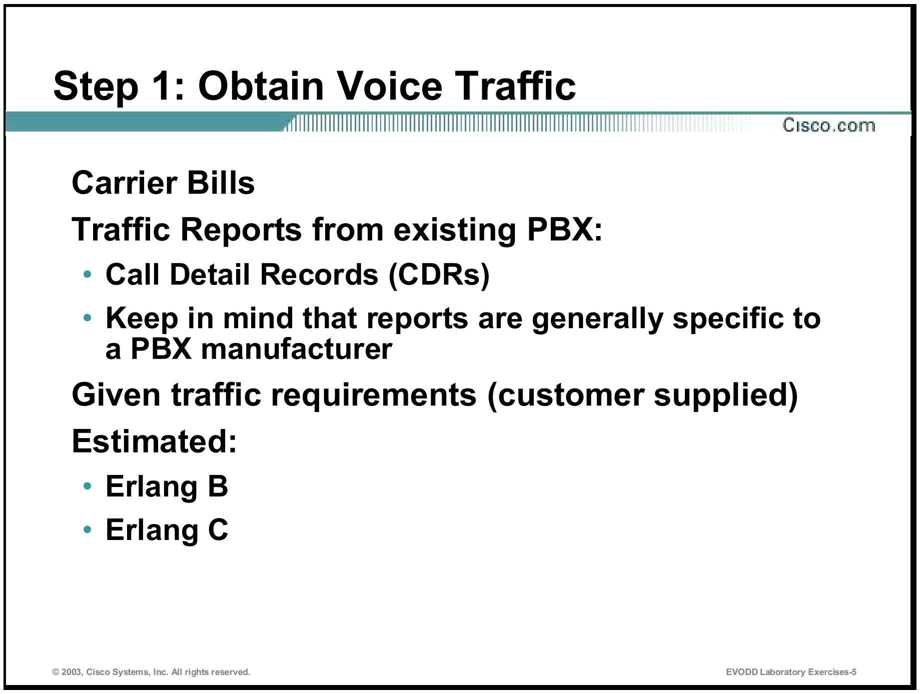
Traffic Engineering is a four-step process:

Step 1 Obtain the customer's current voice traffic pattern.

Step 2 Organize the traffic into common groups.

Step 3 Determine the number of trunks necessary for today's environment. Evaluate the network that the customer already has in place—often this network will be sufficient to meet immediate needs.

Step 4 Plan bandwidth capacity for future VoIP needs—what is the anticipated growth of the company's voice infrastructure?



Data must be collected in order to determine the customer's voice traffic pattern. The costs of collecting information are not always apparent. Data collection is an ongoing effort, which allows past predictions to be checked and, as changes take place, new predictions to be made.

Carrier bills are the first place to start an analysis. Carrier bills, however, fail to account for calls not completed. Call Detail Records (CDRs) are also used to gather information and can provide a means to cross-reference.

Step 2: Organize Traffic

Cisco.com

Inbound vs. Outbound

Call distance and destination

- **Local**
- **Long Distance**
- **International**
- **Service numbers**

Type of call

- **Fax**
- **Data**
- **Voice**

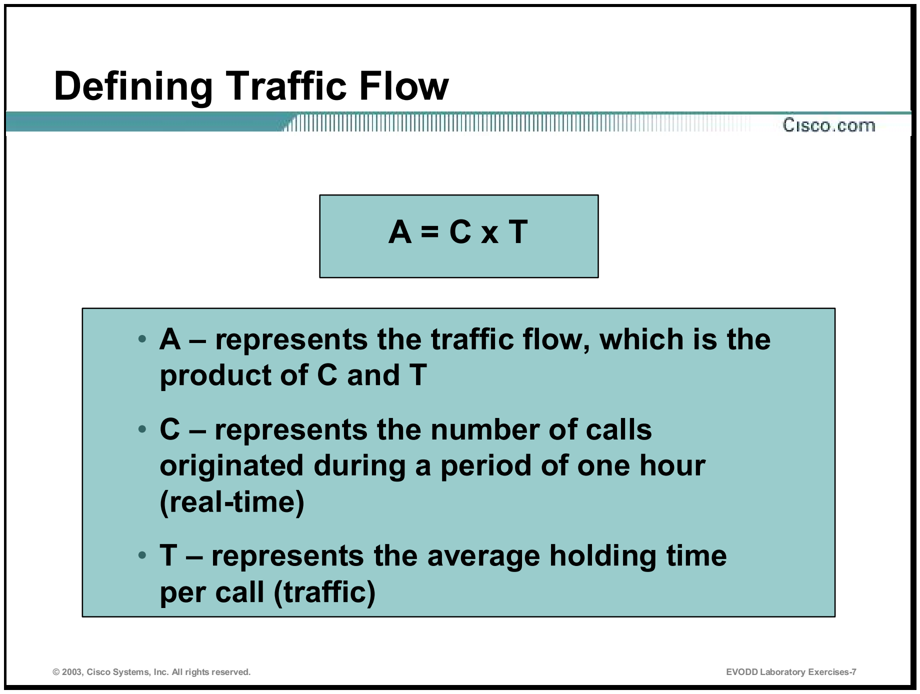
EVODD Laboratory Exercises-6

In addition to classifying the traffic into specific types, it is also usually necessary to determine the type of busy hour traffic you have in the environment.

Switch busy hour is the total of all trunks and stations. Trunk busy hour is the individual trunk group busy hour.

When planning for voice at a facility, there are three distinct areas to evaluate:

- Local traffic that is switched inside the Private Branch Exchange (PBX) or switch – this traffic is considered station-to-station traffic
- The number of trunks between the CO and the PBX
- Tie lines (Tie trunks) between facilities or a depiction of the amount of bandwidth needed to support the expected call load in a VoIP design



"C" is often referred to as offered traffic. Unfortunately, most of the information that is collected is carried traffic, so allowances must be made when field data is used.

"T" is the holding time per call, which is another important element in determining traffic flow. It indicates the time required for dialing and ringing, conversation time, and the time required to end the call. Automatic message accounting typically only monitors the conversation time, therefore, provisions must be made.

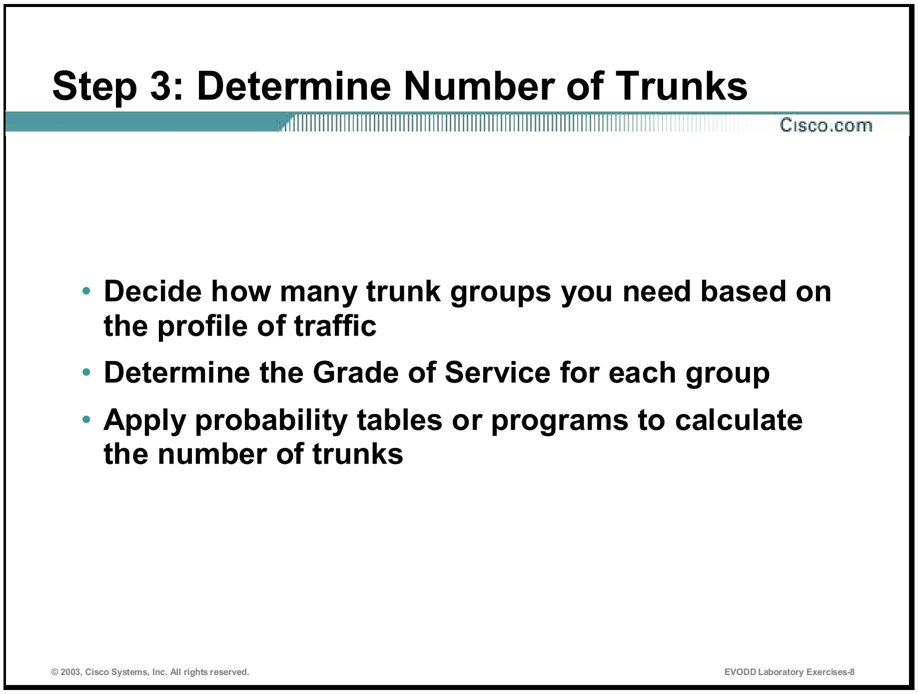
When defining traffic flow it usually helps to place the information into a common format. In some cases, the format may be call-minutes instead of call-hours. To place the information into the call-hours format, simply divide by 60. When measuring this quantity Erlangs are typically used. An Erlang is defined as the continuous use of a circuit for one hour. Another common measurement is Centum Call Seconds (CCS), also referred to as 100 call seconds.

1 Erlang= 36 CCS

Thus, common conversions would look like this:

1 Erlang= 60 minutes= 36 CCS= 3600 seconds

All of the above values are measuring the same amount of time. Calls of less than 30 seconds are typically rounded to 30 seconds.



The Grade of Service is based on the probability of a call receiving a busy tone. This value is based on the Busy Hour value, which can vary among days, weeks, and even months. The Grade of Service can differ between PBXs and Automatic Call Distribution (ACD) centers. An acceptable Grade of Service is usually unique to each organization.

- Typically, a value of B.01 for COs, DID, and T1 super trunks is acceptable. A blockage value of B.01 indicates that one out of every 100 callers may experience a busy tone.

- For more expensive T1 connections to Sprint/MCI/LD, a value of B.05 may be used. A blockage value of B.05 indicates that five out of every 100 callers may experience a busy tone.

- For Tie Lines or Tie Trunks, a value of B.10 is commonly used. A blockage value of B.10 indicates that 10 out of every 100 callers may experience a busy tone.

The standard table commonly used to represent the amount of traffic that goes through an enterprise type system in an hour is Erlang B. The standard table commonly used to represent the amount of traffic that goes through a call center in an hour is Erlang C.

To determine the total number of trunks required, the Busy Hour Traffic must be determined. Usually, the busy hour is based on a 22-day month. The percent of the traffic that occurs in the busy hour is usually 17 percent. The busy hour is always used to determine the required number of trunks.

Use the formula described here to estimate the busy hour. First, take the number of calls and multiply by the average duration of all calls (usually measured in minutes). Then, take that value and divide it by 60 to transform it into hours per month. Take the hours per month and divide by the measured month value (usually 22 days). This gives you the busy hour value. Once you have the busy hour, multiply this result by the percentage of traffic that occurs during this time (usually 17 percent). This will give you the Erlang value for trunk usage and may be cross-referenced on the Erlang table according to your applied blockage (Grade of Service).

The Busy Hour formula is shown here:

number of calls * average duration of calls / 60 = hours per month / (measured month) = Busy Hour * percent of traffic = Erlangs

Practice

Assuming 18 trunks are carrying nine Erlangs of traffic with an average call duration of three minutes, determine the following:

Q1) What is the average number of busy trunks?

Q2) What is the number of calls that can originate in an hour?

Q3) How long does it take to complete all the calls?

- Customer Traffic Information:
  - Incoming calls: 7964
  - Outgoing calls: 11233
  - Both incoming and outgoing calls are 2.5 minutes
  - 21 days/month measure
  - 17 percent Busy Hour Traffic
- Incoming Analysis:

Step 1 Determine total call hours per month (Total/month = Total incoming calls * minutes)

Step 2 Determine total call hours per day (Total/day = Total/month / #days per month)

Step 3 Determine busy hour to get Erlangs (Erlangs for incoming = Total/day * busy hour %)

- Outgoing Analysis:

Step 4 Determine total call hours per month (Total/month = Total outgoing calls * minutes)

Step 5 Determine total call hours per day (Total/day = Total/month / #days per month)

Step 6 Determine busy hour to get Erlangs (Erlangs for outgoing = Total/day * busy hour %)

- Trunk Calculation:
  - Using the calculated Erlang values, use the Erlang B table (n) to cross-reference the required number of trunks.
- Voice Capacity Planning:
  - Determine if the customer currently has enough trunk capacity to handle existing traffic flow
  - Make recommendations for future customer needs

Lesson Review

This practice exercise reviews what you have learned in this lesson.

Q1) What is the correct order for the Voice over Data deployment stages?

A) Implement, Operate, Design, and Plan

B) Plan, Design, Operate, and Implement

C) Design, Plan, Implement, and Operate

D) Plan, Design, Implement, and Operate

Q2) Which two of the following factors might be motivators for a customer who decides on a Voice over Data solution? (Choose two.)

A) cost savings

B) an opportunity to use the existing digital PBX telephones for the customer

C) IP telephony applications

D) an opportunity to increase departmental staffing

Q3) What is the Cisco recommendation regarding connecting IP Phones to a LAN?

A) use no more than three IP Phones per hub

B) connect IP Phones to switches

C) connect IP Phones to routers

D) connect IP Phones to gateways

Q4) What is an Erlang?

A) a compression mechanism for T.120 white board collaboration

B) a coder-decoder (codec) for very low-speed WAN links

C) an H.323 measure of quality

D) 3600 seconds of calls on the same circuit

Q5) Which three of the following costs are considered current costs for the customer? (Choose three.)

A) LEC

B) backup power

C) staffing

D) average busy hour call volume

Q6) Which two of the following switch models would typically be upgraded during a Voice over Data migration? (Choose two.)

A) Cisco Catalyst 1900

B) Cisco Catalyst 3500

C) Cisco Catalyst 5500

D) Cisco Catalyst 6500

Q7) Which four of the following costs should be considered when calculating staff costs? (Choose four.)

A) internal voice staff costs

B) internal data staff costs

C) contracted staff costs

D) outsourced staff costs

E) telephone provider staff costs

Q8) What is the default time that the Cisco Networking Feasibility Tool uses when estimating a project life?

A) 2 years

B) 3 years

C) 4 years

D) 5 years

Q9) What does a voice sensitivity analysis show?

A) the average daily call minutes required to achieve savings

B) the ROI for the project

C) the IRR for the project

D) the NPV for the project

2

Standards for Voice Over Data Signaling

Overview

Implementing Voice over IP (VoIP) or Voice over Frame Relay (VoFR) presents the need for new signaling methodologies, supported by protocol development. H.323 is a key protocol in VoIP networks, and FRF.11 is a key suite of protocols for VoFR. Understanding the various uses and components of H.323 and FRF.11 will aid in proper network design.

Upon completing this module, you will be able to:

- Describe the functions of H.323 standard components, signaling, and gateway protocols
- Describe VoFR design considerations, encapsulation standards, and challenges

Outline

The module contains these lessons:

- VoIP Gateway Protocols
- Voice Over Frame Relay Design Considerations

VoIP Gateway Protocols

Overview

The emergence of Voice over IP (VoIP) applications and IP telephony has paved the way for revising the H.323 specification. This lesson will teach you about IP telephony signaling and protocols. You will learn about the development of the various versions of H.323.

Importance

H.323 is a key protocol in Voice over Data networks. Proper design of Voice over Data networks requires an understanding of H.323 functionality and features, as well as the differences among H.323 versions.

Objectives

Upon completing this lesson, you will be able to:

- Describe the major differences between legacy and IP telephony signaling
- Describe the current versions of H.323
- Identify H.323 standard components
- Describe the function of each H.323 standard component
- Describe H.323 signaling protocols
- Describe call control signaling and setup
- Identify gateway control protocols
- Describe Media Gateway Control Protocol
- Describe session initiation protocol

Learner Skills and Knowledge

To fully benefit from this lesson, you must have these prerequisite skills and knowledge:

- A fundamental understanding of TCP and the Open System Interconnection (OSI) protocol stack
- A fundamental understanding of Cisco Architecture for Voice, Video and Integrated Data (AVVID)

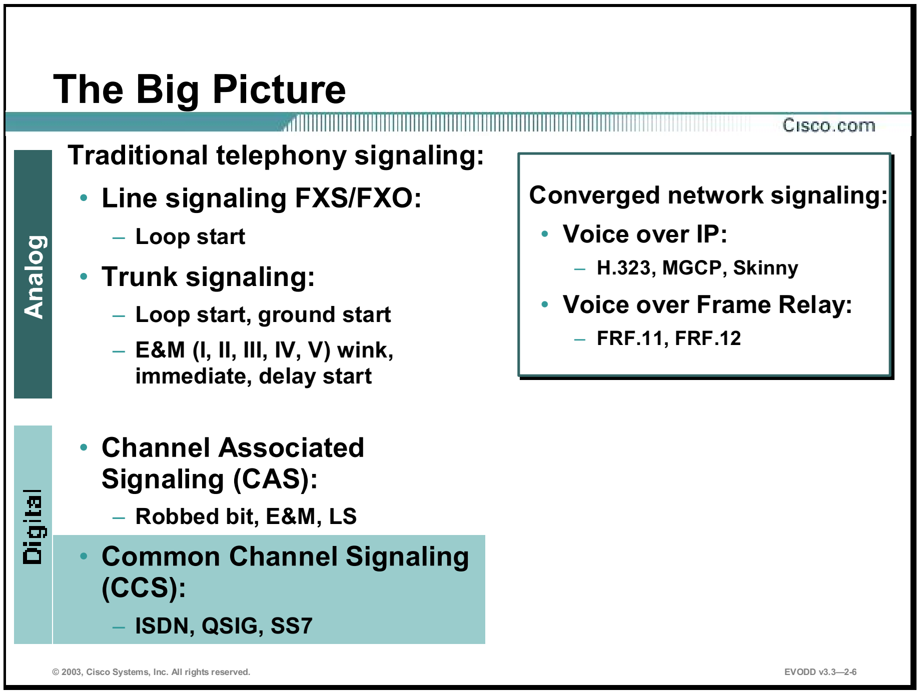
Outline

This lesson includes these topics:

- Overview
- Introduction
- H.323 Versions
- Components
- Component Details
- Registration, Admission, and Status
- Call Control and Setup
- Protocols
- Media Gateway Control Protocol
- Session Initiation Protocol
- Summary
- Lesson Review

Introduction

This topic examines network signaling.

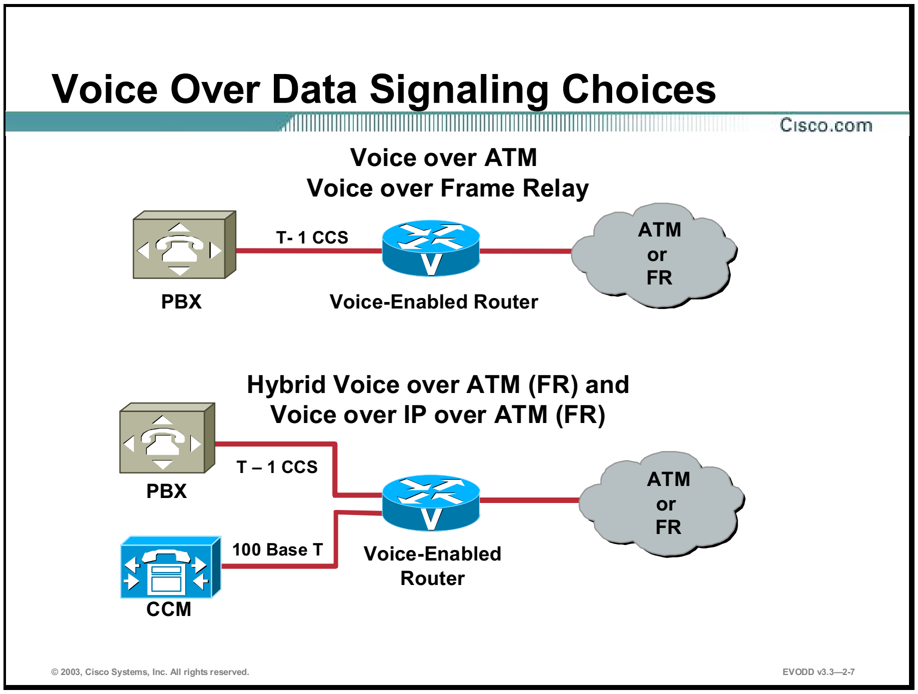


Digital signaling in the voice network is handled in one of two ways: channel associated signaling (CAS) or common channel signaling (CCS). CAS, which is also known as robbed-bit signaling or in-band signaling, is a bit-oriented signaling scheme. CCS is a message-oriented scheme in which signaling is sent out of a band in a message format.

Note	Messages are simply data packets. Signaling is handled in a similar way in VoIP technologies.

The current Cisco Architecture for Voice, Video and Intergrated Data (AVVID) Cisco CallManager (CCM) supports H.323 and Media Gateway Control Protocol (MGCP). Cisco IP Phones use a protocol called Skinny Station protocol, which is similar to the standard Simple Gateway Control Protocol (SGCP).

Reference	You can find more information regarding CAS and CCS at this website: http://www.cisco.com/warp/public/788/voip/voip_cas.htm.



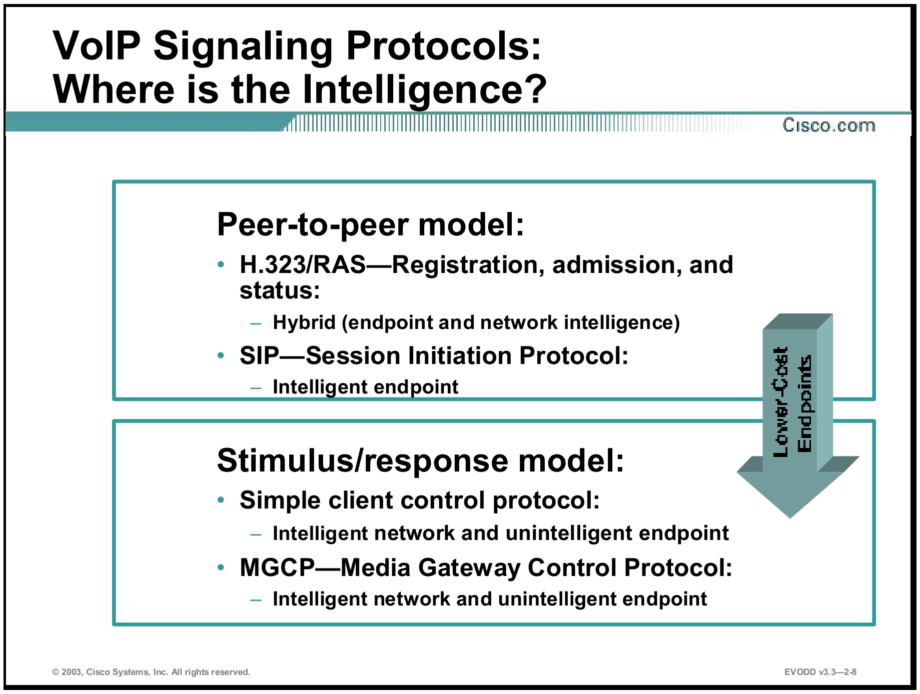
Voice over Data often implies Voice over IP (VoIP); however, Voice over ATM (VoATM) and Voice over Frame Relay (VoFR) are also common ways to implement Voice over Data networks. Additionally, VoIP runs over ATM or Frame Relay in the WAN, which results in a hybrid solution: Voice over IP over ATM (VoIPoATM) or Voice over IP over Frame Relay (VoIPoFR). When you are designing a solution that uses an existing ATM or Frame Relay network, you must decide whether to provide VoATM or VoFR; or VoIPoATM or VoIPoFR.

The VoATM/VoFR solution, shown at the top of the slide, uses a PBX that is connected directly to a voice-enabled router. The router acts as a gateway between the telephony and data worlds. IP is not part of this solution. Calls are placed on static channels. The PBX and the voice-enabled router have no packets and no addressing needs between them.

A hybrid solution is shown at the bottom of the slide. This example uses VoIPoATM Frame Relay and VoATM Frame Relay. The connection between the voice-enabled router and the PBX is similar to the VoATM/VoFR solution because the router acts as a gateway between the telephony and data worlds.

A call that originates from a telephone controlled by CCM, and will ultimately arrive at a telephone connected to the PBX, starts out as a VoIP over Ethernet call. As these packets pass through the router, the call converts to CCS in order to be sent to the PBX. A call that comes from a telephone connected to the PBX, and arrives at a telephone connected to CCM, starts out as a CCS call. When the call reaches the router, it is converted to IP packets and sent to CCM, which switches the call to the proper destination telephone.

Calls crossing the ATM or Frame Relay WAN that originate or terminate at the PBX need VoATM or VoFR service. Calls crossing the ATM/FR WAN that originate or terminate at an IP Phone that is serviced by CCM need VoIPoATM or VoIPoFR service.



You must still use signaling (supervisory, call progress, or addressing) when voice travels over a data network.

Multiple standards-based protocols, intended to solve the signaling function for Voice over Data networks, are in different stages of development. Cisco implements protocols as they are developed and as the marketplace demands them. Currently, Cisco has implemented H.323, MGCP, and session initiation protocol (SIP) in VoIP.

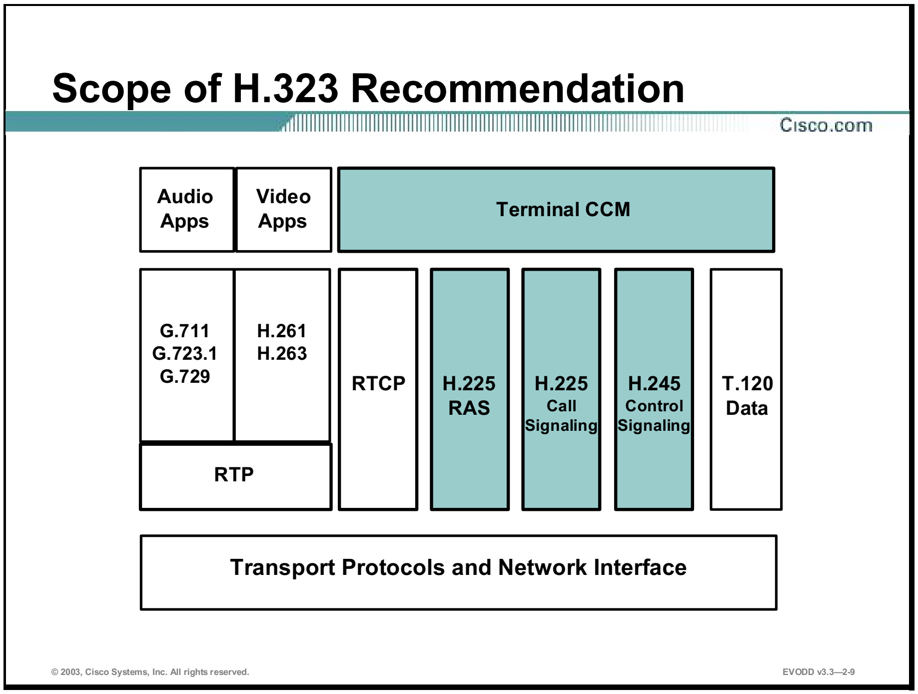
You can characterize the H.323 and SIP protocols as peer-to-peer; both endpoints are intelligent. SIP is used primarily as a *call agent to call agent* protocol. Cisco will support this protocol in future Cisco IOS and Cisco CCM releases.

Simple Client Control Protocol and its successor, MGCP, are based on a stimulus and response model. This model is similar to the client and server model that keeps the intelligence centralized in the network rather than in the endpoints. The main benefit of this model is lower-cost endpoints that can be controlled from intelligent network controllers. The cost of production for the endpoint is lower and the rate at which it becomes obsolete slows. For example, Flash RAM in an IP Phone device can be updated with new software downloaded from a network agent.

CCM implements a proprietary version of Simple Client Control Protocol with CCM communicating to the Cisco IP Phone.

H.323 Versions

This topic covers the details of the original H.323 specification, as well as new versions of H.323. New network requirements are emerging because of the development of VoIP. A new version of H.323 was developed to accommodate these additional requirements, and revisions will be needed as features are added to the H.323 standard.



The slide illustrates the scope of the H.323 recommendation and attempts to provide a context among its components and the related components encountered in IP telephony.

The shaded section delineates the H.323 components.

Note Video compressors and/or decompressors (codecs) are not a concern for this module.

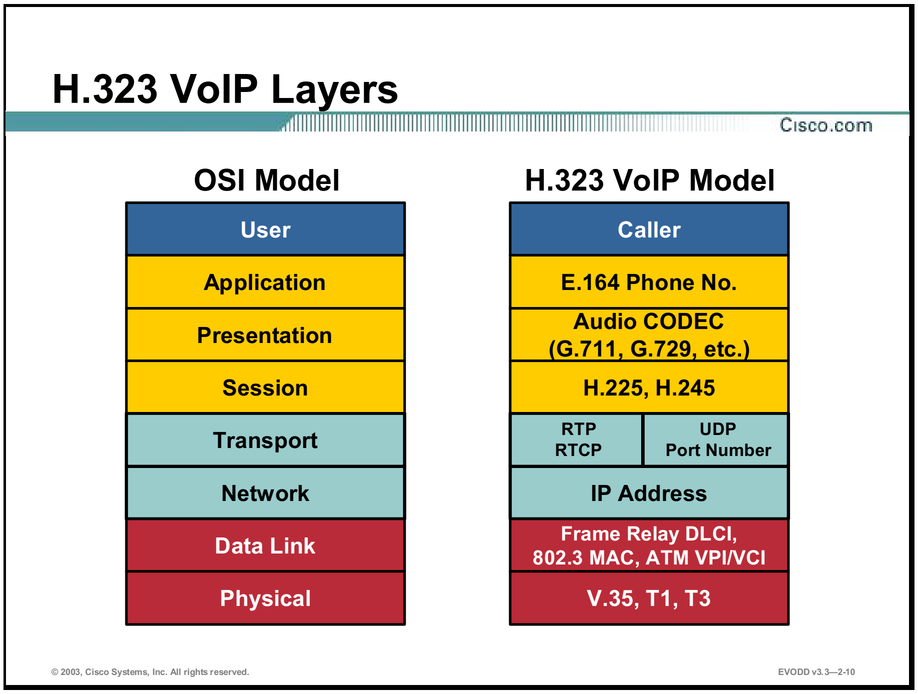
H.323 includes control protocols for the system, and the H.225 registration, admission, and status (RAS) protocol.

Another important component, which is not part of H.323, is the Real-Time Transport Protocol (RTP) and its control protocol, the Real-Time Transport Control Protocol (RTCP). These protocols specify how voice information is carried in IP telephony networks.

RTP provides end-to-end delivery services of real-time audio and video. RTP also provides payload-type identification, sequence-numbering, time-stamping, and delivery-monitoring. RTP is typically used to transport data via the User Datagram Protocol (UDP). RTP, together with UDP, provides transport-protocol functionality. UDP provides multiplexing and checksum services.

Note You can also use RTP with other signaling protocols.

RTCP is the counterpart of RTP that provides control services. The primary function of RTCP is to provide feedback on the quality of the data distribution. The secondary RTCP function is to carry a transport-level identifier for an RTP source, called a canonical name, which receivers use to synchronize audio and video.



The diagram compares the functionality necessary to deliver VoIP in a layered manner, analogous to the Open System Interconnection reference model (OSI).

An example of an application layer can be either a telephone number called from NetMeeting, or an e-mail address from Outlook.

The audio codec is represented at the presentation layer.

The session layer contains the call control functions provided by H.225 and H.245.

The transport layer is a constant for IP telephony: it is always RTP/RTCP and UDP. RTP/RTCP (RFCs 1889/1890) provides the end-to-end network transport function. RTP, a thin protocol, was designed for streaming real-time data, such as VoIP. RTP provides content type identification (voice, video, compression type), loss detection via sequence-numbering, time-stamping, and delivery-monitoring. RTCP provides feedback on the quality of the distribution.

The network layer consists of the IP address and a quality of service (QoS)-enabled network infrastructure, including these Layer 2 data link choices: MAC, Frame Relay, and ATM. You can run these data link choices over a variety of physical interfaces, such as V.35, T1, or T3.

A key point to remember is that H.323 is a peer-to-peer networking protocol, and each endpoint must be fully H.323 compliant. When you review Cisco AVVID architecture, you learn that Cisco uses H.323 between CCMs (think of CCM as an IP version of a PBX), but runs a client/server protocol, called Skinny Station Protocol, between CCM and the IP Phone sets. As a result, the IP Phone endpoint needs a full H.323 stack, which reduces processing requirements and slows obsolescence.

H.323 Versions

- International (ITU-T) standard for packet-based multimedia communications systems
- Version history:
  - Version 1 established in 1996
  - Version 2 in January 1998 (IOS 12.0(4)T)
  - Version 3 in September 1999 (IOS 12.2(2)XA)
  - Version 4 in November 2000
- Cisco supports H.323 version 3

 EVODD v3.3—2-11

The International Telecommunication Union-Telecommunication Standardization Sector (ITU-T) Study Group 16 specified the H.323 standard. Version 1 of the H.323 recommendation—visual telephone systems and equipment for LANs that provide a non-guaranteed QoS—was accepted in October 1996. Version 1 is heavily weighted toward multimedia communications in a LAN environment. Version 1 of the H.323 standard does not provide guaranteed QoS.

The emergence of VoIP applications and IP telephony paved the way for a revision of the H.323 specification. The absence of a standard for VoIP resulted in products that were incompatible. With the development of VoIP, new requirements emerged, such as providing communication between a PC-based telephone and a telephone on a traditional switched circuit network (SCN). These requirements forced the need for an IP telephony standard. Version 2 of H.323—packet-based multimedia communications systems—was defined to accommodate these additional requirements and was accepted in January 1998.

H.323 Version 3 was approved in September 1999. Version 3 made only modest improvements over Version 2, by adding such features as: caller ID, language preference, and remote device control.

The current version of H.323 is Version 4, which was approved in November 2000. Version 4 contains enhancements in a number of important areas, including reliability, scalability, and flexibility. Enhancements include: call intrusion, which is the ability to interrupt an existing call; real-time fax enhancement, which converts a voice call into a fax call; and dual tone multifrequency (DTMF) relay via RTP, which accurately transmits DTMF digits, in-band.

Cisco supports H.323 Version 3, beginning with Cisco IOS release 12.2(2)XA. This list highlights some of the features involved in this upgrade:

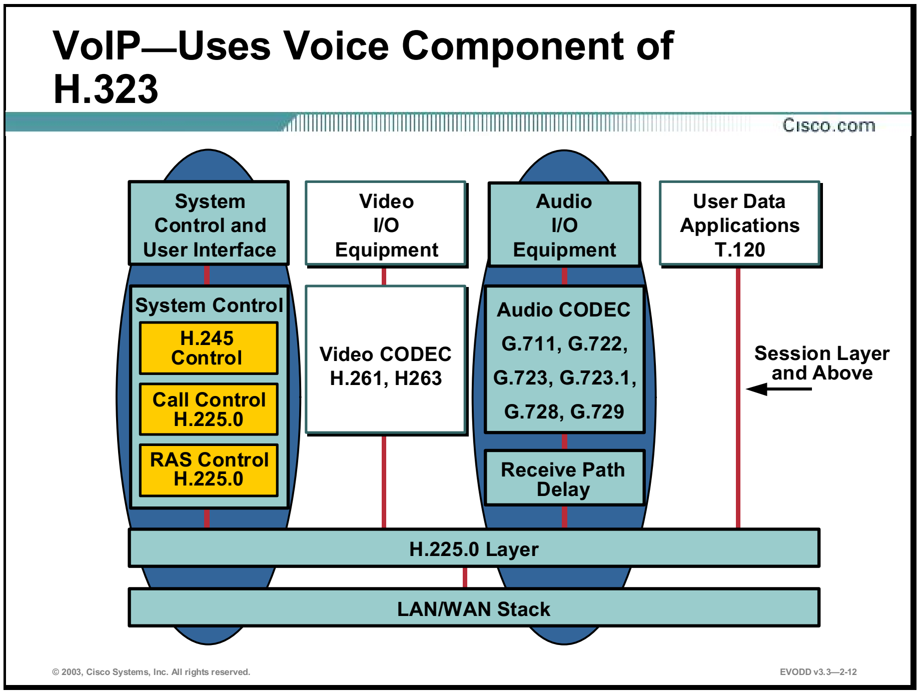
- Support for mandatory H.323 Version 3 elements in the gateway and gatekeeper

- multipleCalls
- maintainConnection
- alternateTransportAddresses
- useSpecifiedTransportSupport for H.225 call signaling over UDP (H.225 messages can be transported over TCP or UDP [as described in Annex E]. During registration, a Cisco H.323 gateway indicates to the gatekeeper whether it is capable of transmitting over both TCP and UDP. If it can transmit over both, then the Cisco H.323 gateway registers both of its TCP and UDP addresses.)

- Address resolution using border elements (BE)

Components

This topic examines the H.323 standard components.

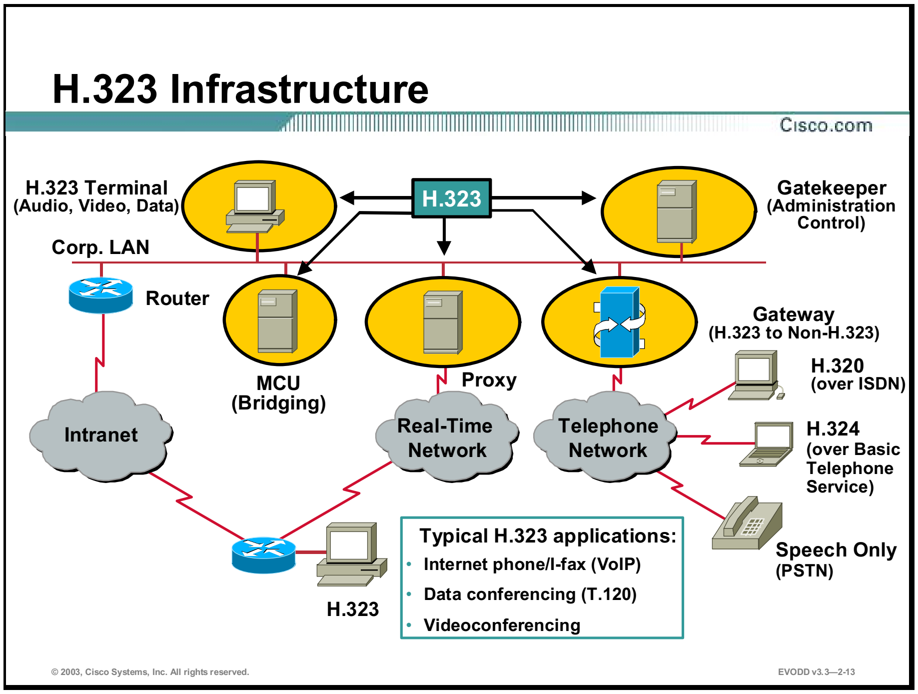


VoIP does not utilize all of the H.323 components. The components delineated by ovals in the illustration are of interest for IP telephony.

Note H.323 is independent of the packet network and the transport protocols, and does not specify them.

Protocols specified by H.323 include the following:

- Audio codecs
- Video codecs
- H.225 RAS
- H.225 call signaling
- H.245 control signaling
- RTP
- RTCP



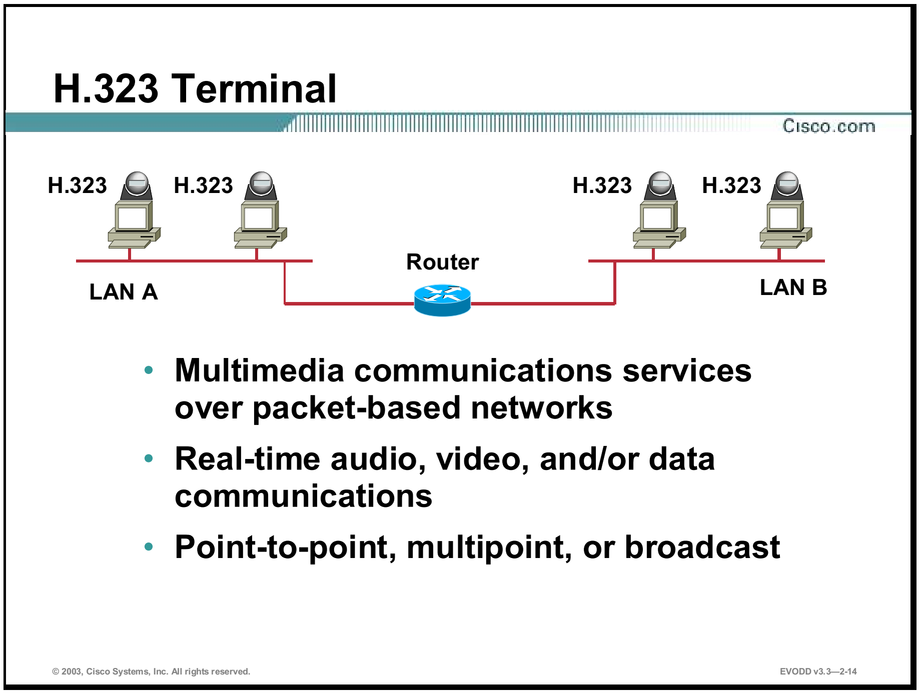
The H.323 standard specifies four kinds of components. When networked together, these components provide point-to-point and point-to-multipoint multimedia communication services. These components are:

- Terminals
- Gateways
- Gatekeepers
- Multipoint control units (MCUs)

The gatekeeper, gateway, and MCU are optional, in the context of the H.323 standard.

Component Details

This topic describes the processes that are handled by each of the four components within the H.323 standard. The components of an H.323 network support many functions within the network and connectivity to external networks as well.



An H.323 terminal is used for real-time, bi-directional multimedia communications. An H.323 terminal can either be a PC running H.323 terminal software or a standalone device. The terminal supports audio communications and can optionally support video or data communications. The basic service of an H.323 terminal is audio communications, which plays a key role in IP telephony services.

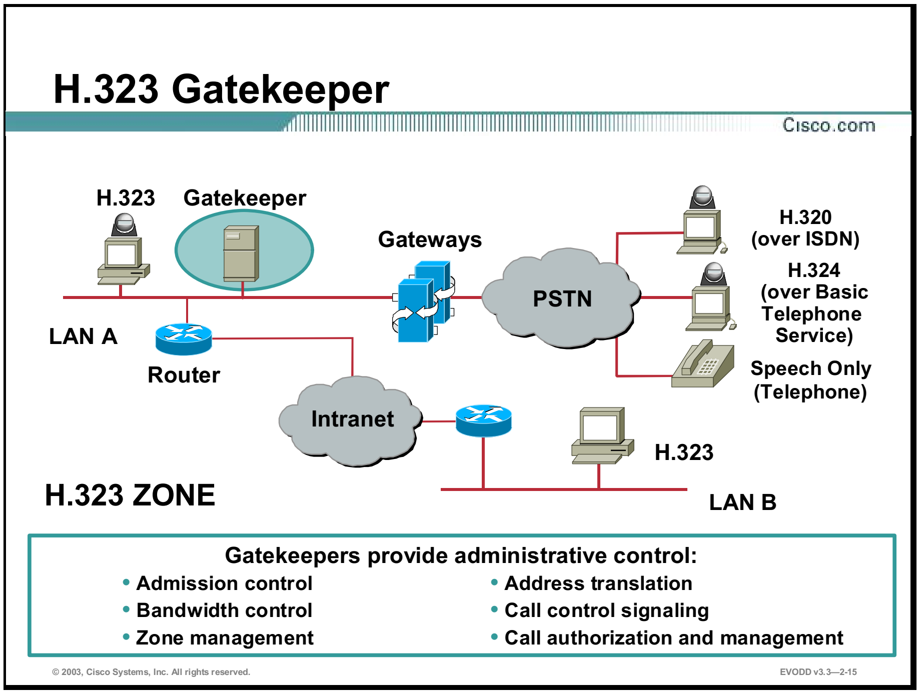
The primary goal of H.323 is to work with other multimedia terminals. H.323 terminals are compatible with H.324 terminals on SCNs and wireless networks, H.310 terminals on Broadband ISDN (BISDN), H.320 terminals on ISDN, H.321 terminals on BISDN, and H.322 terminals on guaranteed bandwidth packet switched networks. You can use H.323 terminals in multipoint conferences.

H.323 terminals must support the following:

- H.245 for exchanging terminal capabilities and creation of media channels
- H.225 for call signaling and call setup
- RAS for registration and other admission control with a gatekeeper
- RTP/RTCP for sequencing audio and video packets

- G.711 audio codec

Note The optional components in an H.323 terminal are video codecs, T.120 data-conferencing protocols, and MCU capabilities.



Calls originating within an H.323 network may use an alias to address the destination terminal. Calls originating outside of the H.323 network and received by a gateway may use an E.164 telephone number (such as, 310-442-9222) to address the destination terminal. The gatekeeper translates this E.164 telephone number, or the alias, into the network address (such as, 204.252.3.56:1720 for an IP–based network, where 1720 is a port number) for the destination terminal. The destination endpoint can be reached using the network address on the H.323 network.

The gatekeeper can control the admission of the endpoints into the H.323 network. The gatekeeper uses RAS messages—Admission Request (ARQ), Admission Confirmation (ACF), and Admission Reject (ARJ)—to achieve this control. Admissions control may also be a null function that admits all endpoints to the H.323 network.

The gatekeeper provides support for bandwidth control by using these RAS messages: Bandwidth Request (BRQ), Bandwidth Confirmation (BCF), and Bandwidth Reject (BRJ). For example, if a network manager has specified a threshold for the number of simultaneous connections on the H.323 network, the gatekeeper can refuse to make any connections once the threshold is reached. The result is to limit the total allocated bandwidth to some fraction of the total bandwidth available, leaving the remaining bandwidth for data applications. Bandwidth control may also be configured to accept all requests for bandwidth changes.

The gatekeeper provides four functions: address translation, admissions control, bandwidth control for endpoints, and zone management.

Mandatory and Optional Gatekeeper Functions

Cisco.com

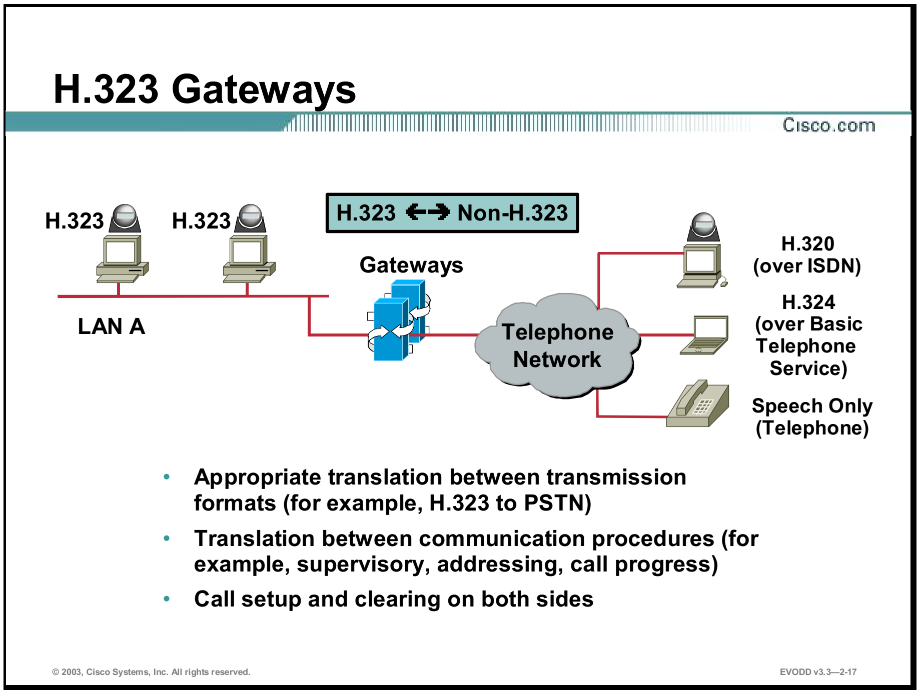
- **Gatekeeper is an optional component of H.323**
- **Required features:**
  - **Address translation (alias to transport within zone)**
  - **Admissions control**
  - **Bandwidth control during the call**
  - **Zone management**
- **Optional features:**
  - **Call control signaling/routing (under Gatekeeper control)**
  - **Call authorization**
  - **Call management (call status, tracking, PBX-like services, etc.)**

 EVODD v3.3—2-16

The gatekeeper can route call-signaling messages between H.323 endpoints. In a point-to-point conference, the gatekeeper may process H.225 call-signaling messages. Alternatively, the gatekeeper may allow the endpoints to send H.225 call-signaling messages directly to each other.

When an endpoint sends call-signaling messages to the gatekeeper, the gatekeeper may accept or reject the call, according to the H.225 specification. The reasons for rejection may include access-based or time-based restrictions to and from particular terminals or gateways.

The gatekeeper may maintain information about all active H.323 calls, which allows the gatekeeper to control its zone by providing the information to the bandwidth management function, or by routing the calls to different endpoints to achieve load balancing.



A gateway connects two dissimilar networks. An H.323 gateway provides connectivity between an H.323 network and a non-H.323 network. For example, an H.323 gateway can connect to, and provide communication between, an H.323 terminal and an SCN. The SCN includes all switched telephony networks, for example, the Public Switched Telephone Network (PSTN). Connecting dissimilar networks is achieved by translating protocols for call setup and release, converting media formats between different networks, and transferring information between the networks connected by the gateway.

Note	A gateway is not required for communication between two terminals on an H.323 network.

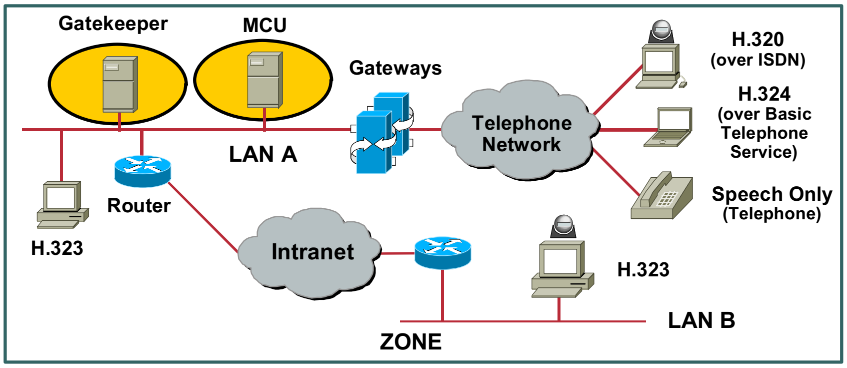
A gateway provides translation of protocols for call setup and release, conversion of media formats between different networks, and the transfer of information between H.323 and non-H.323 networks.

IP telephony contains an application of the H.323 gateway, wherein the H.323 gateway connects an IP network and SCNs (such as the ISDN network). On the H.323 side, a gateway runs H.245 control signaling for exchanging capabilities, H.225 call signaling for call setup and release, and H.225 RAS for registration with the gatekeeper. On the SCN side, a gateway runs SCN-specific protocols (such as, ISDN and SS7 protocols).

Note	In Cisco AVVID IP telephony, H.323 gateways are implemented in the DT24+ and DE30+ products, as well as Cisco IOS gateways and routers. The DT24+ and DE30+ products are now EOL (End of Life).

H.323 Multipoint Control Unit

Cisco.com



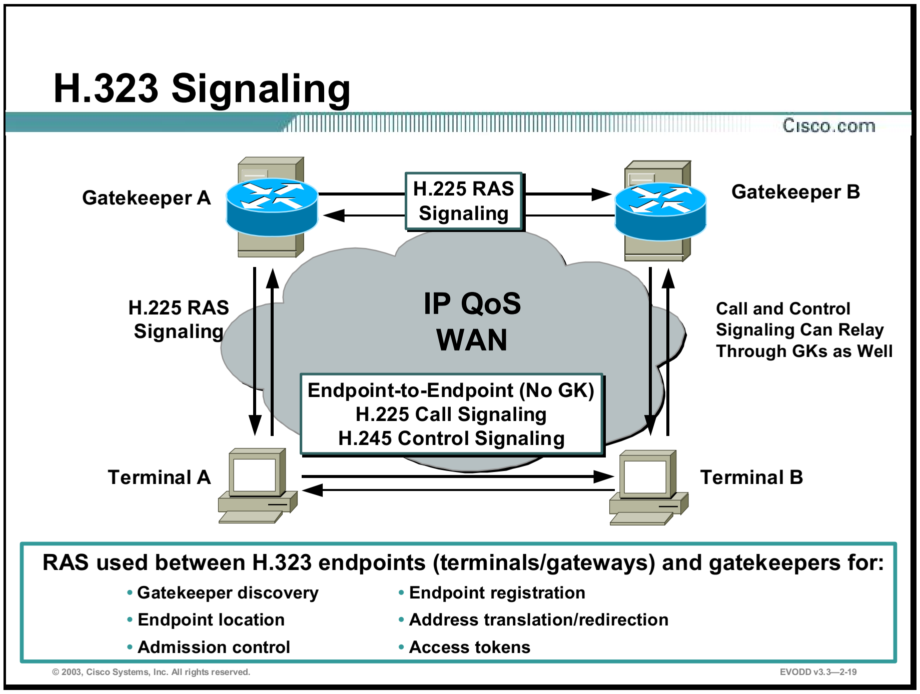
- An endpoint that provides support for multipoint conferences
- An MCU consists of a multipoint controller (MC) and one or more multipoint processors (MP)
- Endpoints establish a point-to-point connection with the MC
- Actual video or audio distribution may be centralized or distributed

EVODD v3.3—2-18

MCUs provide support for conferences of three or more H.323 terminals. All terminals participating in the conference establish a connection with the MCU. The MCU manages conference resources, negotiates between terminals to determine the audio or video codec to use, and may handle the media stream. The gatekeepers, gateways, and MCUs are logically separate components of the H.323 standard, but can be implemented as a single physical device.

Registration, Admission, and Status

This topic describes how the RAS protocol functions. RAS is the protocol between endpoints (terminals and gateways) and gatekeepers.



This illustration provides an overview of H.323 signaling protocols. Signaling can occur between the following:

- Terminals and gatekeepers
- Gatekeepers
- Terminals directly, without gatekeepers

Note Remember that gatekeepers, although common, are not required in H.323.

A RAS channel exchanges RAS messages. The RAS channel can be unreliable, so a RAS message exchange may be associated with timeouts and retry counts.

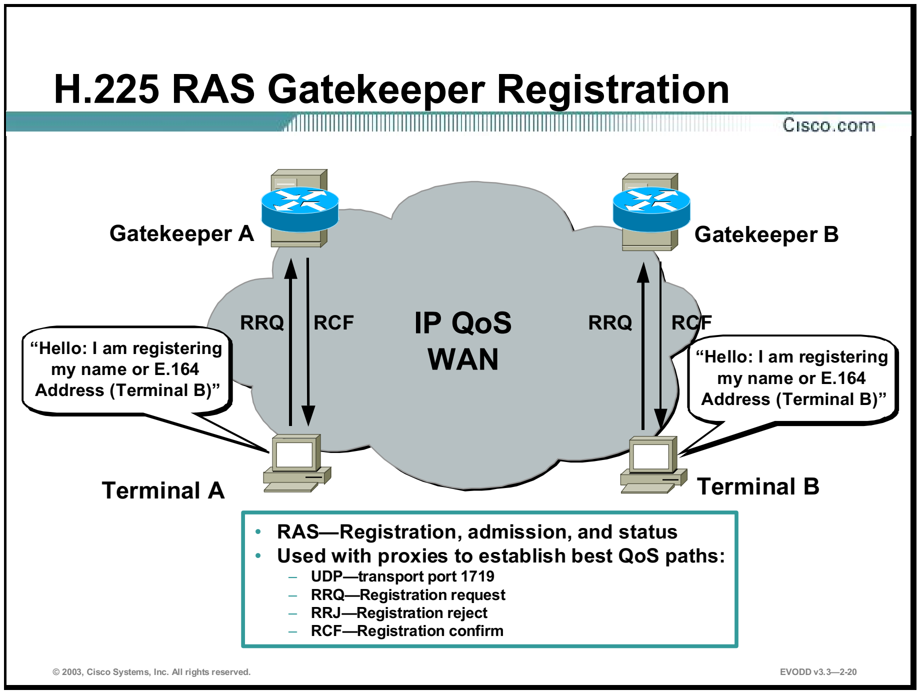
H.323 endpoints use the gatekeeper discovery process to register with a gatekeeper. The discovery can be done statically or dynamically. In static discovery, the endpoint knows the transport address of its gatekeeper beforehand. In dynamic discovery, the endpoint multicasts a Gatekeeper Request (GRQ) message on the discovery multicast address of the gatekeeper: "Who is my gatekeeper?" One or more gatekeepers may respond with a Gatekeeper Confirmation (GCF) message: "I can be your gatekeeper."

Endpoints use the registration process to join a zone and inform the gatekeeper of the transport and alias addresses for that zone. All endpoints register with a gatekeeper as part of their configuration.

The transport address of an endpoint is determined through endpoint location. The endpoint location process assigns the endpoint an alias name or E.164 address.

The RAS channel is used for other kinds of control mechanisms. When an endpoint is disassociated from a gatekeeper and its zone, then the channel can be used for admission control, bandwidth control, disengagement control, and to restrict the entry of an endpoint into a zone.

The Cisco IP telephony AVVID configuration uses H.323 between gatekeepers, MCUs, and gateways, but uses the proprietary Skinny Station Protocol between the MCUs and IP Phones.



This illustration demonstrates call setup and placement between two terminals, each associated with a separate gatekeeper. The terminal sends a Registration Request (RRQ) to the gatekeeper, which sends back a Registration Confirmation (RCF). Registration is the mechanism that allows terminal IP addresses to be known across zones.

Prior to H.323 Version 2, Cisco gateways registered with the gatekeeper every 30 seconds. Each registration renewal used the same process as the initial registration, even though the gateway was already registered with the gatekeeper. This registration process generated considerable overhead at the gatekeeper.

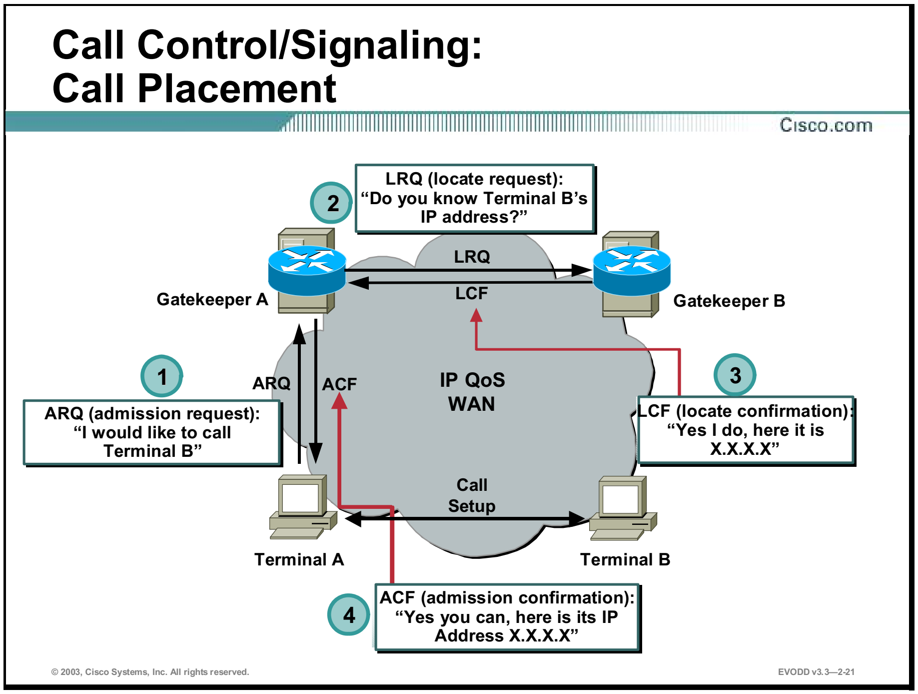
H.323 Version 2 defines a lightweight registration procedure that still requires the full registration process for initial registration, but uses an abbreviated renewal procedure to update the gatekeeper and minimize overhead.

Lightweight registration requires each endpoint to specify a Time to Live (TTL) value in its RRQ message. When a gatekeeper receives an RRQ message with a TTL value, it returns an updated TTL timer value in a RCF message to the endpoint. Shortly before the TTL timer expires, the endpoint sends an RRQ message with the keepalive field set to TRUE, which refreshes the existing registration.

An H.323 Version 2 endpoint is not required to indicate a TTL value in its RRQ message. If the endpoint does not indicate a TTL value, the gatekeeper assigns one and sends it to the gateway in the RCF message. No configuration changes are permitted during a lightweight registration, so all fields—other than the endpointIdentifier, gatekeeperIdentifier, tokens, and TTL—are ignored. In the case of H.323 Version 1 endpoints that cannot process the TTL field in the RCF, the gatekeeper probes the endpoint with Interrupt Requests for a predetermined grace period to see if the endpoint is still alive.

Call Control and Setup

This topic describes call control signaling and setup. The H.323 standard specifies four kinds of components: terminals, gateways, gatekeepers, and MCUs. When networked together, these components provide point-to-point and point-to-multipoint multimedia communication services.

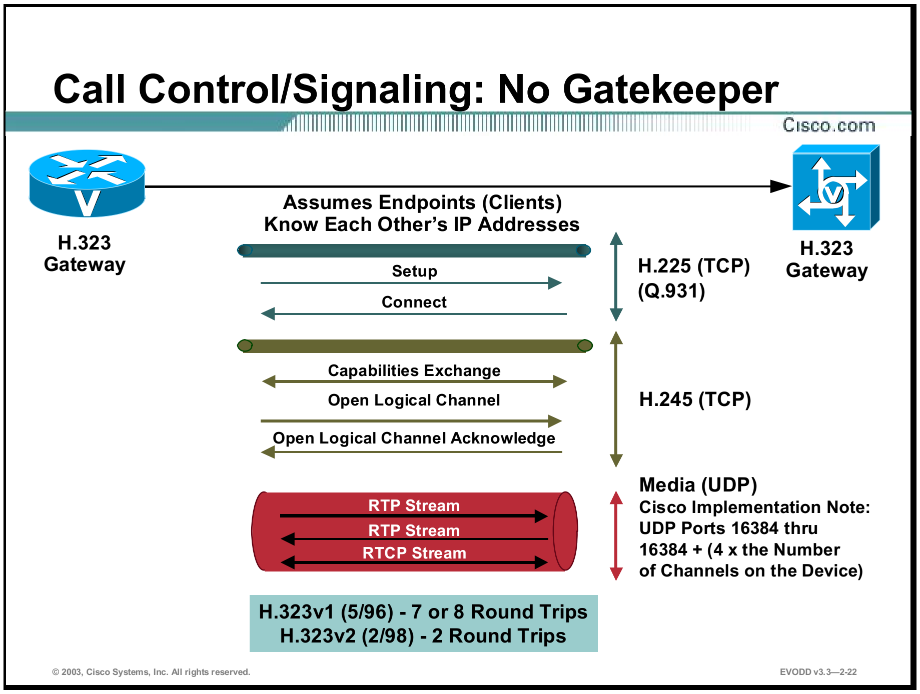


This illustration demonstrates call setup and placement between two terminals, each associated with a separate gatekeeper.

H.225 call signaling is used to set up connections between H.323 endpoints (terminals and gateways). Real-time data is transported over these connections. Call signaling involves the exchange of H.225 protocol messages over a reliable call-signaling channel. For example, H.225 protocol messages are carried over TCP in an IP-based H.323 network. H.225 messages are exchanged between the endpoints if there is no gatekeeper in the H.323 network. When a gatekeeper exists in the network, the H.225 messages are exchanged, either directly between the endpoints (direct call signaling), or between the endpoints after being routed through the gatekeeper (gatekeeper-routed call signaling). The gatekeeper chooses the method during RAS admission message exchange.

The admission messages are exchanged between endpoints and the gatekeeper on RAS channels. The gatekeeper receives the call-signaling messages on the call-signaling channel from one endpoint and then routes them to the other endpoint on its call-signaling channel.

During ACF, the gatekeeper indicates that the endpoints can exchange call-signaling messages directly. Endpoints exchange call signaling on the call-signaling channel.

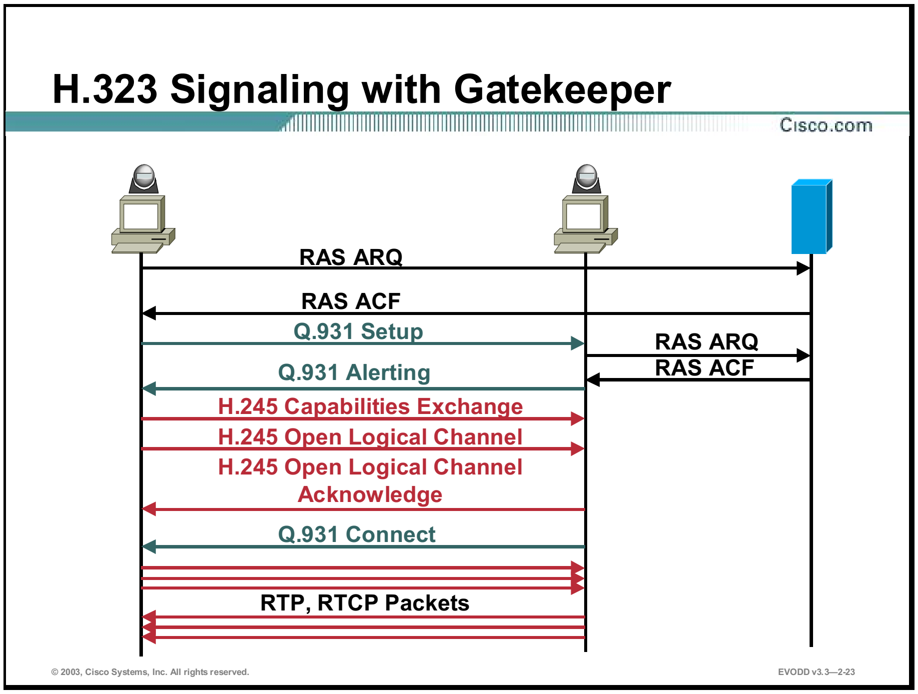


H.323 does not require gatekeepers; therefore signaling can occur directly between terminals.

H.245 control signaling consists of the exchange of end-to-end H.245 messages between communicating H.323 endpoints. The H.245 control messages are carried over H.245 control channels. The H.245 control channel is the logical channel 0 and is open permanently, unlike the media channels. The messages carry the exchange capabilities of terminals, and open and close logical channels.

Capabilities exchange is a process that uses the exchange messages of communicating terminals to provide transmit and receive capabilities to the peer endpoint. Transmit capabilities describe the ability of the terminal to transmit media streams. Receive capabilities describe the ability of the terminal to receive and process incoming media streams.

A logical channel carries information from one endpoint to another endpoint (in the case of a point-to-point conference), or to multiple endpoints (in the case of a point-to-multipoint conference). H.245 provides messages to open or close a logical channel. A logical channel is unidirectional.



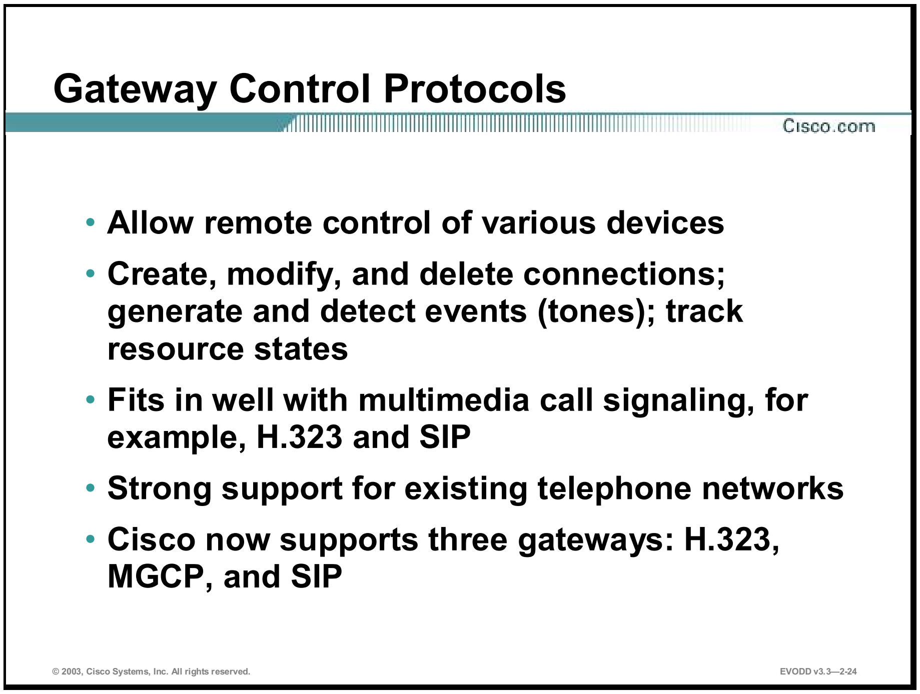
This illustration demonstrates the added protocol interactions that occur when a gatekeeper is involved in the exchange between the two terminals. Both the initiating and responding terminals must exchange RAS ARQ and RAS ACF before capabilities can be exchanged and a logical channel is opened. Once the exchange occurs, the gatekeeper drops out of the exchange.

The gatekeeper provides the following functions:

- **Address Translation:** Translates H.323 aliases (for example, sliu@cisco.com) or E.164 aliases (such as, standard telephone numbers) into IP transport addresses (for example, 10.1.1.1 port 1720)
- **Admissions Control:** Authorizes access to the H.323 network
- **Bandwidth Control:** Manages endpoint bandwidth requirements
- **Zone Management:** Provides these functions to all terminals, gateways, and MCUs that register with the zone

Protocols

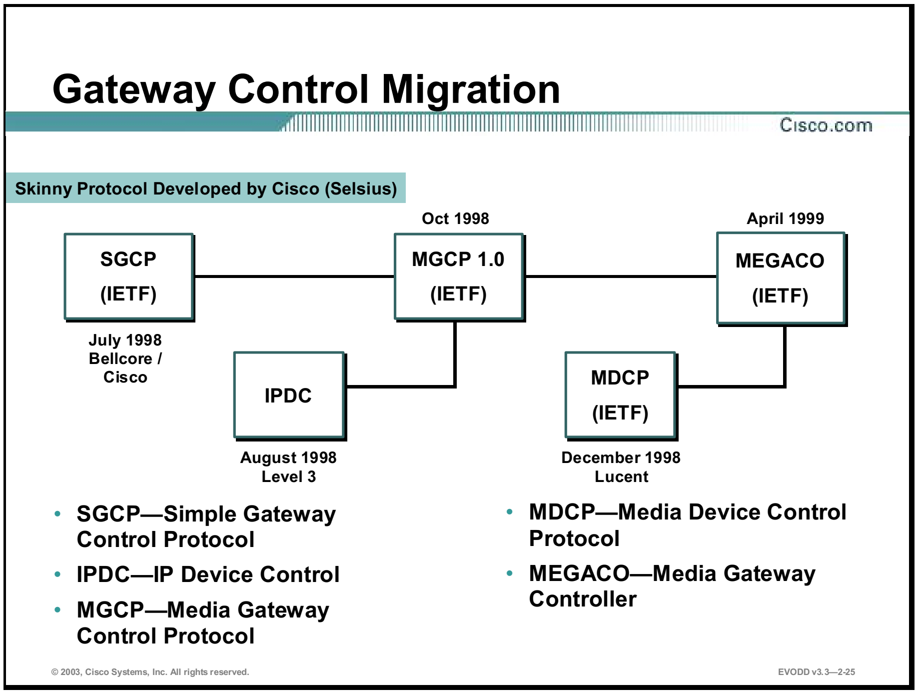
This topic examines gateway control protocols. The protocols used for VoIP calls are evolving at a rapid pace. While certain protocols will be more widely supported than others, it is important to be familiar with all of them.



Gateways bridge H.323 conferences to other networks, communications protocols, and multimedia formats. Typically, gateways are used to connect an IP telephony network with the PSTN.

CCM and Cisco IOS gateways support three gateway protocols: SIP, MGCP, and H.323.

Note	SIP is expected to receive more support in the future. Cisco continues to support SGCP.



The Skinny Station Protocol is the Cisco proprietary protocol that runs between the IP Phones and the IP PBX (CCM). Skinny Station Protocol is similar to SGCP because the intelligence is in CCM, with a master-slave type of relationship between CCM and the IP Phone.

SGCP/MGCP, H.323, and SIP are variants of ways to make VoIP calls. Each protocol solves different problems, in terms of topology and the devices connected to the network. Cisco has plans to support all of these protocols in various platforms where the protocols are necessary for overall solutions.

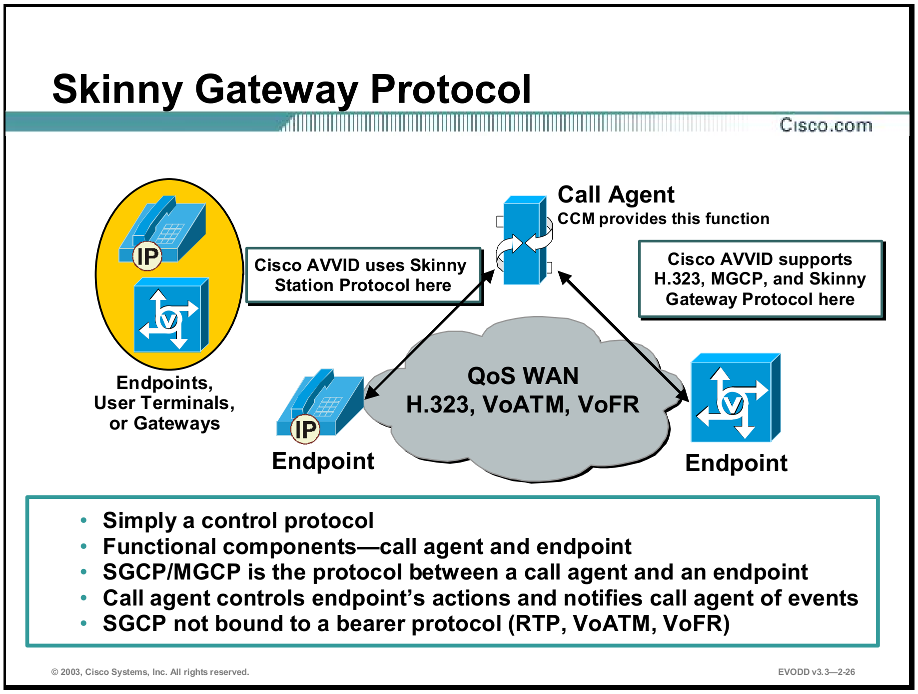
Telcordia (formerly Bellcore) and Cisco developed SGCP as an improvement on H.323 with the intention that SGCP would be more appropriate and scalable for data transmissions. Subsequently, another protocol was developed for Level 3 that defined the handling of voice and multimedia traffic over the Internet. The Level 3 protocol is called Internet Protocol Device Control (IPDC). With urging from the Internet Engineering Task Force (IETF), Telcordia and Level 3 developers announced the merger of their individual specifications into a new protocol called MGCP. The following list provides a protocol background review:

- **H.323:** Original standard approved by the ITU-T
- **SGCP:** Developed by Telcordia (Bellcore) and Cisco to improve on H.323
- **IPDC:** Developed by Level 3 developers to improve on H.323
- **MGCP:** A merger of SGCP and IPDC sponsored IETF
- **MDCP:** Developed by Lucent with the goal of creating a generic, low overhead protocol for the control of media processors

- **MGCP+:** An IETF draft, currently in Version 2, designed to build on the features of MGCP and MDCP

Current Cisco AVVID products run Skinny Station Protocol (Station and Gateway), H.323 (CCM) and MGCP (gateways). Cisco IOS gateways support H.323 and MGCP. Cisco introduced support for several SIP features in Cisco IOS 12.2(2)XB2.

Reference	For more information, go to: http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122newft/122limit/122x/122xb/122xb_2/



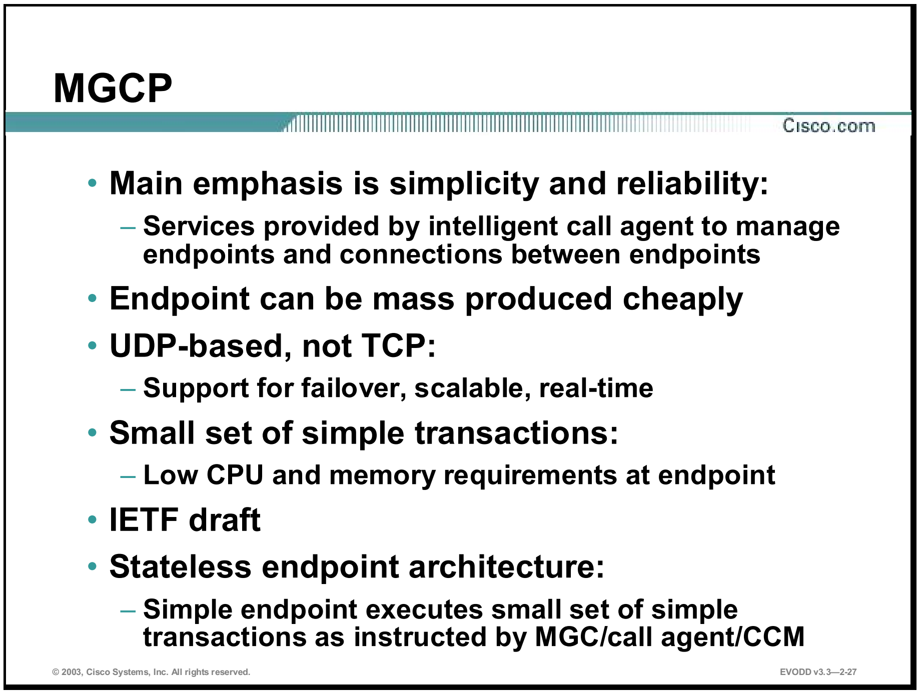
SGCP is a control protocol that has two functional components: a call agent and an endpoint. The call agent has the intelligence and controls the actions of the endpoint. All the endpoint does is notify the call agent of events. SGCP is independent of the protocol that actually carries the voice traffic, such as RTP or VoFR.

As shown in the illustration, the Cisco AVVID implementation uses Skinny Station Protocol between the IP Phone and the call agent (CCM), and uses H.323 between CCM and other CCMs or gateways. In some cases, when communicating with Skinny Station Protocol, SGCP is utilized to control the gateways.

In the SGCP model, the gateways focus on the audio signal translation function, while the call agent handles the signaling functions. As a consequence, the call agent implements the signaling layers of the H.323 standard and presents itself as a gatekeeper to the H.323 systems. Calls are established using the gatekeeper routed-call model.

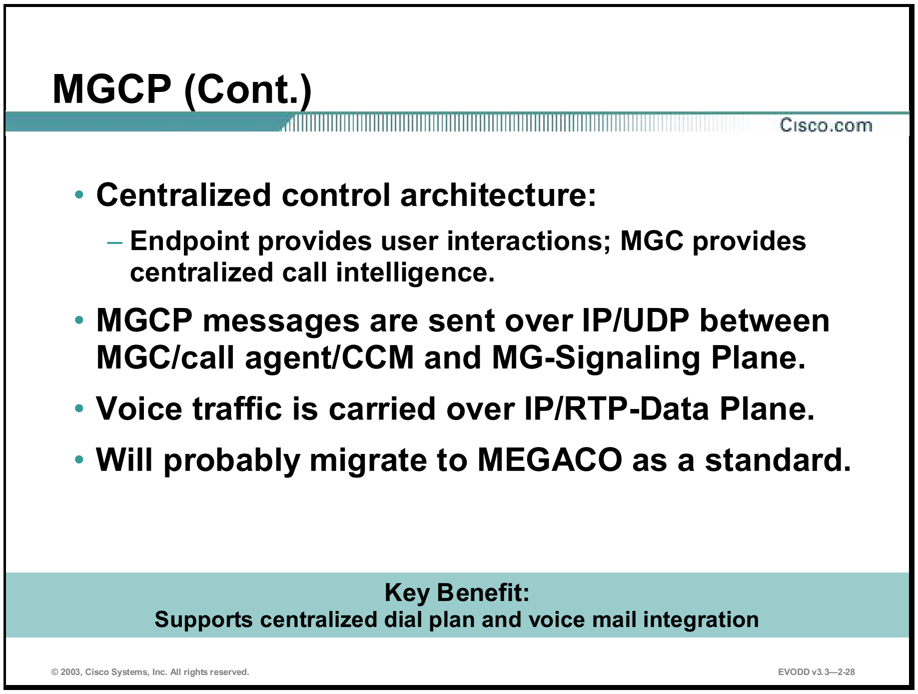
Media Gateway Control Protocol

This topic describes the development and use of MGCP. MGCP is the result of merging SGCP and IPDC.



SGCP was created to provide a simple, reliable protocol based on the client/server model. SGCP relies on an intelligent call agent with unintelligent endpoints. As a result, endpoints can be produced cheaply. A cable modem is an example of a device that would use SGCP. A smaller set of instructions at the client end means lower CPU and memory requirements, which reduce the cost of production. Because SGCP is based on UDP (not TCP) it supports failover, is scalable, and is specifically designed for real-time applications, such as voice.

SGCP has effectively been overtaken by MGCP, which is in draft IETF status. SGCP is unlikely to go much further because the Media Gateway Control (MEGACO), and the MGCP provide MGCP functionality. Pundits expect MEGACO to evolve as the client and/or server protocol of choice, and H.323 and SIP to provide peer-to-peer functionality.



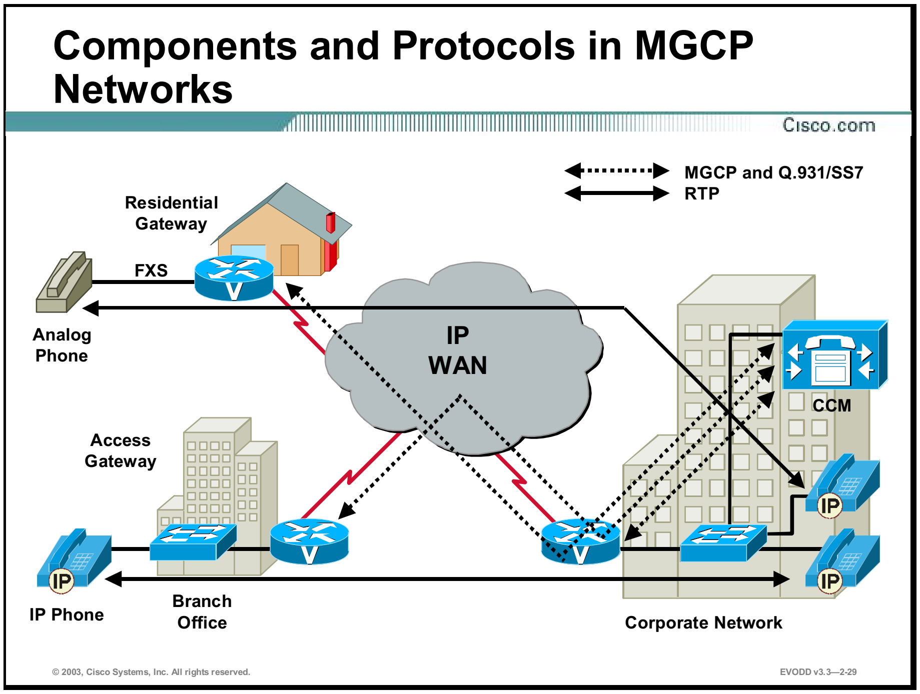
MGCP has a stateless endpoint architecture, similar to Skinny Station Protocol. The endpoint has no real call routing intelligence. The endpoint only executes simple commands as instructed by the Media Gateway Controller (MGC).

MGC is synonymous with call agent, implemented as CCM in the AVVID environment. All true routing and decision-making capabilities are housed within the MGC, call agent, and CCM.

The creation of a centralized dial plan is a benefit of MGCP architecture. A centralized dial plan means that separate dial plans are not residing on separate gateways. For example, you can use this feature for voice integration with an Octel system because the port to which voice mail is sent is known.

Like H.323, in MGCP there is a signaling plane and a data plane. The signaling plane for MGCP is over an IP/UDP channel; the bearer channel is IP/RTP as with H.323.

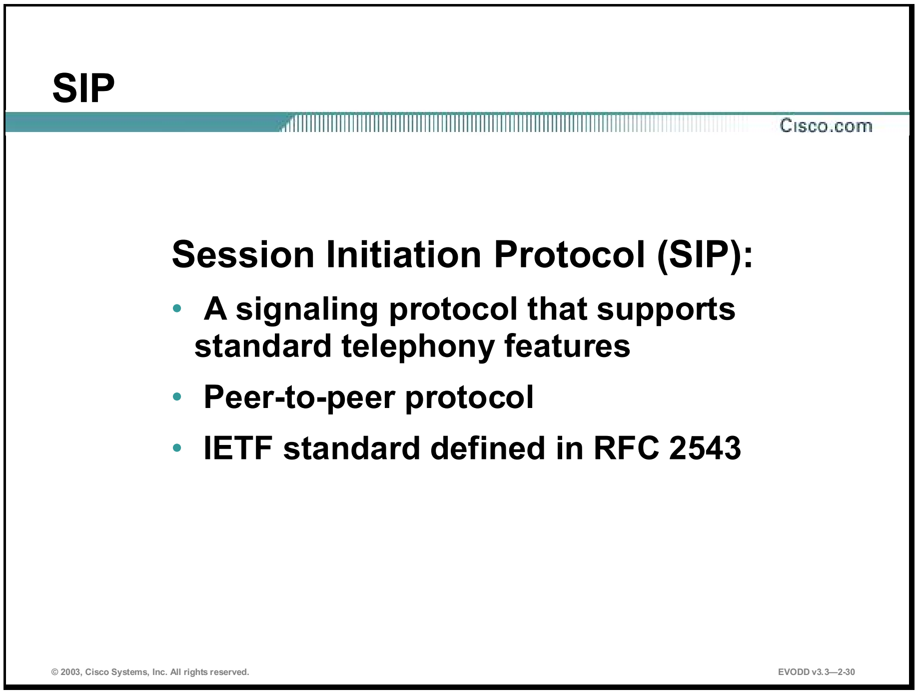
Note MGCP is not an approved standard and implementations are not consistent even within Cisco. For example, the 5300 running MGCP is not compatible with CCM. MGCP will probably mutate or migrate to MEGACO, which will be a standard, with call flow logic similar to MGCP.



The diagram demonstrates the role played by MGCP in a VoIP network. MGCP is configured on Cisco IOS devices, such as routers and gateways. The MGCP configuration instructs these Cisco IOS devices to reference the call agent (CCM) for address resolution. The Q.931 protocol is used for call setup. Once the call is set up, voice traffic is streamed via the RTP protocol.

Session Initiation Protocol

This topic covers the details of SIP, including its interoperability with other signaling protocols. SIP is a signaling protocol for Internet conferencing and telephony.

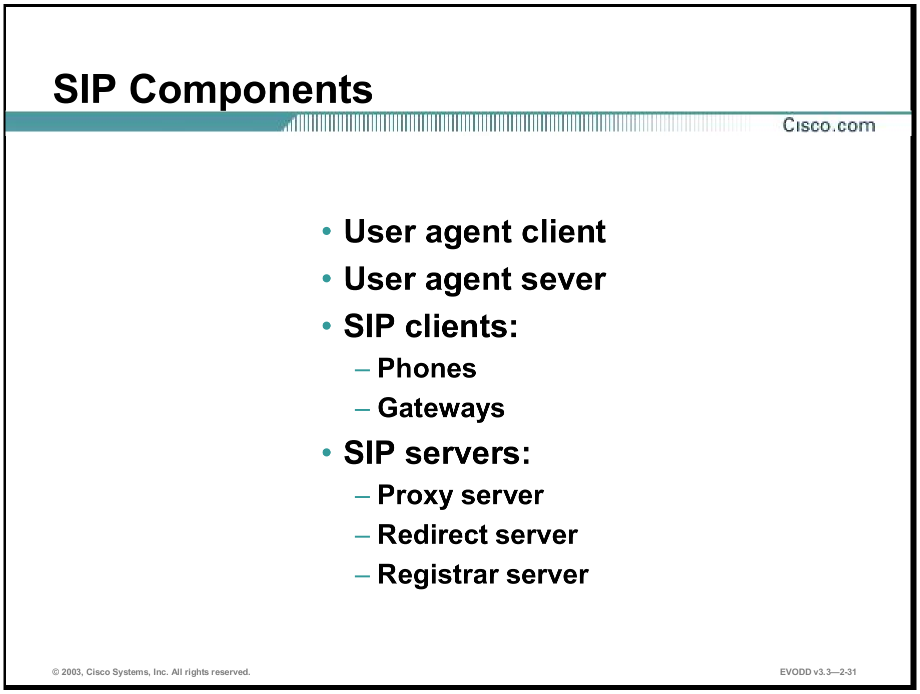


SIP supports standard telephony features, such as: call forwarding (unconditional, busy, etc.), call transfer (call control spec), caller ID, call hold, three-way and multiparty conferencing (call control spec), call return (*69), call park (with NOTIFY), follow-me, find-me, interactive voice response (IVR) systems, multiple line presences, call waiting, camp-on, call queuing, automatic call distribution, and do not disturb. Some services, such as repetitive dialing, station speed dialing, last number redial, and distinctive ringing, are implemented purely in the end system and require no support from the signaling protocol.

The details of combining SIP and MGCP in a system are still being developed. MGCP is a device control protocol in which a master (Media Gateway Controller Call Agent) controls a slave (Media Gateway). SIP may be used between controllers in a peer-to-peer relationship. SIP views the MGC as a node with a large number of connections, but otherwise it is the same as a native SIP device. Similarly, the Media Gateway is completely unaware that the call between MGCs is established via SIP. Only the MGC needs to understand both protocols.

Recent SIP enhancements include features that support communications for deaf, hard-of-hearing and speech-impaired individuals. These requirements address the current difficulties of deaf, hard-of-hearing, and speech-impaired individuals when using communications facilities.

Reference Additional information about SIP can be found at this URL: http://www.cs.columbia.edu/sip/.



SIP is a peer-to-peer protocol. The peers in a session are called user agents (UAs). A user agent can function in one of the following roles:

- **User agent client (UAC):** A client application that initiates the SIP request.

- **User agent server (UAS):** A server application that contacts the user when a SIP request is received and returns a response on behalf of the user.

Typically, a SIP endpoint is capable of functioning as both a UAC and a UAS, but it functions only as one or the other per transaction. Whether the endpoint functions as a UAC or a UAS depends on the UA that initiated the request.

From an architecture standpoint, the physical components of an SIP network can be grouped into two categories: clients and servers.

SIP clients include:

- **Telephones:** Telephones can act as either a UAS or UAC. Softphones (PCs that have telephone capabilities installed) and Cisco SIP IP Phones can initiate and respond to SIP requests.

- **Gateways:** Gateways provide call control. Gateways provide many services, the most common being a translation function between SIP conferencing endpoints and other terminal types. This function includes translation between transmission formats and between communications procedures. In addition, the gateway translates between audio and video codecs and performs call setup and clearing on both the LAN side and the SCN side.

SIP servers include:

- **Proxy server:** The proxy server is an intermediate device that receives SIP requests from a client and then forwards the requests on the behalf of the client. Basically, proxy servers receive SIP messages and forward them to the next SIP server in the network. Proxy servers provide functions, such as authentication, authorization, network access control, routing, reliable request retransmission, and security.

- **Redirect server:** The redirect server provides the client with information about the next hop or hops that a message should take, and then the client contacts the next hop server or UAS directly.

- **Registrar server:** The registrar server processes requests from UACs for registration of their current location. Registrar servers are often located with a redirect or proxy server.

SIP vs. H.323

Cisco.com

Aspect	SIP	H.323
Clients	Intelligent	Intelligent
Network intelligence and services	Provided by servers (Proxy, Redirect, Registrar)	Provided by gatekeepers
Model used	Internet/WWW	Telephony/QSIG
Signaling protocol	UDP or TCP	TCP (UDP is optional in Version 3)
Media protocol	RTP	RTP
Code basis	ASCII	Binary (ASN.1 encoding)
Other protocols used	IETF/IP protocols, such as SDP, HTTP/1.1, IPmc, and MIME	ITU / ISDN protocols, such as H.225, H.245, and H.450
Vendor interoperability	Wide-spread	Limited

 EVODD v3.3—2-32

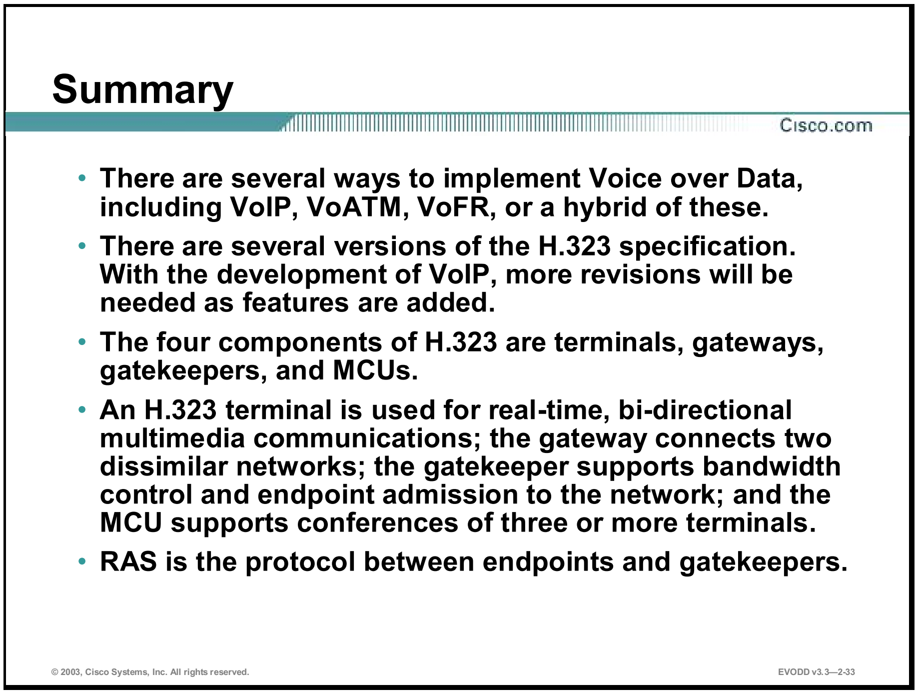
Both SIP and H.323 are peer-to-peer protocols. This slide illustrates the similarities and the differences between the protocols.

SIP and H.323 were designed to address session control and signaling functions in distributed call control architectures. Although SIP and H.323 can also be used to communicate to limited intelligence endpoints, they are especially well suited for communication with intelligent endpoints.

Although SIP messages are not directly compatible with H.323, both protocols can coexist in the same packet telephony network if a device that supports the interoperability is available. For example, a call agent could use H.323 to communicate with gateways and use SIP for inter-call agent signaling. After the bearer connection is set up, the bearer information flows between the different gateways as an RTP stream.

Summary

This topic summarizes the key points discussed in this lesson.



Summary (Cont.)

Cisco.com

- **H.225 call signaling is used to set up connections between H.323 endpoints and H.245 control signaling consists of the exchange of end-to-end H.245 messages between communicating H.323 endpoints.**
- **CCM and Cisco IOS gateways support three gateway protocols: SIP, MGCP, and H.323.**
- **MGCP is a control protocol that provides centralized call control. The endpoint provides the user interaction, and the MGC provides the centralized call intelligence.**
- **SIP is a signaling protocol for Internet conferencing and telephony.**

EVODD v3.3—2-34

Next Steps

After completing this lesson, go to:

- Voice over Frame Relay Design Considerations lesson

References

For additional information, refer to these resources:

- H.323 Standards:

  http://www.packetizer.com/iptel/h323/

- Cisco IOS SIP Configuration:

  http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122cgcr/fvvfax_c/vvfsip.htm

- MDCP:

  http://www.ietf.org/proceedings/98dec/slides/megaco-mdcp-98dec.pdf

- MGCP+:

  http://www.ietf.org/internet-drafts/draft-ietf-megaco-h248v2-01.txt

- SIP:

  http://www.cs.columbia.edu/sip/

Lesson Review

This practice exercise reviews what you have learned in this lesson.

Q1) Which two of the following signaling protocols are considered peer-to-peer protocols? (Choose two.)

A) H.323

B) Simple Client Control Protocol

C) SIP

D) MGCP

Q2) What is the current version of H.323?

A) 2

B) 3

C) 4

D) 5

Q3) In the context of the H.323 standard, which of the following components is considered mandatory?

A) terminals

B) gateways

C) gatekeepers

D) MCUs

Q4) Which codec must be supported by H.323 terminals?

A) G.729

B) G.711

C) G.729 Annex A

D) G.723

Q5) Which of the following statements best describes registration, admission, and status (RAS)?

A) RAS is a protocol that runs between endpoints (terminals and gateways) and gatekeepers.

B) RAS is a protocol that runs between Cisco IP Phones and other IP Phones.

C) RAS is a protocol that runs between terminals and gateways.

D) RAS is a protocol that runs between Cisco IP Phones and CCMs.

Q6) Which of the following protocols is used to set up connections between H.323 endpoints, over which real-time data can be transported?

A) RTCP

B) RTP

C) H.245

D) H.225

Q7) Which protocol does Cisco AVVID use between Cisco IP Phones and CCMs?

A) H.245

B) Skinny Station Protocol

C) RTCP

D) H.225

Q8) Which of the following statements are true concerning MGCP?

A) MGCP uses a large set of complex transactions for maximum vendor interoperability.

B) MGCP uses TCP instead of UDP because TCP is a reliable transport protocol.

C) Endpoints can be mass-produced cheaply.

D) MGCP was developed by Lucent.

Q9) Which three of the following components belong to SIP? (Choose three.)

A) user agent client

B) user server client

C) SIP client

D) SIP server

Voice over Frame Relay Design Considerations

Overview

Voice over Frame Relay (VoFR) technology consolidates voice and voice-band data with data services over the Frame Relay network. VoFR provides end users with greater efficiencies in bandwidth use and other network resources. This lesson will teach you about the benefits of VoFR and the common voice applications used over Frame Relay. You will learn about VoFR design considerations, specifically the various encapsulation standards and common challenges encountered when implementing VoFR.

Importance

This lesson benefits those students who want to increase their understanding of VoFR design. This lesson provides information on Frame Relay enhancements, which aid in preserving the quality of streaming media, such as voice and video.

Objectives

Upon completing this lesson, you will be able to:

- Describe the benefits of VoFR
- Identify common voice applications used over Frame Relay
- Describe VoFR encapsulation methods
- Describe the FRF.11 implementation agreement
- Identify Cisco VoFR connection types
- Describe design strategies for integrated multiservice network architecture
- Calculate VoFR bandwidth

Learner Skills and Knowledge

To fully benefit from this lesson, you must have these prerequisite skills and knowledge:

- *Interconnecting Cisco Network Devices* (ICND) course
- *Designing Cisco Networks* (DCN) course

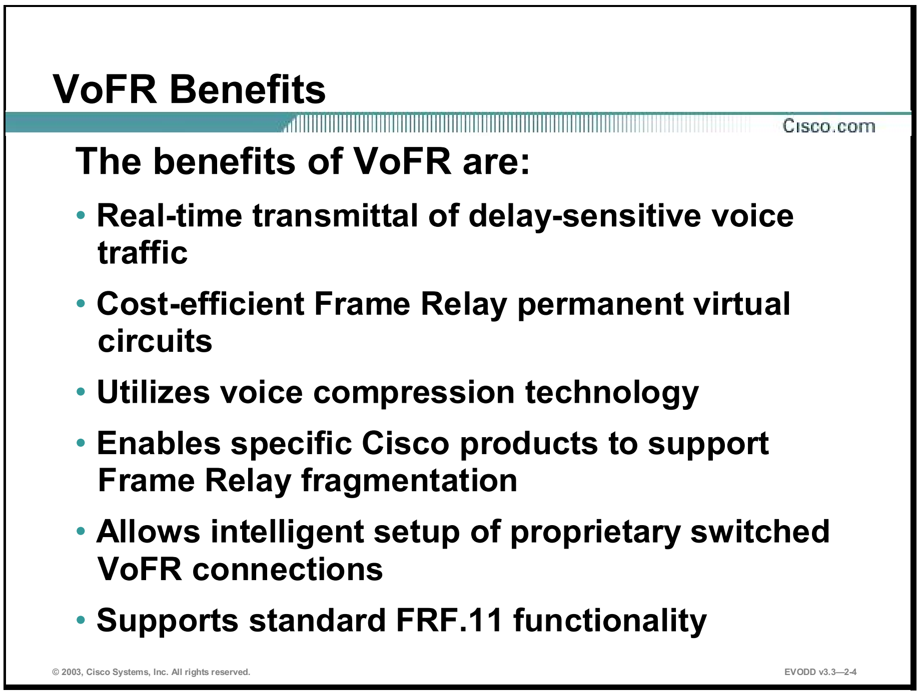
Outline

This lesson includes these topics:

- Overview
- VoFR Benefits
- Common Applications
- Encapsulation Methods
- FRF.11
- Cisco VoFR Connection Types
- Minimizing Delay and Loss
- Bandwidth Calculations
- Summary
- Laboratory Exercise: Voice over Frame Relay
- Lesson Review

VoFR Benefits

This topic describes the benefits of Voice over Frame Relay (VoFR). VoFR has the potential to provide end users with greater efficiencies in accessing bandwidth by functionally integrating voice, data, and fax over a single access link. In addition, VoFR provides end users with a cost-effective option for transporting voice traffic between company locations.

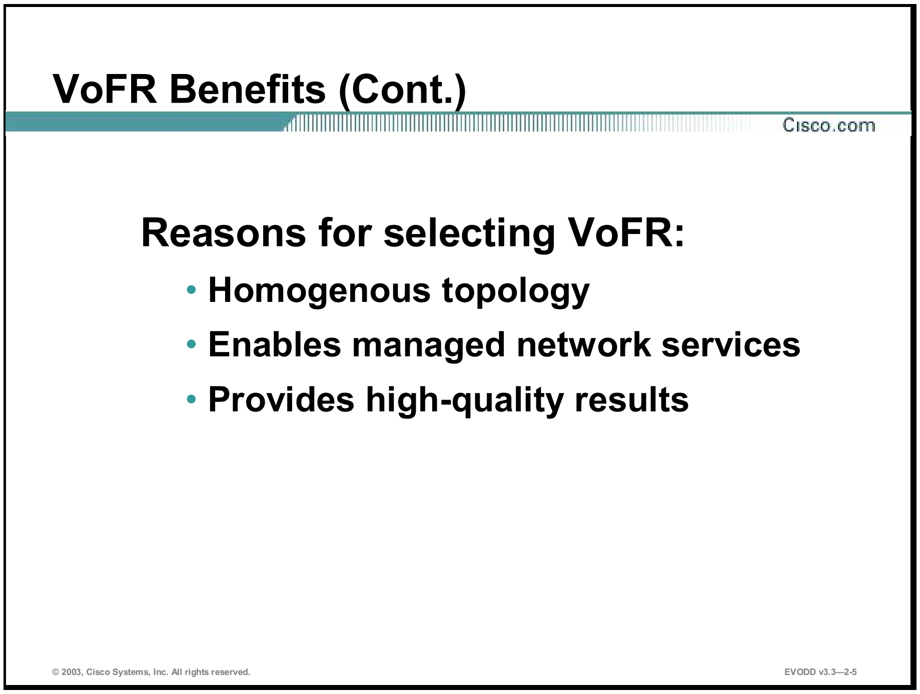


The implementation of VoFR technology provides the following benefits to an existing Frame Relay network:

- Enables real-time, delay-sensitive voice traffic to be carried over slow Frame Relay links
- Allows dedicated 64-kbps time-division multiplexing (TDM) telephony circuits to be replaced by more economical Frame Relay permanent, virtual circuits
- Allows voice-enabled routers from multiple remote sites to be multiplexed into a central site router, through Frame Relay links
- Utilizes voice compression technology that conforms to International Telecommunication Union-Telecommunication Standardization Sector (ITU-T) specifications
- Enables Cisco 2600 and 3600 series routers and the Cisco MC3810 multiservice access concentrator to support Frame Relay fragmentation

- Allows intelligent setup of proprietary switched VoFR connections between two VoFR endpoints, saving the extensive configuration overhead associated with the Frame Relay Forum implementation agreement for VoFR (FRF.11) implementations

- Supports standard FRF.11 functionality, allowing Cisco routers to interconnect with other equipment supporting this specification



Besides the numerous benefits that VoFR provides to an existing Frame Relay network, the main reasons for selecting Frame Relay as a Voice over Data transport include the following:

- Integrates voice and data on a homogenous topology
- Enables managed network services
- Provides high-quality results

VoFR offers a solution for integrating voice and data for customers with a homogenous Frame Relay network. Frame Relay overhead at Layer 2 is smaller than the overhead of Layer 3, which provides greater efficiency on the network.

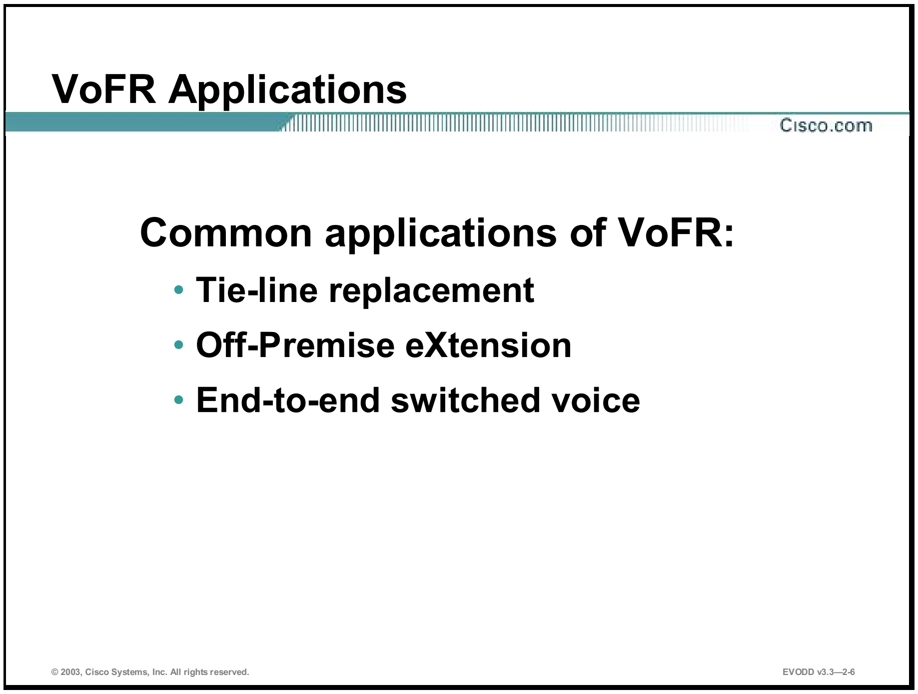
Many service providers, who already offer Frame Relay services, are looking for service differentiators. VoFR can be that differentiator. Companies such as Equant already offer services based on Cisco VoFR products. Some of these services over managed networks include videoconferencing, electronic white boards, and voice calls placed from World Wide Web pages.

Other telecommunication and Internet-related services are voice and fax over IP, domestic and international long distance, digital subscriber line (DSL), web portal and hosting, interactive voice response (IVR) applications, fax broadcast services, application service provider (ASP) and Internet service provider (ISP) services.

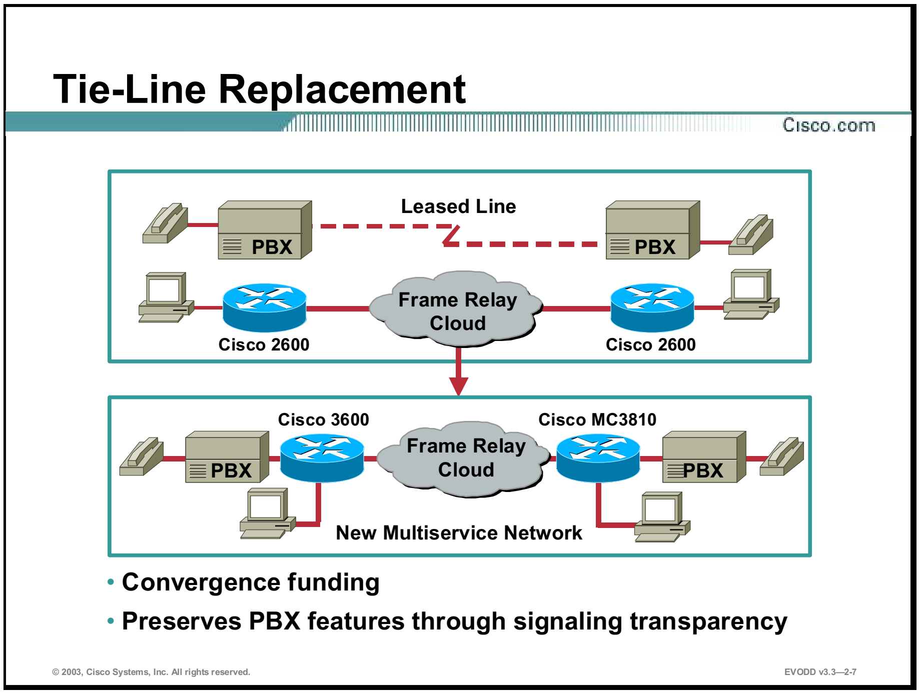
Cisco VoFR products are designed to provide the best in voice quality and offer a variety of compression techniques, which enable customers to adjust for optimal bandwidth and quality. Advanced Cisco IOS router-based queuing techniques ensure that the voice/fax packets are given the highest priority.

Common Applications

This topic examines common voice applications that are used over Frame Relay.



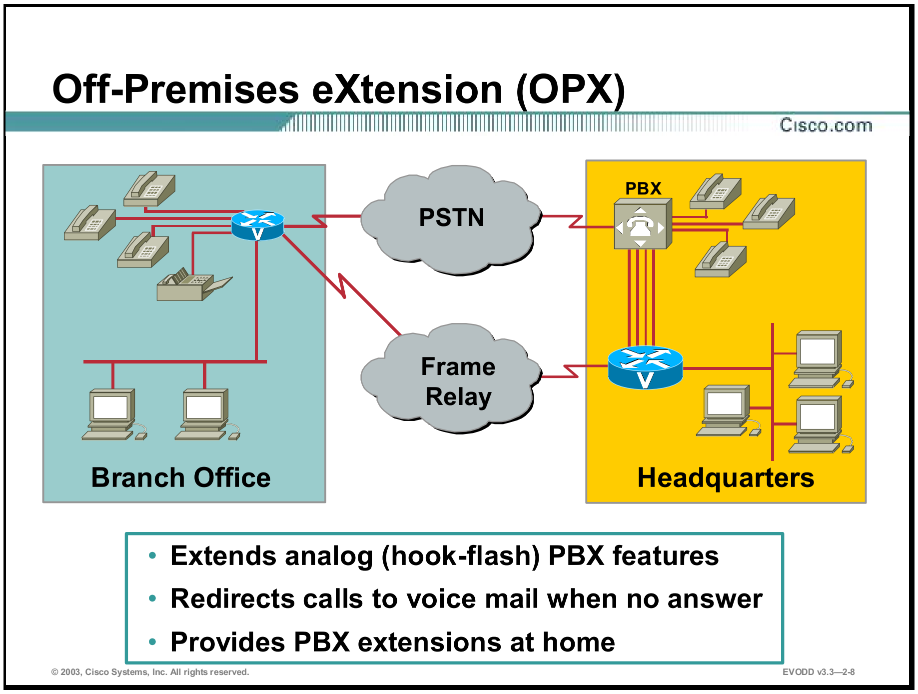
VoFR allows you to implement the applications shown here. By moving these services to VoFR, you can deliver them while reducing cost.



Tie-line replacement is a typical application for FRF.11 trunks. For example, a Frame Relay permanent virtual circuit (PVC) can replace an expensive dedicated 64 kbps circuit.

This type of VoFR application provides the following benefits:

- Convergence funding
- Preserves PBX features through signaling transparency



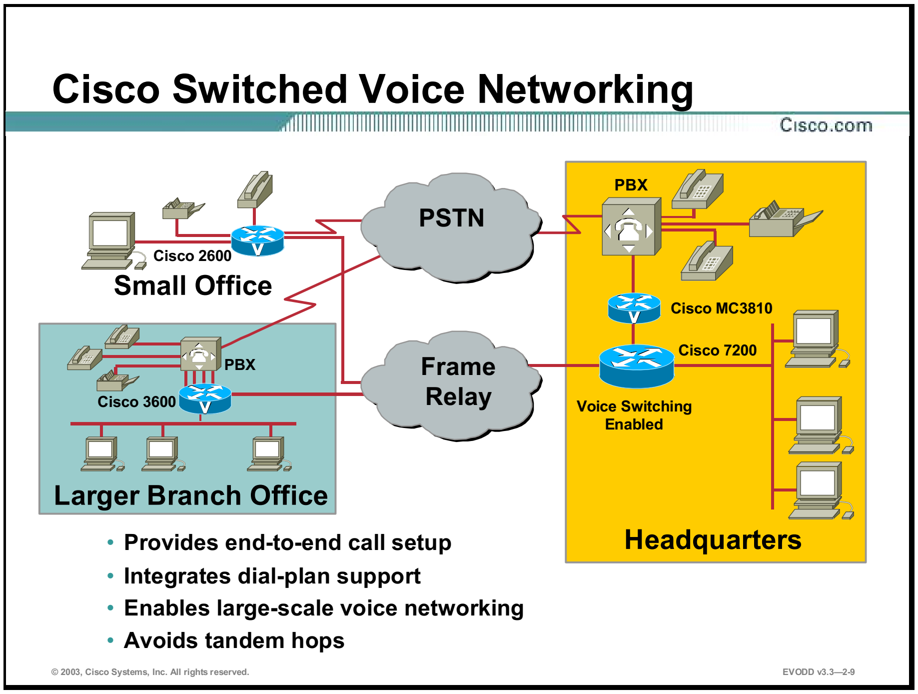
Off-Premises eXtension (OPX) is an application used with FRF.11 trunks. In networks where a trunk connection is used to support an OPX, you will have the following results:

- A remote telephone acts as if it is connected directly to the PBX and the connection is always active.
- Every remote telephone requires one channel at the PBX (no over subscription of ports).
- Transparent signaling allows any signaling protocol, although the signaling protocol would be Foreign Exchange Office (FXO) to Foreign Exchange Station (FXS).
- Carries hook-switch flashes, mid-call signaling, and dual-tone multifrequency (DTMF) tones without distortion.
- Supports preemptive busy-back if far-end is Out-of-Service (OOS), or network connection fails.

In networks where VoFR is utilized to support an OPX, you will have the following results:

- The remote telephone acts as if it is connected directly to the PBX and the connection is active only when either end is off-hook.
- Hunt groups may be used to oversubscribe remote ports to PBX ports.
- Signaling supported by the router must be used.

- Carries hook-switch flashes, mid-call signaling, and DTMF tones without distortion.
- The network answers calls presented by the PBX, so that they cannot call forward or hunt to voice mail (unless using Cisco MC3810 with connection private line, automatic ringdown-Off-Premises eXtension [plar-opx], a feature known as answer supervision).
- Does not support preemptive busy-back if far-end is OOS, or network connection fails.



The application of VoFR technology as a Cisco switched voice networking solution supports corporate dialing plans and provides the following benefits:

- Provides an end-to-end call setup
- Integrates a dial-plan support
- Enables large-scale voice networking
- Avoids tandem hops

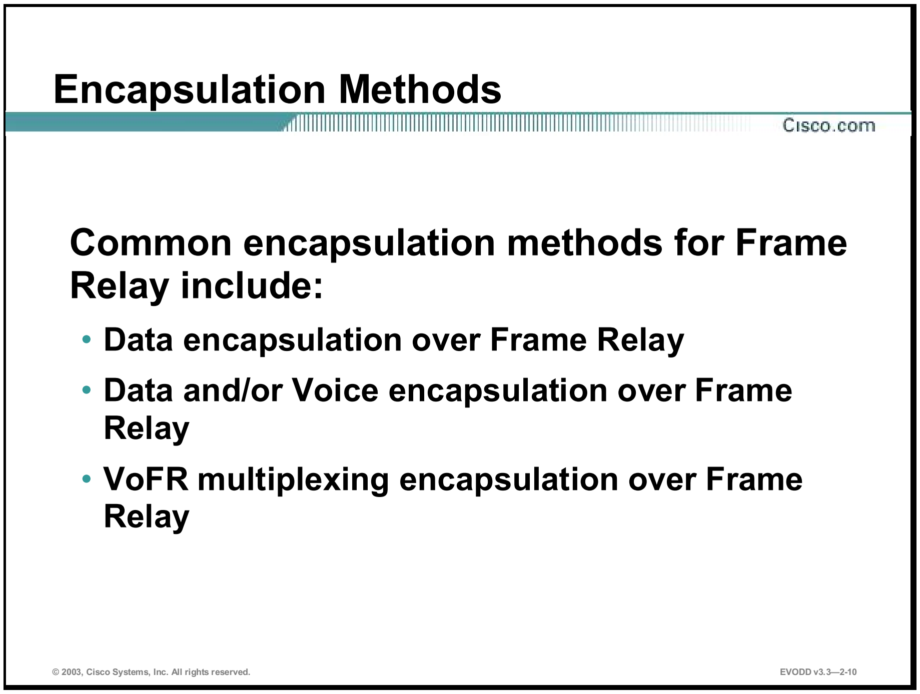
A telephone receives a dial tone from the adjacent router. The router then receives the digits dialed by the telephone, and sets up a voice switched virtual circuit (SVC) connection with the destination.

Cisco routers can accept extra digits, if required, and play them out at the far-end for Direct Inward Dialing (DID) or voice-mail access.

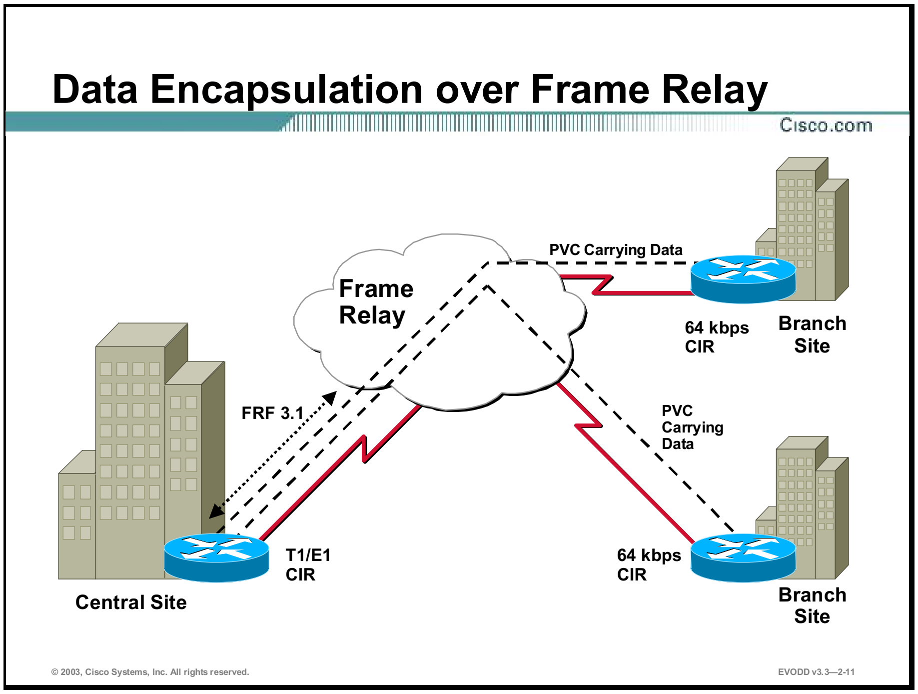
Once the connection is established, DTMF is detected and regenerated at the far-end according to FRF.11.

Encapsulation Methods

This topic describes VoFR encapsulation methods.



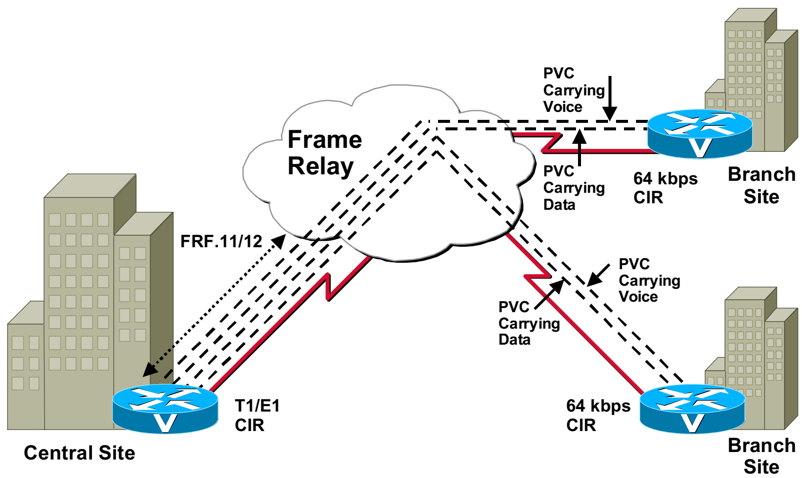
Encapsulation methods affect network performance and decisions regarding network design.



The illustration shows a typical data over Frame Relay network without voice connections. Data is transported with the FRF 3.1 data encapsulation standard.

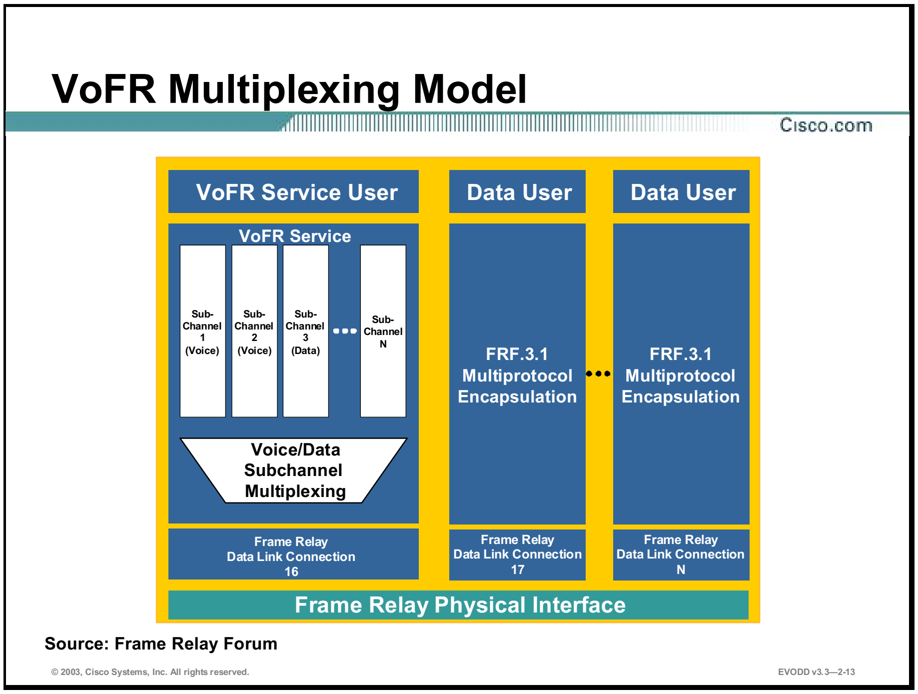
Data/Voice Encapsulation over Frame Relay

Cisco.com



 EVODD v3.3—2-12

The illustration shows a data over Frame Relay network modified to carry both voice and data. Voice and data routers have been added, as well as new PVCs that are designed to accommodate the aggregate voice and data bandwidth requirements.



The illustration shows a Frame Relay multiplexing model. The two columns on the right demonstrate a typical data over Frame Relay network, where data is transported via the FRF 3.1 encapsulation.

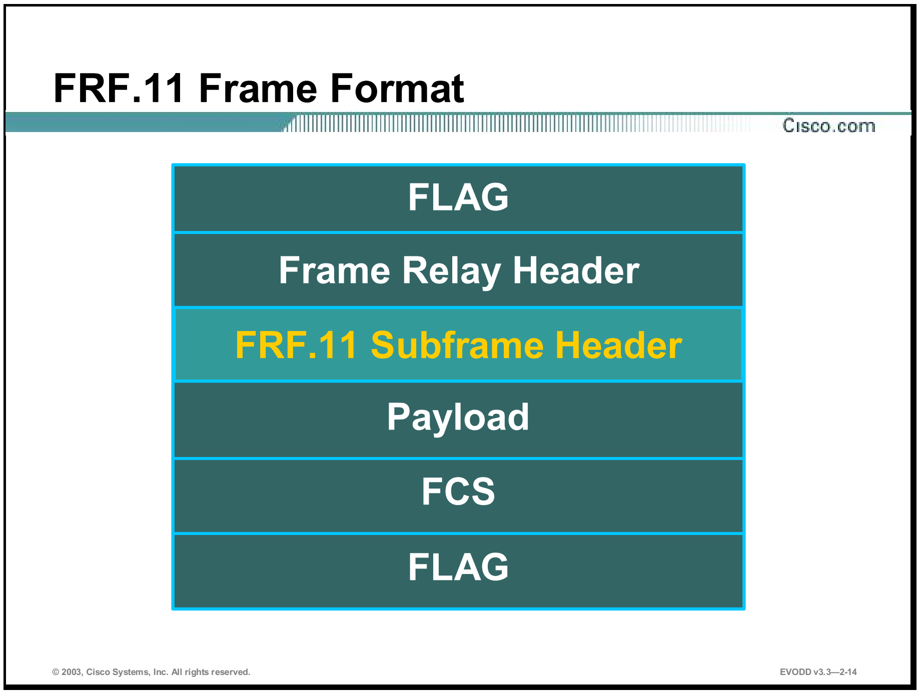
Note Voice encapsulation requires additional VoFR services.

New syntax was added to the Frame Relay reference model, which supports new VoFR services. The VoFR service column represents the FRF.11 syntax and the capabilities for multiple subchannels within FRF.11. Both FRF.11 and FRF 3.1 encapsulations may be present on the same physical link.

The only application where an FRF.11 trunk is preferred to a Cisco trunk is when the Cisco router is connected to another vendor’s FRF.11/12-compliant device. Using FRF.11/12 limits capabilities (point-to-point—no switched connections, negotiation, or rerouting). The Cisco device and the foreign Frame Relay Access Device (FRAD) typically share only a subset of options.

FRF.11

This topic describes the FRF.11 implementation agreement. The FRF.11 frame format allows for subchannels, which provide a better transport mechanism for voice.



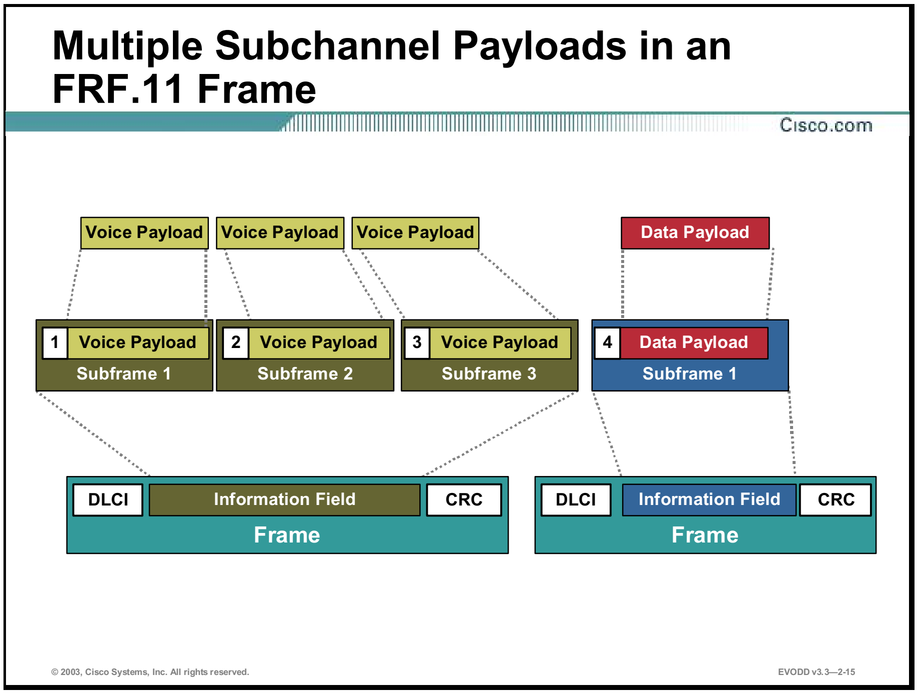
The subframe header is added to the Frame Relay frame for FRF.11 services. The subframe header is from one to three octets in length, and includes fields for the subchannel ID (CID), the payload type, and the payload length.

FRF.11 allows for up to 255 subchannels on a data-link connection identifier (DLCI) to be configured for voice and one subchannel of this DLCI to be used for data protocols. Subchannels 0 to 3 are reserved.

Subchannel 4 is the default data subchannel in Cisco FRF.11 implementation (this subchannel is configurable). Cisco switched VoFR protocol for inter-node communication also uses subchannel 5.

An FRF.11 trunk is static. The FRF.11 trunk is established at router bootup, or when configured. The trunk remains up until the voice port or serial port is shut down, or until a network disruption occurs. In an FRF.11 trunk, signaling frames are transparently transported through the network. On the Cisco 2600 and 3600, this signaling terminates on the router, which means that the router interprets the signaling frames that it receives and relays the on-hook/off-hook information to the hardware.

The Cisco 2600 and 3600 can currently interpret EIA-464 (North America) signaling. However, they cannot detect on-hook/off-hook conditions sent from a device generating signaling frames of a different type.



In the illustration, a single DLCI supports three voice channels and one data channel. The first frame contains three subframes, with three voice payloads. The second frame contains a data payload.

Voice and data payloads are multiplexed using FRF.11 subchannels over a VoFR connection.

Each payload is packaged as a subframe within the information field of a frame. Subframes may be combined within a single frame, to improve processing and transport efficiencies.

Each subframe contains a header and payload. The subframe header identifies the voice/data subchannel, and when required, the payload type and length.

FRF.11 Concept

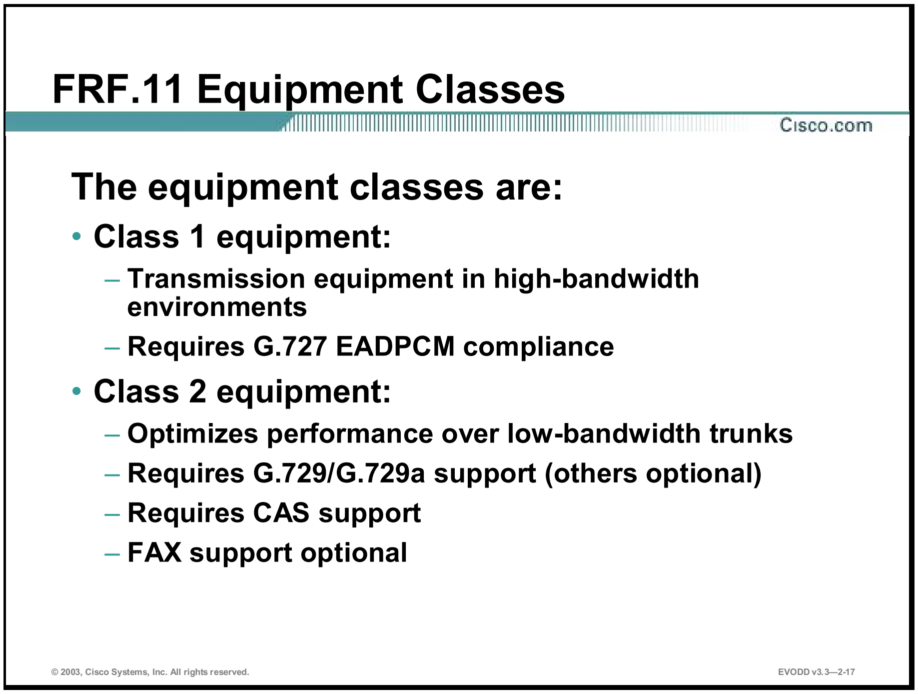
Cisco.com

The FRF.11 concept includes:

- Extension of Frame Relay application support for compressed voice
- Multiplexing of up to 255 subchannels
- Support of multiple payloads
- Support of data subchannel

EVODD v3.3—2-16

FRF.11 adds new syntax to the Frame Relay reference model of voice frames. The FRF.11 syntax may contain voice, data, or other payloads. The slide lists FRF.11 benefits.



A VoFR implementation is compliant with the FRF.11 agreement if the requirements for at least one of the following two classes of FRF.11 equipment are met. Class 2 compliant devices support capabilities that enable optimal performance over low bit-rate Frame Relay interfaces.

Class 1 compliant devices support capabilities suitable for high bit-rate interfaces. These include:

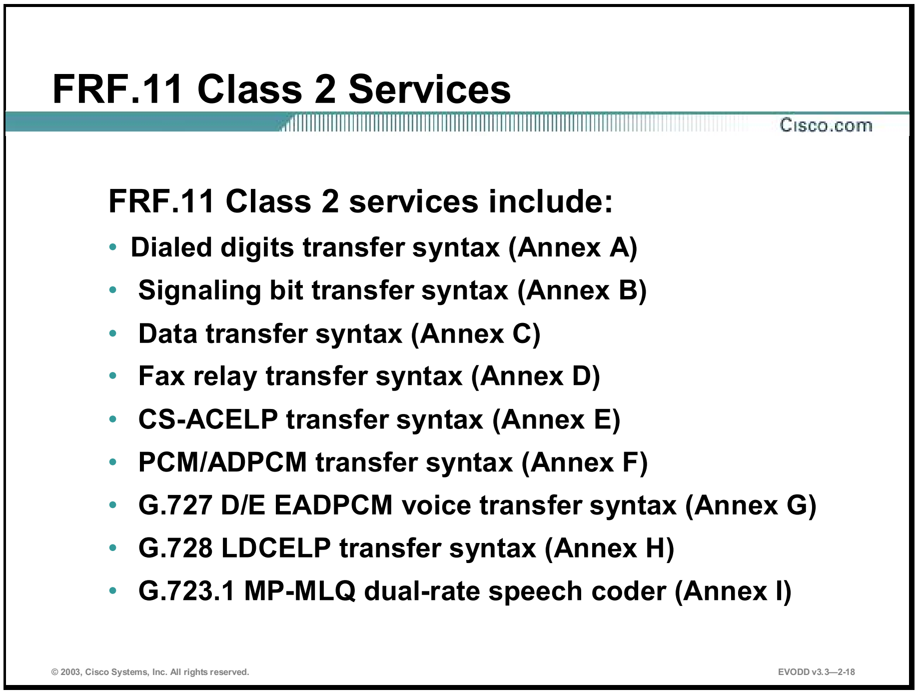
- **Primary payload types:** Support of G.727, as described in Annex F, is mandatory. Support of other voice coders (vocoders) described in Annex F is optional. A transmit rate of 32 kbps, with support for rates of 32 kbps, 24 kbps, and 16 kbps are mandatory at the receiver. Support for other primary payload transfer syntax definitions (for example, fax) are optional.

- **Signaled payload types:** Support for the dialed digit signaled payload type is optional. Support for the signaling bits signaled payload type channel associated signaling (CAS) and alarm indication signal (AIS) is mandatory. Support for the encoded fax signaled payload type is optional.

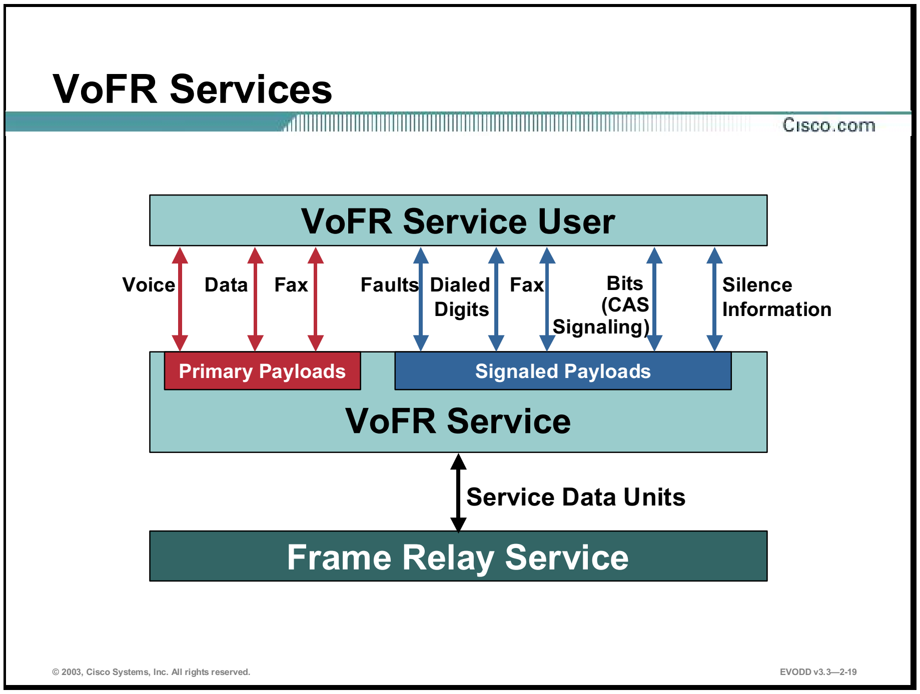
Class 2 compliant devices support capabilities that enable optimal performance over low bit-rate Frame Relay interfaces. These include

- **Primary payload types:** Support for Annex E Conjugate Structure Algebraic Code Excited Linear Prediction (CS-ACELP) G.729 or G.729A voice transfer syntax is mandatory. Support for other primary payload transfer syntax definitions (for example, fax) is optional.

- **Signaled payload types:** Support for the dialed digit signaled payload type and the signaling bits signaled payload type (CAS and AIS) is mandatory. Support for the encoded fax signaled payload type is optional.



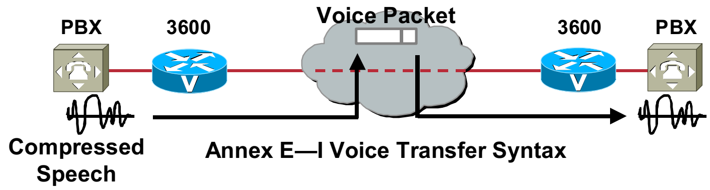
FRF.11 includes the Class 2 services shown. Primary payload requires Annex E CS-ACELP G7.29 or G7.29A voice transfer syntax. Signaled payload types require support for the signal bits and dialed digits transfer syntax.



The relationship between the VoFR service user, the VoFR service, and the Frame Relay service is shown in this illustration.

The two payload types are primary payload and signaled payload. A primary payload is used to transport the encoded bits from voice, fax, or data. Signaled payload is used to transport dialed digits, silence information, and other signaling information.

Voice Payload Transfer: FRF.11 Annex E, F, G, H, I



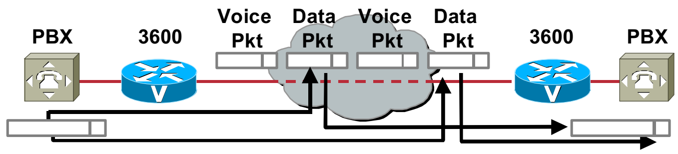
- **CODEC syntax specified per annex:**
  - **Annex E—CS-ACELP G.729**
  - **Annex F—Generic PCM/ADPCM G.711, G.726, G.727**
  - **Annex G—X.25 over Frame Relay**
  - **Annex H—LDCELP G.728**
  - **Annex I—G.723.1**
- **Sequence number and coding type are optional**

 EVODD v3.3—2-20

FRF.11 annexes *E, F, G, H, and I* specify the FRF.11 syntax for voice payload transfers. For example, Annex E is specific for G.729. In each specific annex, the codec from the voice sample generates the layout for the bytes. The layout for the bytes is mapped to the byte position in the payload.

Data Payload Transfer: FRF.11 Annex C

Cisco.com



Annex C—Data Transfer Syntax

Annex C includes:

- **FRF.11 subframe headers allow voice and data subchannels within a PVC**
- **Each packet contains a whole data frame or a fragment of an original data frame**
- **Original frames smaller than the fragmentation threshold are encapsulated with both the B and E bits set**

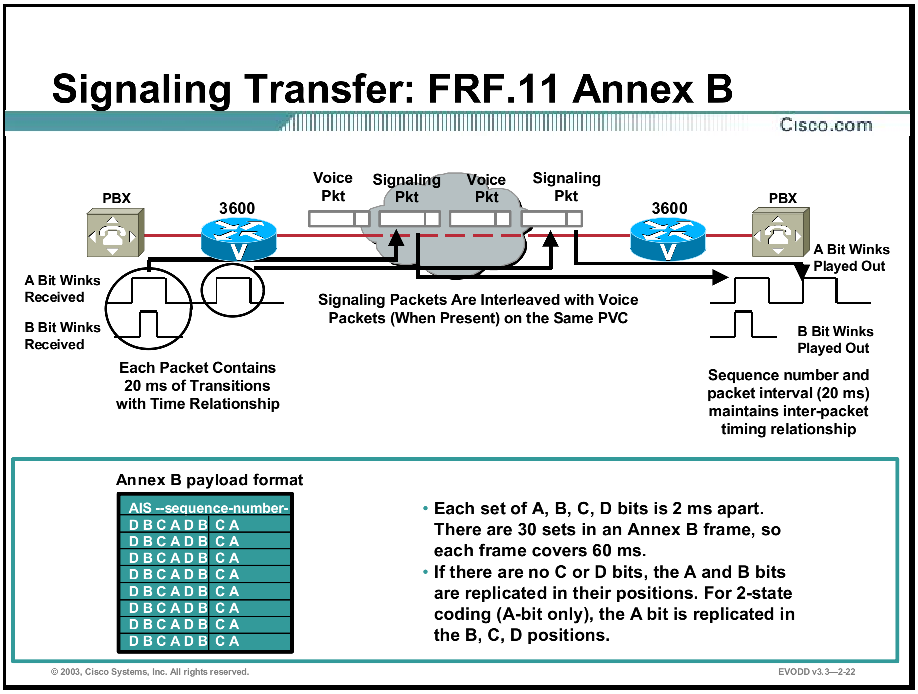
 EVODD v3.3—2-21

FRF.11 Annex C specifies the syntax for data transfer.

The subchannel payload type can specify either voice or data; therefore, a single DLCI can transport a mixed payload of voice and data subchannels. Cisco does not support voice and data in a subchannel, and FRF.11 is used only for voice.

The supported data packet is fragmented, based on the FRF.11 fragmentation threshold. The first fragment has the B bit set, the last fragment has the E bit set, and fragments smaller than the fragmentation threshold have both the B and E bits sets.

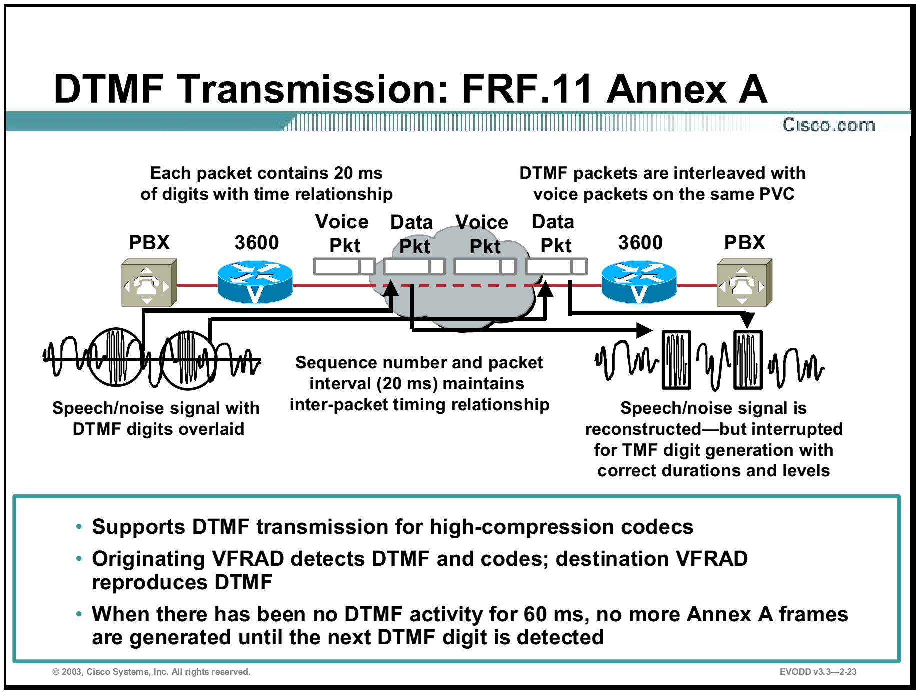
A sequence number is added to each subframe for correct reassembly at the remote end.



FRF.11 Annex B specifies the syntax for signaling transfer.

If there has not been a signaling bit change for 500 ms, the packet reverts to a keepalive packet every five seconds. The sequence number does not change. Cisco bounces signaling using a 10 ms timer.

Analog terminations have signaling transitions converted to A, B, C, and D bit equivalents.



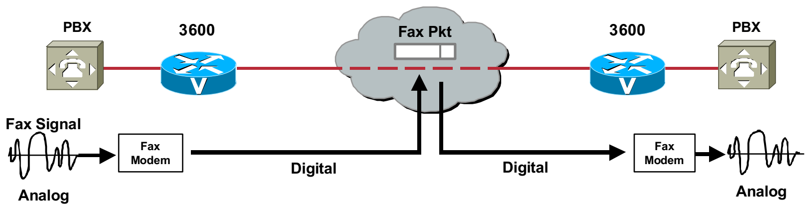
FRF.11 Annex A specifies the transfer syntax for dialed digits. Some of the higher compression, lower-quality compressor/decompressors (codecs) do not always accurately detect, compress, and reproduce DTMF digits.

Note Recall that the higher-compression codecs are tuned for the voice range.

Annex A specifies the syntax for the originating Voice Frame Relay Access Device (VFRAD) to detect the DTMF digits and encode the digits as data. The destination VFRAD decodes the data and reproduces the DTMF digits.

Fax Relay: FRF.11 Annex D

Cisco.com



Payload types:
- Modulation turn-on
- Modulation turn-off
- T.30 payload
- T.4 payload

Relay commands:
- Modulation off
- Modulation on
- Data
- HDLC end-of-frame
- HDLC frame-abort

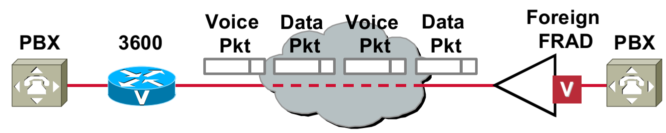
Supported rates:
- Single frequency tone
- V.21 300 bps
- V.27ter 2400 bps
- V.27ter 4800 bps
- V.29 7200
- V.29 9200
- V.33 12000
- V.33 14400
- V.17 7200
- V.17 9600

EVODD v3.3—2-24

Annex D specifies the fax relay syntax. Annex D is comprised of the fax relay transfer procedure and the fax relay payload format.

FRF.11 Inoperability

Cisco.com



- In theory, FRADs from different vendors can interoperate with FRF.11
- Requires industry-standard FRF.11 connection

EVODD v3.3—2-25

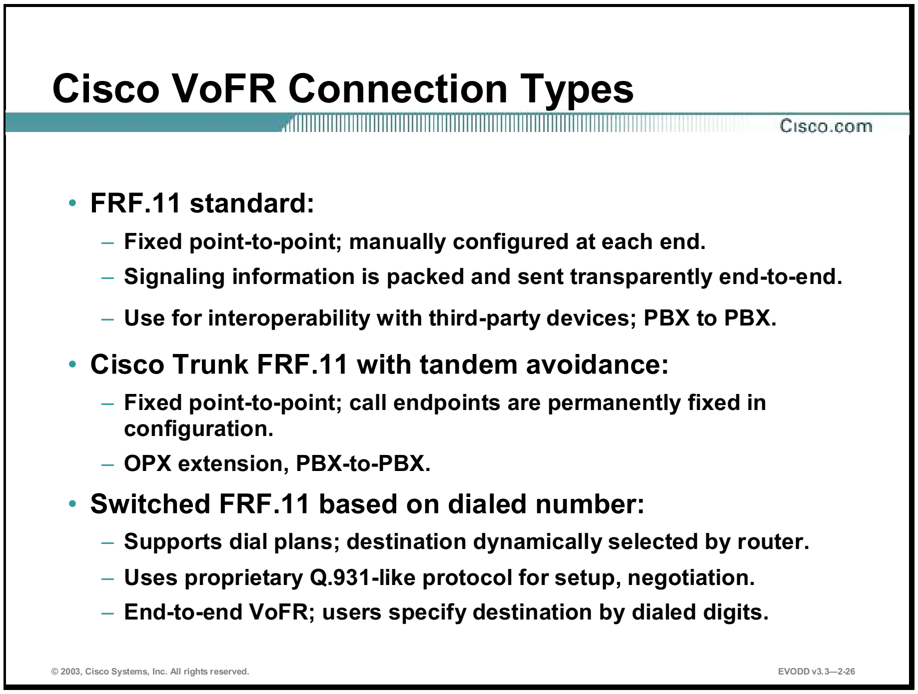
In theory, Frame Relay DTE from different vendors can interoperate using FRF.11. However, proprietary vendor extensions will not interoperate; only the minimal common subset of supported syntax will interoperate.

Use FRF.11 for interoperation with other vendor equipment only. For Cisco to Cisco equipment, use Cisco trunk.

Interoperability, even with FRF.11, is not guaranteed at any level.

Cisco VoFR Connection Types

This topic examines VoFR connection types. FRF.11 allows for multiple connection types. Each connection type has unique characteristics. Various factors determine the connection type that you use.



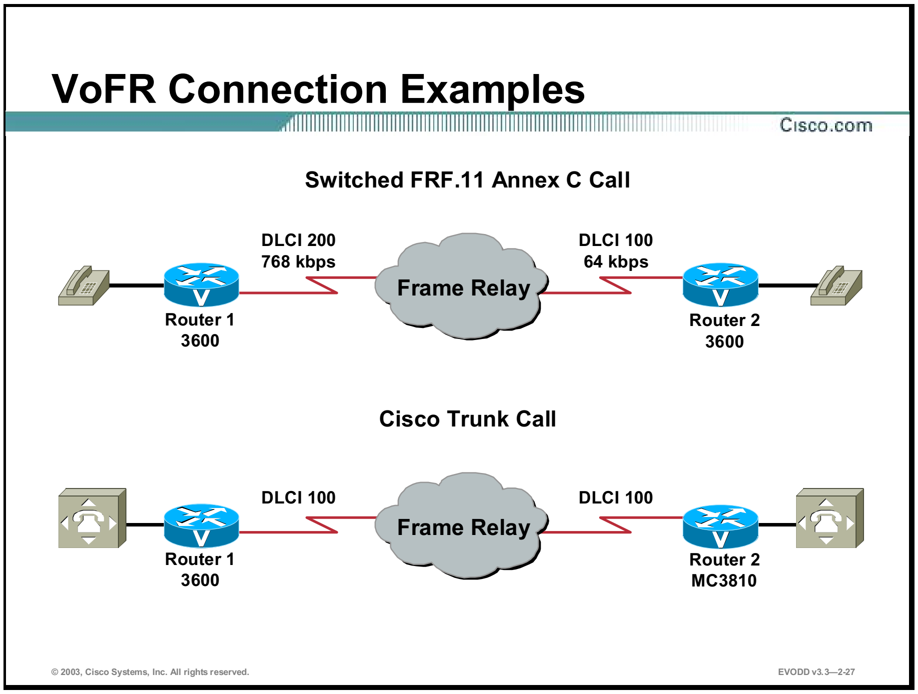
Static FRF.11-trunks allow for standards-based vendor interoperability by specifying the frame format and coder types used when transmitting voice through a Frame Relay network. However, FRF.11 has no specifications for end-to-end negotiation, call setup process, or any other form of communication between Frame Relay nodes. As a result, static FRF.11 trunks must be set up by manually configuring each router in the voice trunk path with compatible parameters. A voice port and a specific subchannel on a DLCI are explicitly bound on each end router. Signaling information is packed and sent transparently end-to-end.

Once configured, a static FRF.11-trunk remains active until the voice port or serial port is shut down, or until a network disruption occurs. The FRF.11 specification does not include any standardized methods for performing Operation, Administration, and Maintenance (OAM) functions. There is no standard protocol for detecting faults and providing rerouting of connection paths.

Cisco switched VoFR addresses the lack of end-to-end call parameter negotiation and call setup syntax in FRF.11 with a Q.931-like session protocol running on a user-configurable CID of an FRF.11-format multiplexed DLCI. This proprietary Cisco switched VoFR protocol handles call setup and parameter negotiation for both endpoints and intermediate nodes within the (multihop) call path. The call setup mechanism, originally implemented in the Cisco MC3810, is used for either permanent switched (Cisco trunk) or dynamic switched calls. The Cisco switched VoFR protocol includes forwarding the called telephone number and supports tandem switching of the call over multiple Frame Relay PVC hops.

Cisco trunk (private line) calls are basically normal, dynamic switched calls of indefinite duration that use a fixed destination telephone number and include optional transparent end-to-end signaling. The telephone number of the destination endpoint is permanently configured into the router so that it always selects a fixed destination. Once established, either at bootup or when configured, the call stays active until one of the voice ports or network ports is shut down, or until a network disruption occurs.

Dynamic switched calls are regular telephone calls in which the dial-plan-based call switching is performed by the Cisco router. The router, based on the dialed telephone number and the dial-plan information contained in the dial-peer configuration entries, selects the destination endpoint of the call. Contrast this implementation with permanent calls (Cisco trunk calls), in which the call endpoints are permanently fixed at configuration time.



This illustration shows examples of VoFR connections. The top network topology depicts a switched VoFR call between two telephones. When one of the telephones dials the other telephone, the local router examines its dial plan information to determine which DLCI it should send the voice call, in order to reach the appropriate destination telephone number. After the call is completed (for example, when one of the parties hangs up their telephone) the circuit is torn down. The call is considered a "switched" call, because the circuit is not continuously in a connected state.

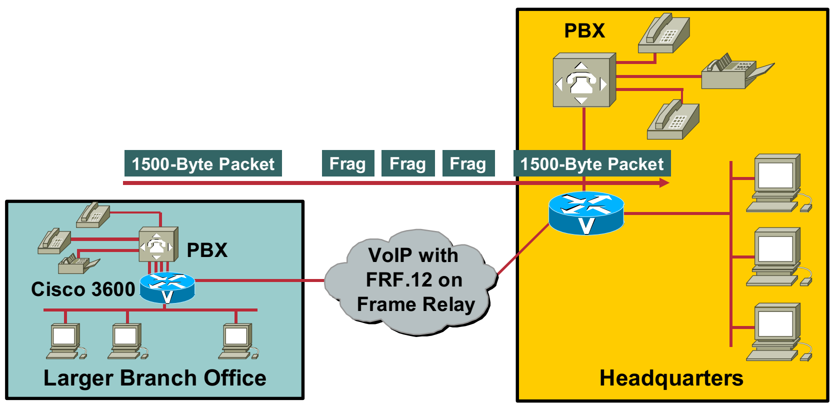
The bottom network topology depicts two Cisco routers interconnected via a Cisco trunk. A trunk, as opposed to a switched circuit, is always connected. A typical application for a trunk call is the interconnection of PBXs at remote sites. When one PBX extension dials a remote PBX extension, the call setup time is minimal because a trunk is always connected. Also, by interconnecting PBXs in this fashion, your customer may not need a tie-line.

Minimizing Delay and Loss

This topic examines how to minimize delay and loss. Due to Frame Relay characteristics, delay and loss are major concerns. Voice traffic does not deal well with delay and loss.

Minimizing Delay: Fragmentation with FRF.12

Cisco.com



- Multiple PVCs, some with FRF3.1 data on same physical interface
- Can be used with VoIP
- Fragments and interweaves large data frames with voice packets

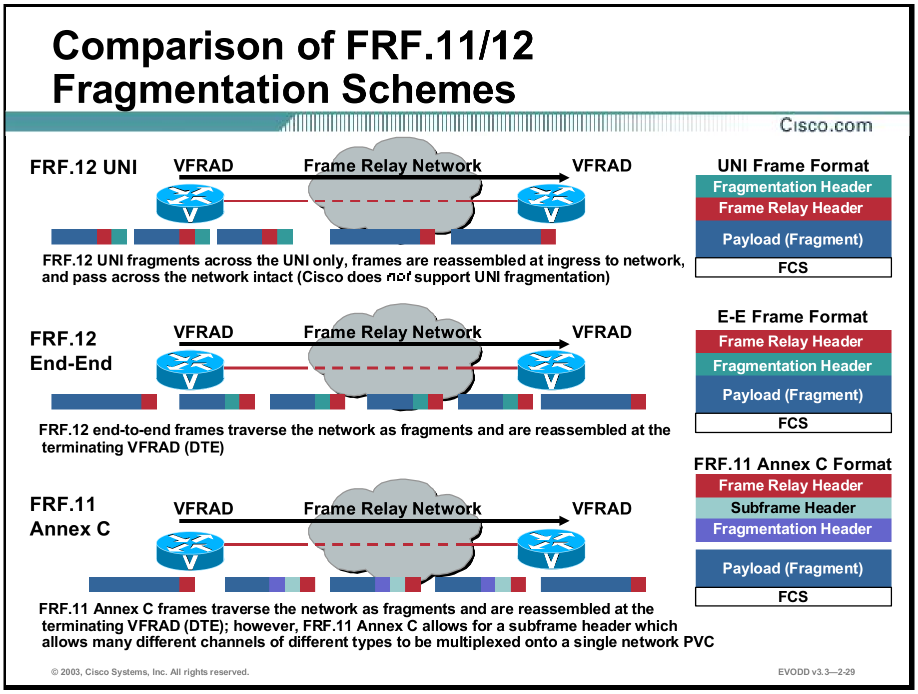
 EVODD v3.3—2-28

While FRF.11 fragments data, FRF.12 fragments the non-FRF.11 frames (normally FRF.3.1 frames) that exceed the configurable threshold.

The Frame Relay multiplexing model with FRF.11 also allows support of FRF.3.1 PVCs and DLCIs on the same physical interface. One of the non-FRF.11 PVCs transmitting a large data frame would cause a voice frame to be delayed.

FRF.12 is transparent to the data frame; it simply fragments any frame on the interface larger than a configurable packet size. For this reason, FRF.12 alone is not enough and networks need some form of queuing to interweave the voice frames with the data frame fragments.

Because FRF.12 is transparent to the data and fragments based on the packet size only, it works well when VoIP is the encapsulation method (versus VoFR).



There are two forms of FRF.12:

- End-to-end
- User-Network Interface (UNI)

In the end-to-end procedure, the network is unaware that the data packet has been fragmented. The network nodes transparently handle the fragments from the data packet and the interwoven voice frames. The remote Frame Relay DTE receives and holds the first data fragment until the remaining data fragments arrive. Voice frames (or interwoven data frames) are forwarded immediately.

In the end-to-end form of FRF.12, the fragmentation header is *inside* the Frame Relay header, and is passed transparently by the network.

In the UNI form of FRF.12, the network actually participates in the FRF.12 procedure. The Frame Relay DTE fragments data packets exceeding the threshold, and interweaves voice packets. The Frame Relay data communications equipment (DCE), the network node, receives the fragmented data segments with the voice frames in between. In the UNI procedure, the Frame Relay DCE network node holds the first data fragment and reassembles the data frame to be forwarded across the network. Voice frames (or any other frames received smaller than the fragmentation threshold) are forwarded ahead of the reassembled data frame.

In the UNI form of FRF.12, the fragmentation header is *before* the Frame Relay header, and is stripped before the frame is forwarded across the network.

Note The UNI form of FRF.12 is a special case, in which it is actually desirable for the network to forward frames out of order, or in priority order.

Fragmentation Frame Size Matrix

Cisco.com

Real-Time Packet Interval

Link Speed	10 ms	20 ms	30 ms	40 ms	50 ms	100 ms	200 ms
56 kbps	70 bytes	140 bytes	210 bytes	280 bytes	350 bytes	700 bytes	1400 bytes
64 kbps	80 bytes	160 bytes	240 bytes	320 bytes	400 bytes	800 bytes	1600 bytes X
128 kbps	160 bytes	320 bytes	480 bytes	640 bytes	800 bytes	1600 bytes X	3200 bytes X
256 kbps	320 bytes	640 bytes	960 bytes	1280 bytes	1600 bytes X	3200 bytes X	6400 bytes X
512 kbps	640 bytes	1280 bytes	1920 bytes X	2560 bytes X	3200 bytes X	6400 bytes X	12800 bytes X
768 kbps	1000 bytes	2000 bytes X	3000 bytes X	4000 bytes X	5000 bytes X	10000 bytes X	20000 bytes X
1536 kbps	2000 bytes X	4000 bytes X	6000 bytes X	8000 bytes X	10000 bytes X	20000 bytes X	40000 bytes X

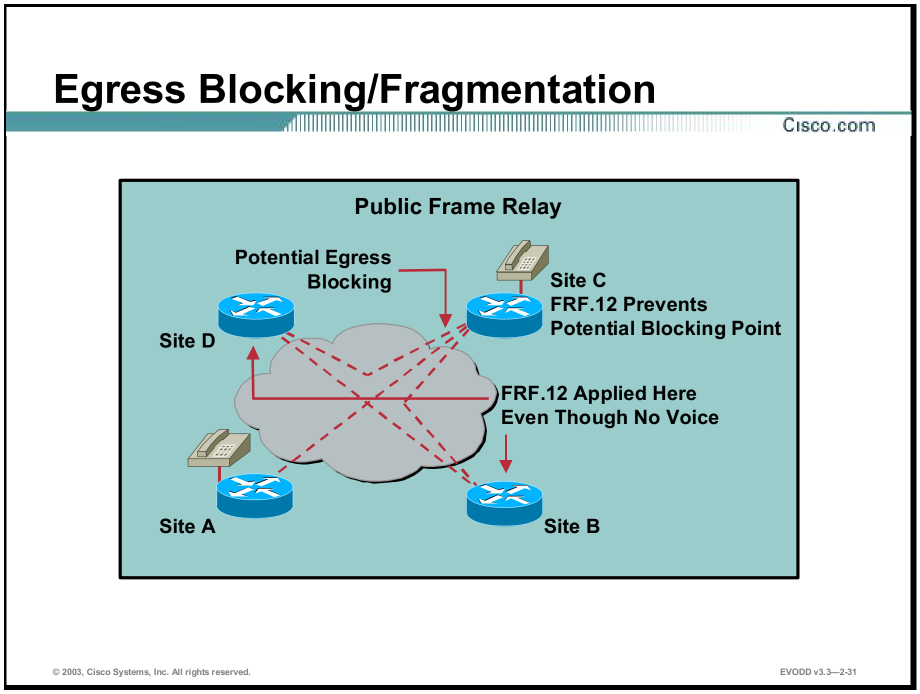
X—Fragmentation not an issue due to bandwidth + interval combination

EVODD v3.3—2-30

Large frames need to be fragmented because when these frames precede voice frames on a low-speed link, the residency or occupancy of the large data frame will cause a significant penalty in the delay budget. Fragmentation of large data frames allows voice frames to be interwoven with the data fragments.

The chart shown here is a guide for calculating the fragment size based on the maximum delay the low-speed link may add to the delay budget. For example, if the maximum acceptable delay that the low-speed link may add to the budget is 20 ms, and the link speed is 64 kbps, then the maximum fragment size may be no more than 160 bytes. The blocks marked with a red "X" are generally of sufficient speed that fragmentation is not required or even desirable, because when a frame is fragmented, each fragment has its own header information.

Note This chart, in conjunction with the delay budget chart, is a powerful tool set for designing a VoFR network.

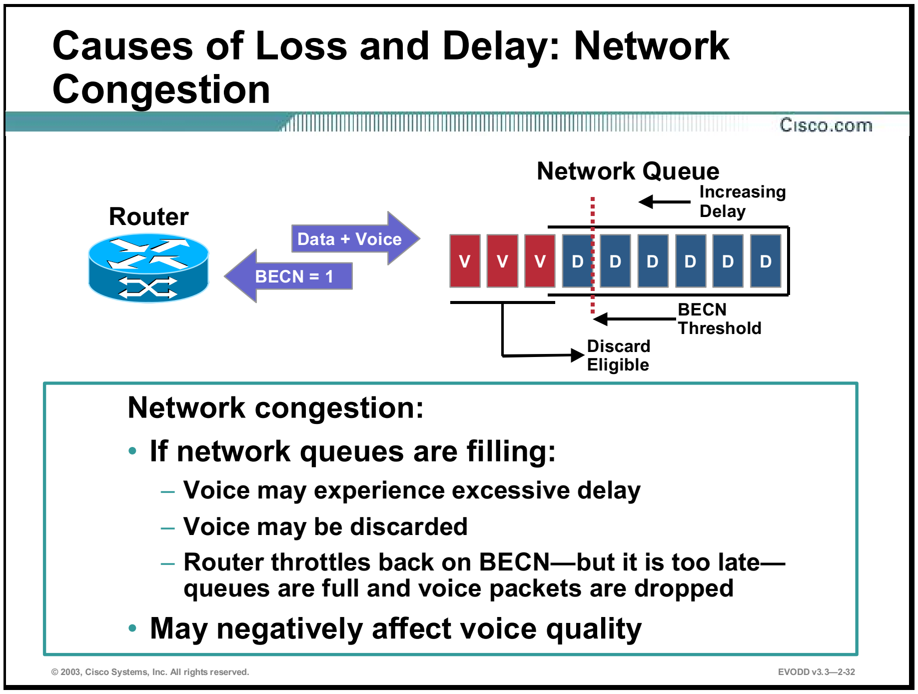


This network depicts one of the frequent problems encountered when designing VoFR networks.

If residency of a large data frame can block a voice frame from being placed on the link at the ingress to the network, then large data frames from non-voice sites can block voice frames being placed on low-speed links at the egress of the Frame Relay network to the destination customer premises equipment (CPE).

This problem is frequently referred to as egress blocking. Fragmenting non-voice-enabled sites will not solve the problem entirely, but it does allow voice frames to be queued and transmitted in between data fragments from a large data frame.

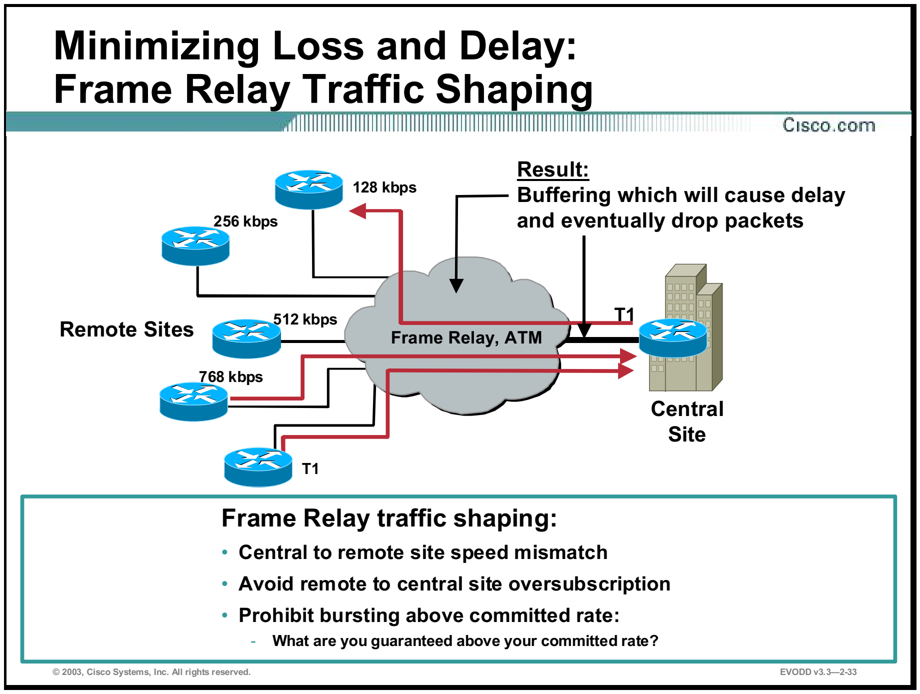
Voice frames can still be trapped in a queue behind a large number of small data fragments due to multiple sites transmitting data concurrently and the higher speed of the backbone trunks relative to the egress link speed. If non-voice sites are bursting to the voice egress, the trapping of these voice frames can result in significant delays. Priority queuing from the service provider can also be effective in helping to solve the problem of egress blocking.



Another frequent problem in VoFR network design is network congestion, which affects Frame Relay service users and providers.

Many users have configured the committed information rate (CIR) of their routers to the link speed instead of the CIR that is contracted with the service provider. This configuration is used with the idea of bursting to the link speed for long periods when the network is not congested. Users find that they need a higher CIR to achieve voice quality in their VoFR network.

However, to achieve the voice quality required, users will need to match their traffic to their contracted service. This process may not improve throughput significantly because of the aggressive back-off mechanisms of TCP.



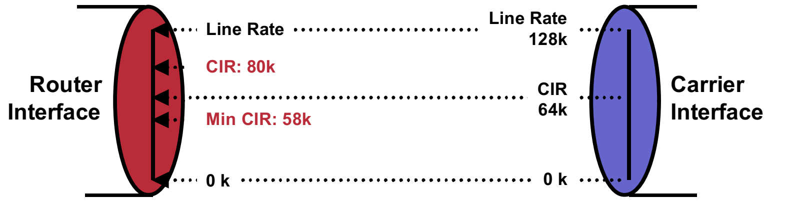
A priority queuing mechanism and Frame Relay traffic shaping are necessary to minimize delay and loss. In addition to some form of priority queuing mechanism, Frame Relay traffic shaping allows better utilization of the contracted parameters by staying with the CIR and burst parameters.

In combination, priority queuing and Frame Relay traffic shaping generally provide an acceptable level of quality of service (QoS) for voice, even in single-PVC environments.

Note In poorly designed networks, Frame Relay traffic shaping and priority queuing may be insufficient for the preservation of voice quality.

Single-PVC Design Considerations

Cisco.com



The following are Single-PVC design considerations:

- **Goal is best voice quality, *not* data volume**
- **Contract CIR to meet calculated requirements**
- **Shape conservatively**
- **Have carrier set BECN threshold to a smaller setting**
- **Set traffic shaping parameters for finest granularity**

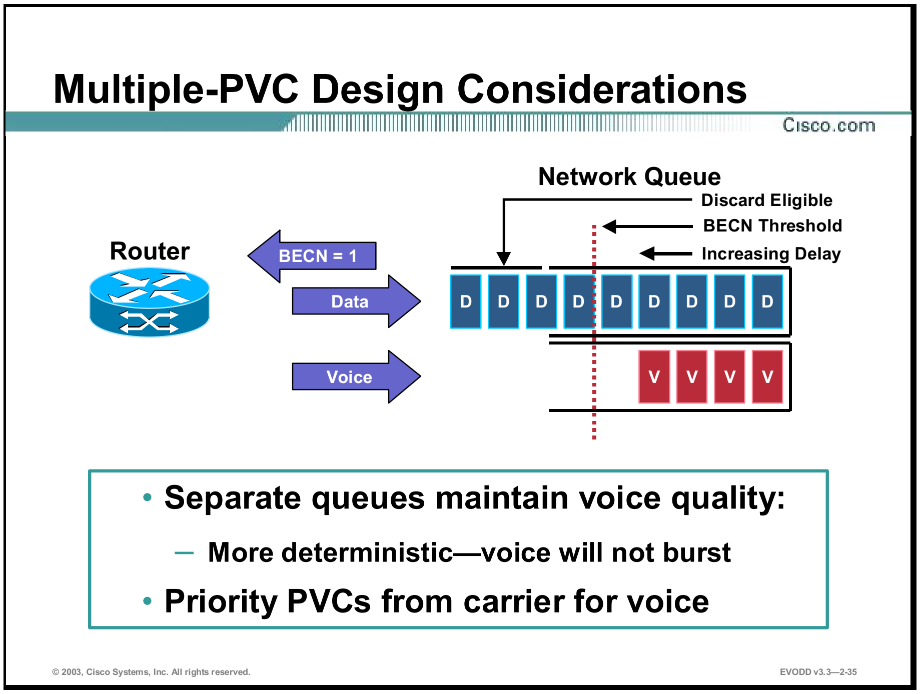
EVODD v3.3—2-34

In designing a VoFR network, the number one priority is voice quality. In transitioning to an integrated network, voice quality standards must meet or exceed the present environment, or the service will be perceived as poor quality.

Note Conservative engineering is needed to meet this requirement in a single PVC configuration.

The CIR should be sized to accommodate good voice quality. In star configurations with single PVCs, the CIR needs to accommodate not only voice and data from the hub to the peripheral sites, but also any voice and data frames that are switched through the central site.

Some form of priority queuing is an absolute requirement, not only for voice frames that originated at the site, but also for voice frames that are switched through the site. Set traffic shaping parameters to the finest granularity possible, and shape traffic conservatively.



In configurations with separate PVCs for voice and data, the potential voice load is better understood and is partially isolated from the data load.

The CIR from the voice load can be closely calculated, even in a star configuration with calls switched through the hub.

Network behavior becomes easier to determine with the correct CIR for the voice PVCs. Voice frames and data frames are not competing for the same bandwidth under the CIR.

You can implement priority queuing end-to-end if priority PVCs are available from the service provider.

Number of PVCs Criteria

Cisco.com

Choose the number of PVCs using these criteria:

- **Single PVC for voice and data:**
  - Compelling if PVCs mean extra $$
  - Limit bursting to committed rate (CIR), or use PQ-CBWFQ
  - Fragment if link on either side of the network is slow speed
  - Easier IP routing configuration for VoIPovFR
  - Data can leverage BW when little or no voice is present
- **Separate PVCs for voice and data:**
  - Possible solution if you own your FR backbone and PVCs can be added at no cost
  - PVC for data can burst above CIR; voice can be shaped to CIR
  - Bandwidth must be dedicated to voice or data
  - Fragmentation rules still apply for data PVC
  - IP routing becomes complicated if VoIPoFR

EVODD v3.3—2-36

A network design with single PVC for voice and data has the following characteristics:

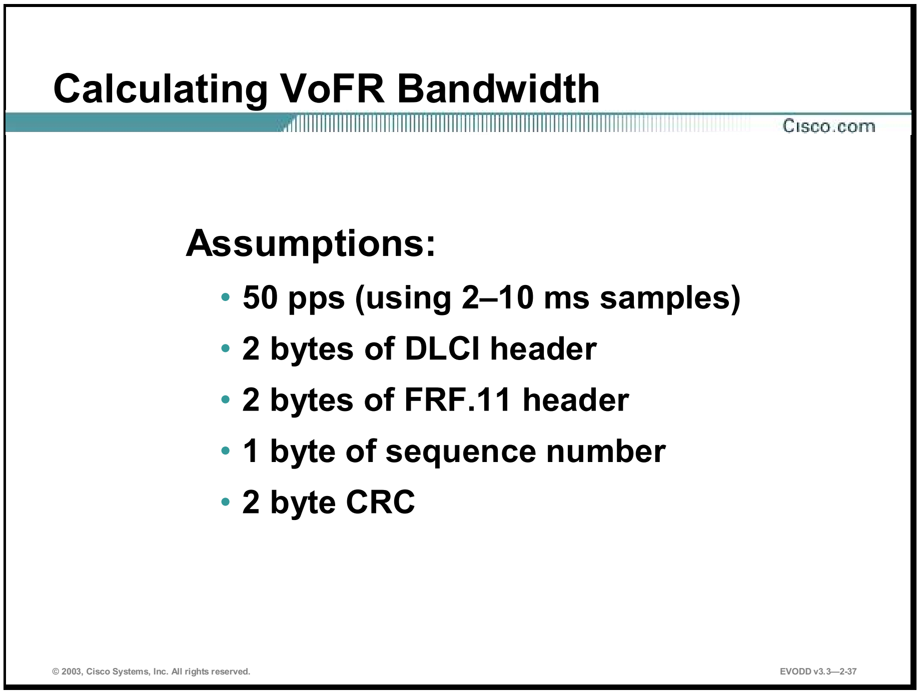
- Compelling if the Frame Relay provider charges based on the number of PVCs
- Limit bursting to CIR, or use priority queuing and/or class-based weighted fair queuing (PQ/CBWFQ), also referred to as Low Latency Queuing
- Fragment if the link on either side of the network is slow speed
- Easier IP routing configuration for VoIP over Frame Relay
- Data can leverage bandwidth when little or no voice is present

A network design with separate PVCs for voice and data has the following characteristics:

- Possible solution if Frame Relay backbone is owned and PVCs can be added at no cost
- PVC for data can burst above CIR; voice can be shaped to CIR
- Bandwidth must be dedicated to voice or data
- Fragmentation rules still apply for data PVC
- IP routing becomes complicated for VoIP over Frame Relay

Bandwidth Calculations

This topic describes how to make bandwidth calculations. To properly size the CIR for voice PVCs, or for the voice portion of the combined voice and data PVC, you must understand the bandwidth that each voice call will require.



In this example, a G.729 codec with a 20 ms sample time is used.

The remaining assumptions regarding Frame Relay and FRF.11 overhead will be the same, regardless of the codec and sample rate:

- 50 packets per second (pps), using 2–10 ms samples
- 2 bytes of FRF.11 header
- 1 byte of sequence number
- 2 byte cyclic redundancy check (CRC)

Note The pps stream generated by the codec will always be the inverse of the sample time, in this case 1/20 ms.

For the purpose of this calculation, it is assumed that there is no silence suppression. The bandwidth that is being calculated assumes continuous transmission, such as that of a radio being played into the telephone. In most real-life telephone conversations, people take turns speaking, so bandwidth is gained from silence suppression.

Calculating VoFR Bandwidth (Cont.)

Cisco.com

Voice payload calculation:

- **20 ms voice sample x 8 kbps (for G.729)/8 bits/byte = 20 bytes**
- **Note: To derive the payload for G.711, substitute 64 kbps = 160 bytes**

Packet size calculations:

- **20 byte payload + 7-byte header = 27 bytes (header = DLCI/FRF.11/seqn/CRC)**

Bandwidth calculations:

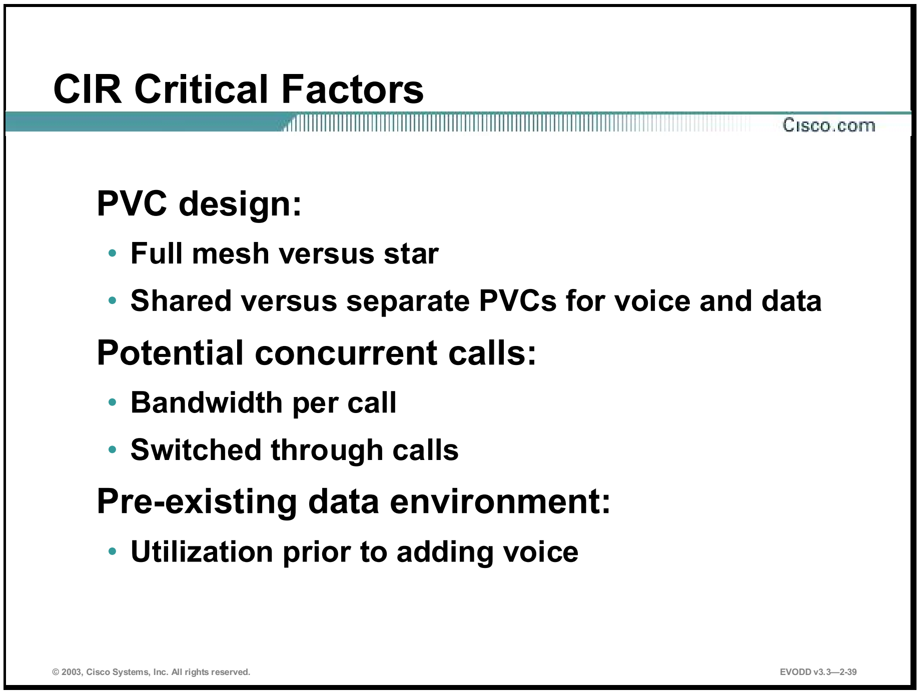
- **27 byte/voice packet x 8 bits/byte x 50 pps = 10.8 kbps per call**

EVODD v3.3—2-38

In this example, voice payload will be 20 bytes. With overhead, total packet size will be 27 bytes.

If the codec is G.711 using pulse code modulation (PCM), the payload calculation equals 160 bytes, and the packet size equals 167 bytes. For G.711, bandwidth is approximately 67 kbps.

Note	You may have a marginal rate of error based on the look-ahead or error recovery bytes being omitted from the payload calculation; G.711 is actually more accurate because it does not use a look-ahead.



After understanding the bandwidth for a call, you must calculate the CIR for PVC. For this calculation, you must understand the number of calls on the PVC. You must also consider whether the PVC will be reserved for voice, or shared with data.

If the PVC is shared with data, you must consider the preexisting data environment.

You must also consider the call matrix. For example, you need to know if the design is a full mesh of PVCs, or if calls can be switched through some sites.

Important elements of proper VoFR network design include priority queuing, Frame Relay traffic shaping, and the correct PVC design. PVC design elements include using a single PVC for voice and data, or using separate PVCs for voice and data. Sharing a PVC for voice and data is more economical, but separate PVCs yield better voice quality. The other element of PVC design is the configuration of the PVCs into either a full mesh or star design. The configuration of PVCs for voice will usually depend on the telephone call usage matrix.

Avoid multiple compression and decompression cycles of voice, especially with high-compression codecs, due to voice-quality issues. You should calculate the elements of the delay budget and calculate the total end-to-end delay to determine the network design.

Non-voice sites may send large data blocks that block voice frames at the egress to the network, requiring network implementation of FRF.12 at non-voice-enabled sites.

You can calculate the bandwidth required for voice calls as one element of the CIR, based on the relationship between the CIR and the PVC design.

Summary

This topic summarizes the key points discussed in this lesson.

Summary

Cisco.com

- The benefits of VoFR include increased optimal utilization of bandwidth, and additional telecommunication and Internet-related services for customers.
- Common voice applications used over Frame Relay include tie-line replacement, Off-Premise eXtensions, and end-to-end switched voice.
- FRF.11 is the Frame Relay forum standard for voice.
- FRF.12 is the Frame Relay forum standard for fragmentation of data packets.
- Cisco VoFR connection types include FRF.11 standard, Cisco Trunk FRF.11 with tandem avoidance, and switched FRF.11 based on dialed number.
- Design strategies include priority queuing, Frame Relay traffic shaping, and correct PVC design.
- It is necessary to understand and calculate the bandwidth that each voice call requires.

 EVODD v3.3—2-40

Next Steps

After completing this lesson, go to:

- Cisco CallManager Overview and Gateway Selection module

References

For additional information, refer to these resources:

- Frame Relay Fragmentation for Voice:

  http://www.cisco.com/warp/public/788/vofr/fr_frag.html

- VoIP Over Frame Relay with QoS:

  http://www.cisco.com/warp/public/788/voice-qos/voip-ov-fr-qos.html

- Frame Relay Forum:

  http://www.frforum.com/

Laboratory Exercise: Voice over Frame Relay

The laboratory exercises are designed to reinforce concepts discussed throughout the course. This laboratory exercise highlights design considerations for Voice over Frame Relay networks.

Exercise Objective

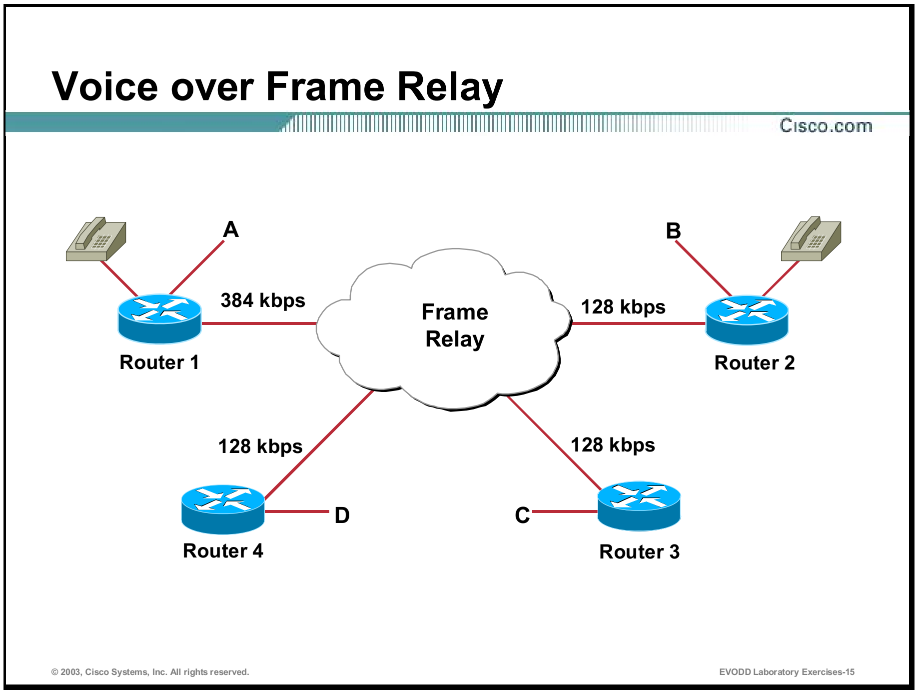
In this exercise, you will answer questions based on a Voice over Frame Relay topology.

After completing this exercise, you will be able to:

- Determine the proper protocols to use in various situations on a Voice over Frame Relay topology
- Determine when traffic shaping is necessary

Exercise Procedure

Study this topology and answer the following questions.



Practice

Both voice and data will use the Frame Relay ports in the diagram.

Q1) What protocol should be implemented to keep data from placing too much delay on the voice?

Q2) What link speed would provide enough bandwidth so that data would not have to be fragmented?

Q3) Which ports should FRF.12 be implemented on?

Q4) If third-party vendor interoperability is required for voice, then what fragmentation method should be placed on the ports?

Q5) If the call endpoints are permanently fixed and no third-party interoperability is required, then what fragmentation method should be placed on the ports that voice will be crossing?

Q6) If dynamic switched voice is required, then what fragmentation method should be placed on the ports?

Q7) Will traffic shaping be required for this network?

Lesson Review

This practice exercise reviews what you have learned in this lesson.

Q1) Which two of the following benefits apply to VoFR? (Choose two.)

A) integrates voice and data in a homogeneous topology

B) enables managed network services

C) adds the performance benefits of tandem switching

D) utilizes Frame Relays built-in QoS service classes

Q2) What type of VoFR trunk is used for tie-line replacement between PBXs?

A) FRF.8

B) FRF 3.1

C) FRF.11

D) FRF.12

Q3) What encapsulation method is used to transport data across a Frame Relay network?

A) FRF.11

B) FRF.12

C) FRF.5

D) FRF 3.1

Q4) How many subchannels does FRF.11 support on a single DLCI?

A) 1

B) 23

C) 127

D) 255

Q5) Which two of the following applications would be used for a Cisco trunk FRF.11 connection with tandem avoidance? (Choose two.)

A) OPX extension

B) interoperability with third-party FRADs

C) PBX-to-PBX connection

C) end-to-end VoFR (for example, when users specify dialed digits)

Q6) In a hub and spoke Frame Relay network, where some routers carry VoIPoFR and data traffic, and some routers carry only data, where should FRF.12 be implemented?

A) only on the hub router

B) only on routers carrying voice traffic

C) on all spoke routers

D) on all routers

Q7) Approximately how much bandwidth is required for a VoFR call using the G.729 codec (assuming 50 voice samples per second, a 2 byte FRF.11 header, and a 1 byte sequence number)?

A) 26.4 kbps

B) 8 kbps

C) 32.6 kbps

D) 10.8 kbps

3

Cisco CallManager Overview and Gateway Selection

Overview

The Cisco CallManager (CCM) is the primary software application that extends enterprise telephony features and functions to enterprise packet telephony network devices. CCM is installed on a server-class PC and provides basic call processing, signaling, and connection services to configured devices.

This module will teach you about the architecture and scalability of CCM clusters. You will learn about digital signal processors (DSPs), Call Admission Control (CAC), and various CCM deployments. Additionally, the CCM needs access to devices other than Cisco IP Phones, such as the Public Switched Telephone Network (PSTN). Therefore, this module also details the available gateway hardware for connecting off of the IP LAN.

Upon completing this module, you will be able to:

- Describe CCM clusters
- Describe scalability planning
- Explain how to use DSPs for transcoding, conferencing, and media termination points (MTPs)
- Explain how to handle call admission
- Describe the four deployment types
- Identify the critical gateway features for Cisco IP telephony designs
- Describe the functionality and features of H.323, VG200, VG248, 6500, and 4000 gateways
- Describe the features and functionality of the IP telephony service
- Describe the features and functionality of survivable remote site telephony (SRST)

- Determine the proper gateway for a given environment

Outline

The module contains these lessons:

- Cisco CallManager Architecture
- Gateway Types

Cisco CallManager Architecture

Overview

Cisco Architecture for Voice, Video, and Integrated Data (AVVID) defines a framework for building and evolving customer networks that support Internet business solutions. Cisco AVVID, the only industry enterprise-wide, standards-based network architecture, provides a roadmap for combining business and technology strategies into one cohesive model. In order to deliver a complete end-to-end Cisco IP telephony solution, you must include a Cisco CallManager (CCM) in the design. This lesson will teach you about Cisco IP telephony. You will learn about IP telephony benefits, network elements or devices, network infrastructures, and IP telephony compatible software.

Importance

Adding CCM requires additional resources, so you must also understand the architecture of CCM to incorporate these resources and to ensure proper design.

Objectives

Upon completing this lesson, you will be able to:

- Describe the migration path to an IP telephony solution
- Describe the operation of CCMs in a cluster
- Describe how to scale CCM to support thousands of devices
- Evaluate the impact of different devices on the resources of a CCM
- Describe how CCM clusters communicate and function with one another
- Identify how CCM supports supplementary services (for example, conference calling)
- Describe Call Admission Control (CAC) techniques and how they assist in preserving quality of service (QoS)

- Identify the differences between possible CCM cluster deployments

Learner Skills and Knowledge

To fully benefit from this lesson, you must have these prerequisite skills and knowledge:

- A fundamental understanding of the services provided by legacy telephony systems
- A fundamental understanding of IP networking concepts

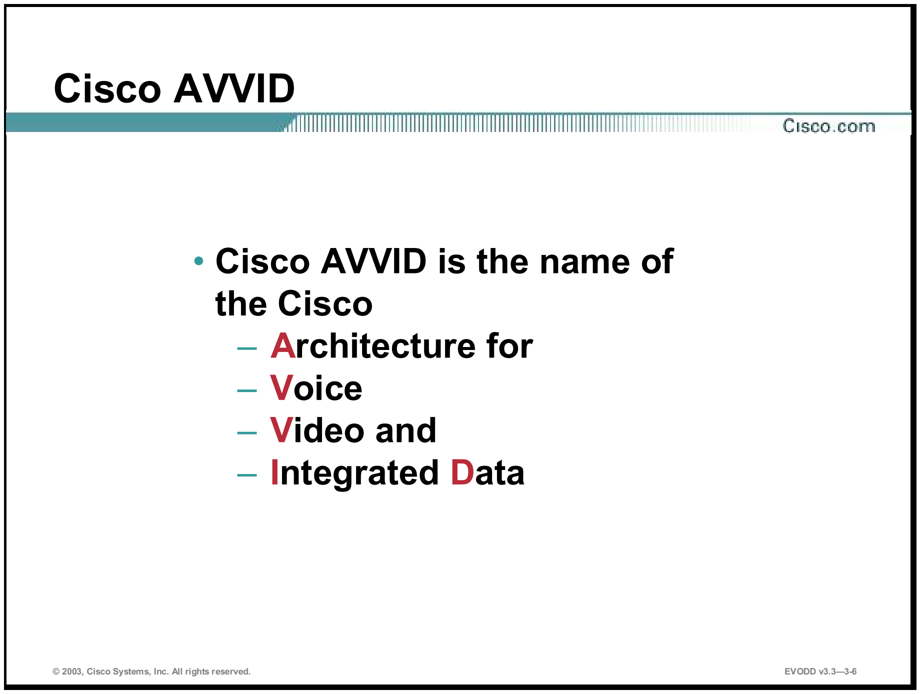
Outline

This lesson includes these topics:

- Overview
- IP Telephony Concepts
- CCM Clusters
- CCM Scalability
- Devices and Weights
- Intercluster Communication
- DSP Services
- CAC
- CCM Deployments
- Summary
- Laboratory Exercise: CCM
- Lesson Review

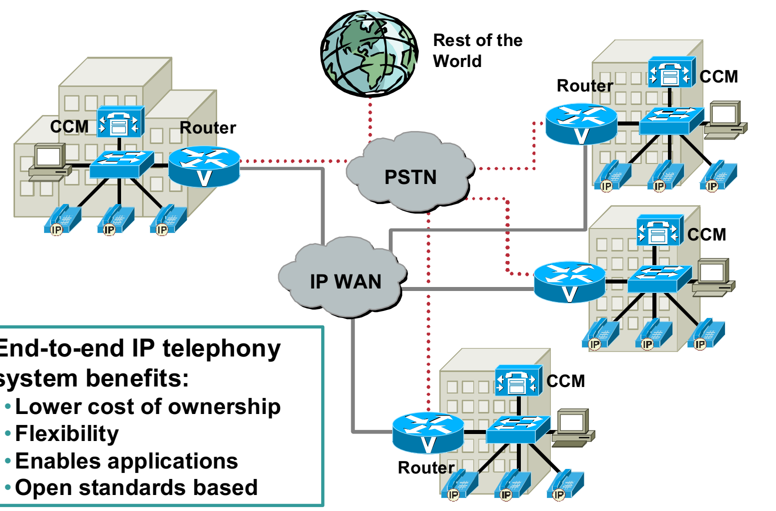
IP Telephony Concepts

This topic describes the main driving forces behind demand for the Cisco IP telephony solution and how they unify new applications. These driving forces are cost efficiency, flexibility, and better technology.



Cisco AVVID includes Cisco IP telephony offerings. IP telephony includes customer premises equipment (CPE) components, such as IP Phones, CCMs, gateways, gatekeepers, and unified messaging or voice-mail servers.

The Big Picture: IP Telephony Architecture Goals



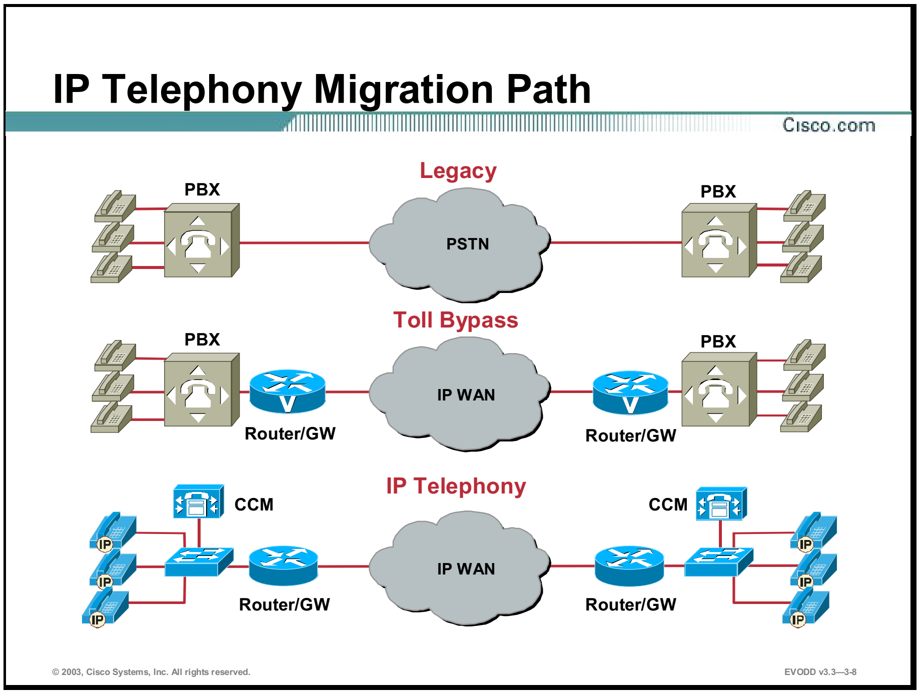
End-to-end IP telephony system benefits:
- Lower cost of ownership
- Flexibility
- Enables applications
- Open standards based

 EVODD v3.3—3-7

The Cisco IP telephony architecture goals provide end-to-end IP telephony. Cost savings to customers include moves and changes being simple, added productivity with new applications, toll bypass savings, and a lower cost of ownership. Cisco IP telephony goals include helping customers achieve added flexibility by enabling applications, such as unified messaging, virtual call centers, and many potential future applications.

The goal is to create an open, standards-based architecture that has the ability to plug-and-play components from various vendors, much like the architecture of the data environment today. This approach is the primary reason for the growth of the Internet.

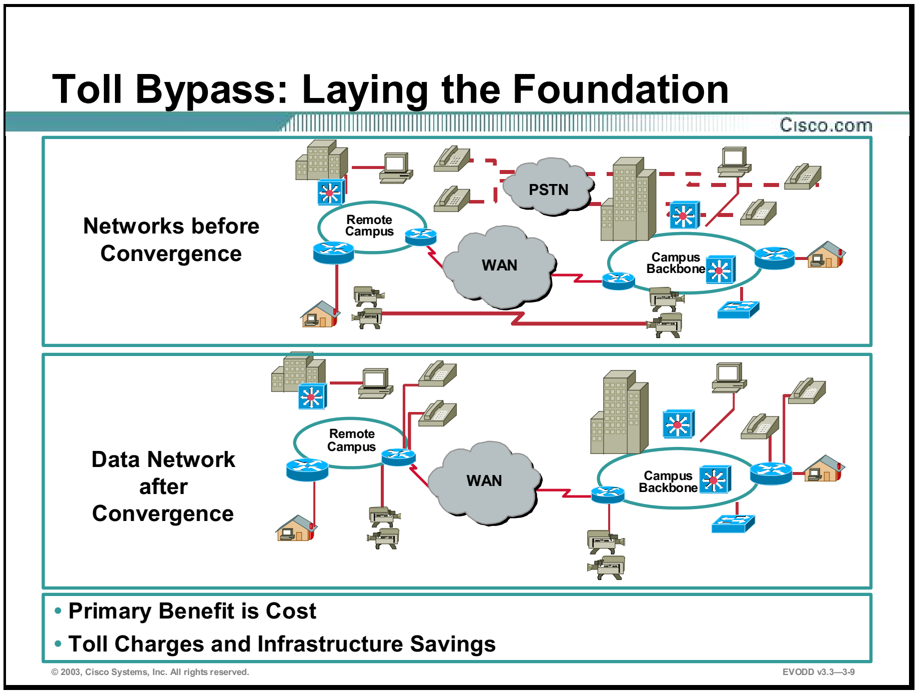
The Cisco IP telephony architecture also includes unified messaging under Cisco Unity.



An IP telephony migration typically starts with voice traffic traversing the Public Switched Telephone Networks (PSTNs) for intersite communications.

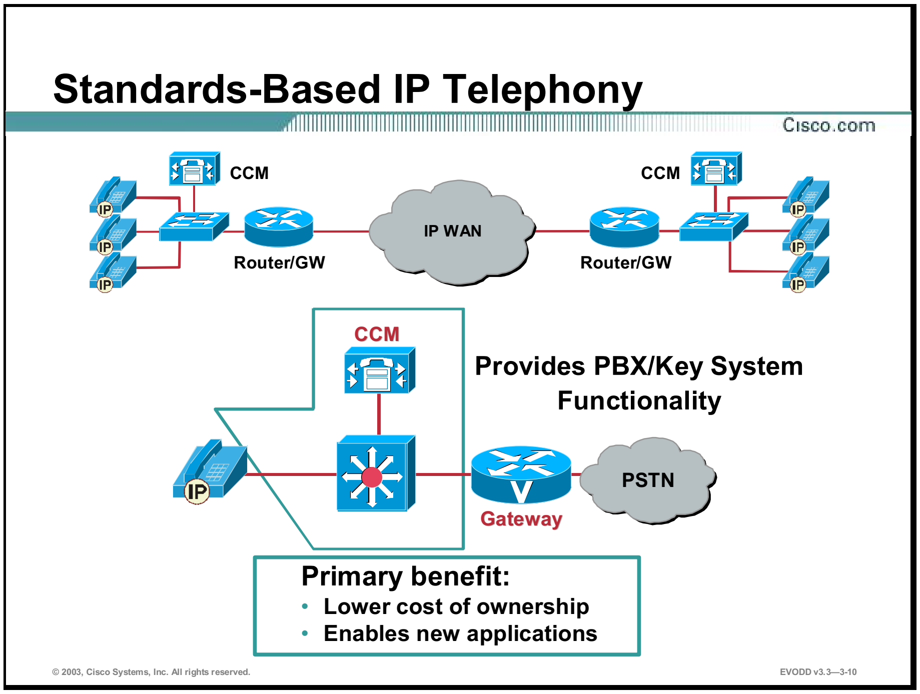
First, customers may migrate to deployments that bypass the PSTN. For example, they can eliminate the existing tie-lines that interconnect PBXs, and they can configure Cisco routers that interconnect via a WAN link to transport voice calls and PBX signaling. This approach eliminates the recurring charge for dedicated tie-lines between the PBXs (such as tie-line replacement).

Customers complete this strategy by migrating to the Cisco IP telephony architecture. The main goal of this phase is to provide end-to-end IP telephony, unified messaging, and other new services.



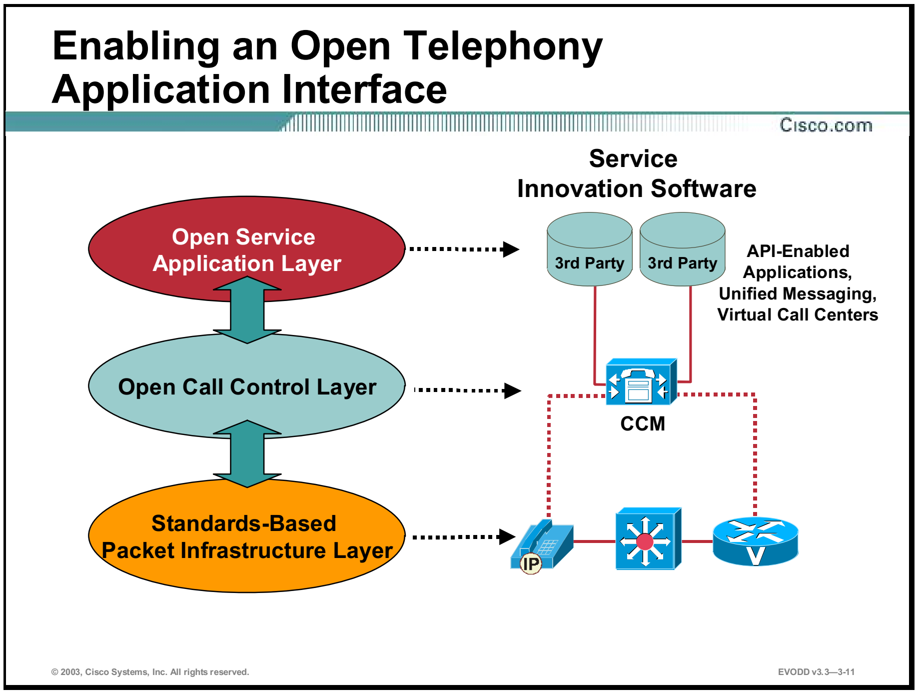
Many Cisco customers have migrated to a toll bypass environment for immediate cost savings. By laying the foundation with the products that participate in toll bypass, customers can install products that eventually become full IP telephony gateways.

Note Customers can use deployments that they originally installed for toll bypass to reap the benefits of the new services, such as unified messaging.



For cost savings and feature requirements, customers may move to standards-based IP telephony architectures (such as the Cisco IP telephony solution), which are more than IP telephony, telephone-to-telephone, and handset-to-handset. For example, the complete solution may include such features as unified messaging and conferencing services.

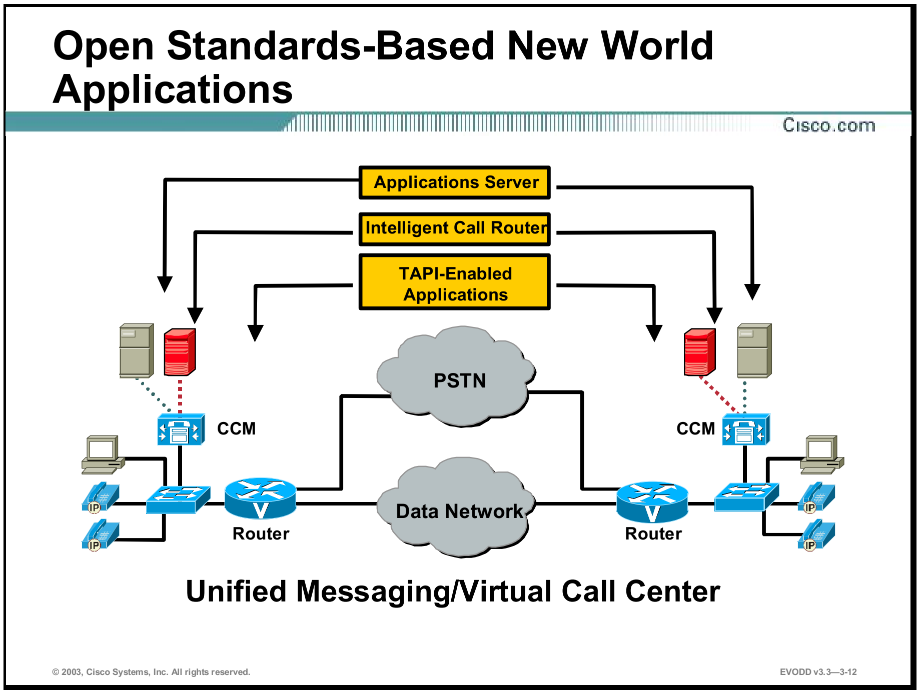
The IP telephony components of Cisco solutions provide the technology for IP Phones, gateways, call processing, and a switch fabric. Together, these devices function like key systems and PBXs.



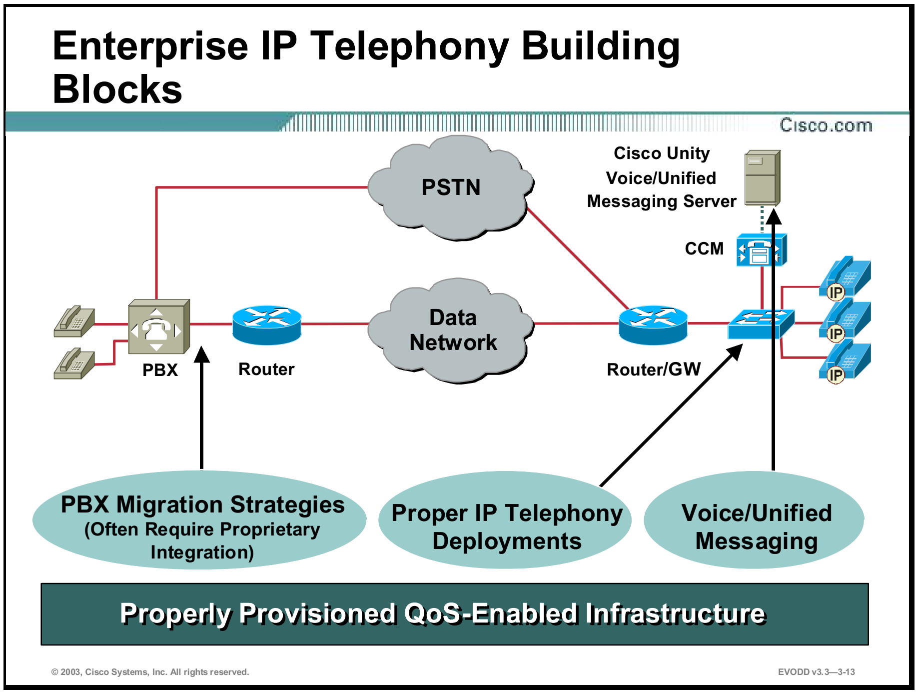
Customers will eventually move away from PBX because of new, open, standards-based applications. For now, however, most will not remove their PBXs. Therefore, an IP telephony design must work with the existing PBX of the customer. Offering public application programming interfaces (APIs) is the key to customer acceptance of IP telephony solutions. Then customers and third party developers can create API-enabled applications, such as virtual call centers, unified messaging, and automated attendant interactive voice response (IVR) and Personal Assistant.

Note Detailed information about the Cisco unified messaging solution, Unity, is available in the Cisco Unity Systems Engineering course.

The Cisco intent is to create a three-tiered layer with an open, standards-based packet infrastructure. Cisco created an open call control layer with standard-based protocols, such as Multimedia Gateway Control Protocol (MGCP) and the Skinny Station Protocol (which Cisco submitted to the standards bodies). For application deployment, the open service application layer supports the standards-based Telephony Application Programming Interface (TAPI), which is a Microsoft initiative, or the Java Telephony Application Programming Interface (JTAPI), which is a Sun Java-based TAPI initiative.



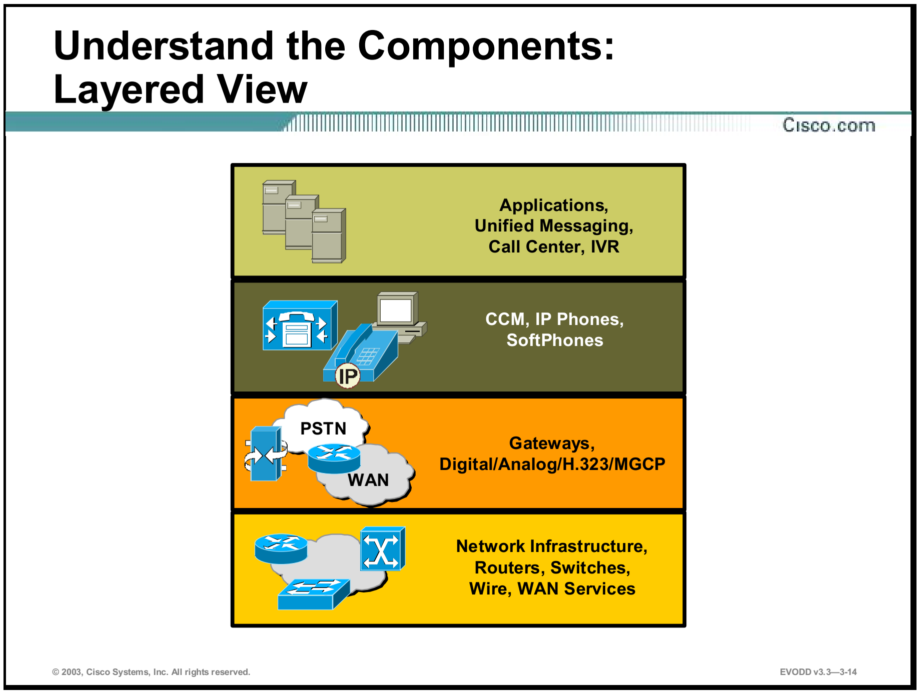
These New World applications are the foundation for open standards-based IP telephony applications, such as unified messaging and the virtual call center.



The primary function of voice telephony is call processing. Unified messaging (Cisco Unity), virtual call center applications, and additional new applications make up the architecture that compels a customer to move from a time-division multiplexing (TDM)-based network. Because these systems must have the correct infrastructure, they are considered critical components of this architecture.

You must understand the guidelines to successfully deploy Cisco IP telephony. The most crucial piece of a successful deployment is a QoS-enabled infrastructure. Cisco IP telephony uses the same QoS-enabled infrastructure as a toll bypass environment.

The CCM and Unity solutions allow a customer to move into larger deployments. A customer requires a migration strategy when moving from a TDM environment to an IP telephony environment. The coexistence of both environments is an essential part of this migration strategy.



One class of IP telephony components contains the CCM and the telephones. From a per-telephone perspective, there are two models to consider. One is peer-to-peer, which is equivalent to a cellular telephone, where the telephone is an autonomous entity and has the ability to make calls. The other is a client-server type of telephone, similar to the Cisco IP Phones, with a client-server model residing between the CCM and the telephone. It uses the Skinny Station Protocol, which is similar to MGCP. The Skinny Station Protocol is a stimulus-response protocol that runs between the two entities.

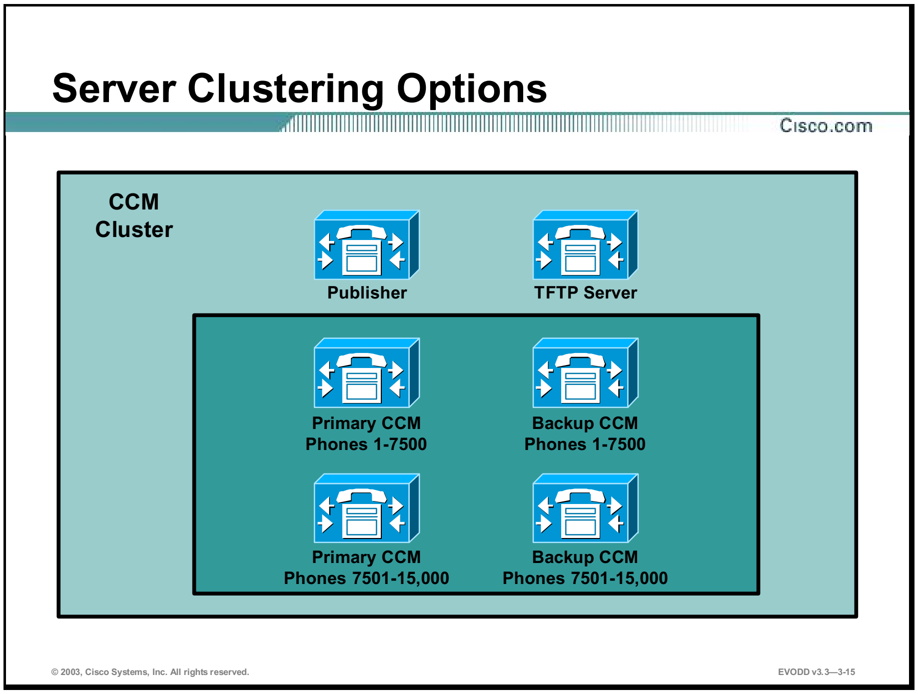
A second major class of IP telephony components is gateways. Gateways interface with PBXs, the PSTN, and other non-IP telephony devices, such as analog telephones.

A third class of components includes the unified messaging, call center, and IVR applications. These perform call processing, support specialized services, enable billing or cost allocation, and support IP telephony-enabled applications.

The fourth class of components includes the network infrastructure. Voice imposes some significant requirements on a network that are not typically present in a “data only” environment.

CCM Clusters

This topic describes how to make CCM redundant and scalable by using CCM clusters in Cisco IP telephony network designs. It also examines CCM clustering concepts.



A design is useful only if it can deliver the same or better reliability as a TDM telephone network. To accomplish this goal, the core of the system must be reliable. CCM clusters provide this reliability. CCM has two types of intracluster communications that provide a means of distributing call processing, thus providing redundancy.

Creating a well-designed CCM cluster is a key element to consider when you initially design an IP telephony network. The CCM cluster not only determines which CCMs have a copy of the Structured Query Language (SQL) database, but also the amount of redundancy designed into your voice network. A CCM cluster is a grouping of CCM servers supporting a designated group of Cisco IP Phones and gateways. When you look closer at the cluster functions, you see that there is a complex database replication and management structure designed to create redundancy and load balancing in the voice network. Because of this complex database replication and management structure, you must plan and design your cluster solution to meet the needs of the end users and to scale as their network expands and more services are required.

Your first design consideration is that CCM 3.3 supports a maximum of nine CCM servers that support the SQL database: one SQL publisher and up to eight SQL subscribers. These CCM database servers are responsible for the actual IP Phone registration and call processing. The number and type of CCM database servers installed in a cluster directly affects the number of telephones that the cluster is able to support. The SQL publisher server is responsible for managing the only writable copy of the SQL database; therefore, it typically does not participate in managing the call processing aspects of the network unless the IP telephony network contains less than 1000 IP Phones. You then have up to eight SQL subscriber servers to support the IP Phones and gateways.

The type of media convergence server (MCS) that you choose for these SQL database servers determines the number of devices the cluster is able to support. For example, using entry-level MCS 7815 servers in a cluster limits you to a maximum of 200 IP Phones per server. You can significantly increase this capacity by specifying MCS 7835 or 7845 servers in your design.

The second major design consideration is redundancy options. Cisco IP Phones support triple call processing redundancy so you can configure them to register with a primary, secondary, and tertiary CCM server. If the primary server fails, the telephone immediately tries to connect to the secondary server, and if the secondary server fails, the telephone attempts to contact the tertiary. The number of redundant servers in the network also directly affects the number of telephones that a cluster can support. With 1:1 redundancy, each primary server has a dedicated backup server.

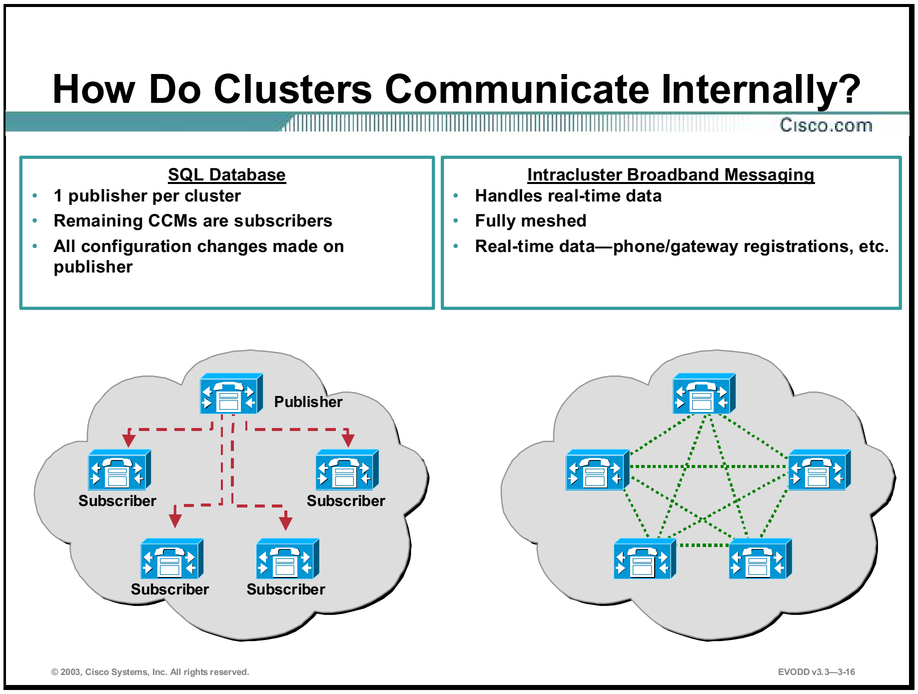
While this method guarantees that IP Phone registrations will never overwhelm the backup server, even if multiple primary servers fail, the 1:1 redundancy design limits the maximum cluster size considerably and is not cost effective.

A 2:1 redundancy design is another option. This design allocates a single backup server for every two primary servers and supports more IP Phones, but there is the risk of overwhelming the backup server if multiple primary servers fail.

If you use a 1:1 redundancy design, Cisco recommends that you configure load balancing across the devices (i.e., 50 percent to 50 percent), versus a 2:1 scheme where the backup does not have any registered devices and is waiting for devices to fail over. Therefore, the 2:1 redundancy scheme is 100 percent to 0 percent, while the 1:1 redundancy scheme uses 50 percent to 50 percent device load balancing.

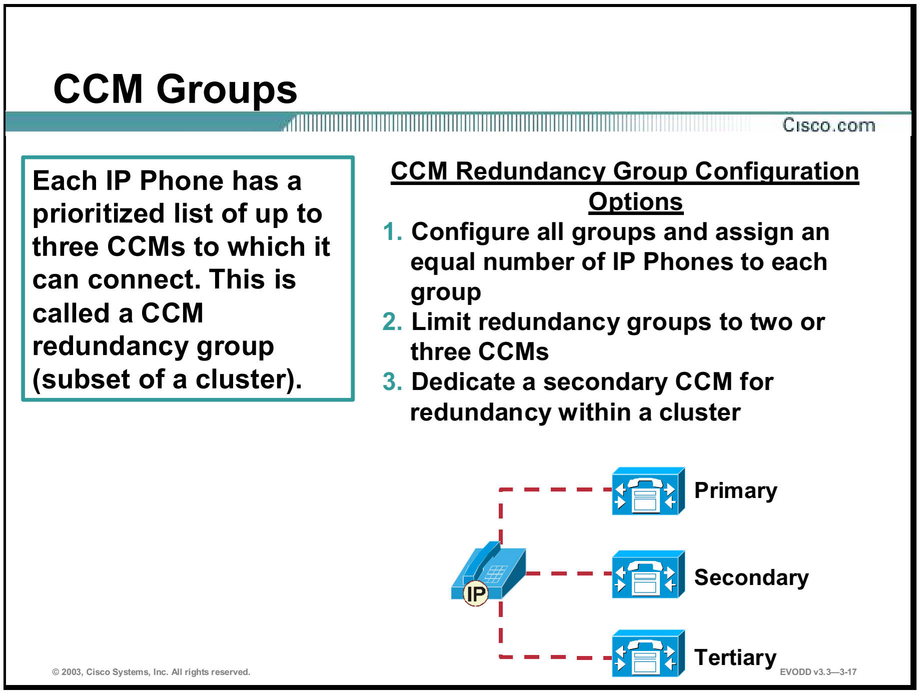
Each cluster must also have a designated TFTP server. Depending on the number of devices a server is supporting, this TFTP server can combine functionality with the publisher or subscriber CCM servers, or you can deploy the TFTP functionality on a separate, standalone server. The TFTP server is responsible for delivering IP Phone configuration files to each telephone, along with media streamed files, such as music on hold (MOH) and ring files, which can cause the TFTP server to experience considerable network and processor load.

Note The maximum number of IP Phones supported by each server or by a CCM cluster can change significantly depending on the other processes running on each MCS server and the version of CCM that you are using. It is important to design your network for scalability and redundancy. For the latest server and cluster load capabilities, visit www.cisco.com/warp/public/cc/pd/nemnsw/callmn/index.shtml.



There are two types of intracluster communication. One type is a mechanism for distributing the database that contains all of the device configuration information. The configuration SQL database is stored on the publisher and replicated to the subscriber members of the cluster. Changes made on the publisher database are communicated to the subscriber databases, ensuring that the configuration is consistent across all of the members of the cluster and facilitating the spatial redundancy of the database.

The second type of intracluster communication is the propagation and replication of run-time data, such as registration of IP Phones, gateways, and digital signal processor (DSP) resources. This information is shared with all of the members of a cluster and assures the optimum routing of calls between members of the cluster and the associated gateways.

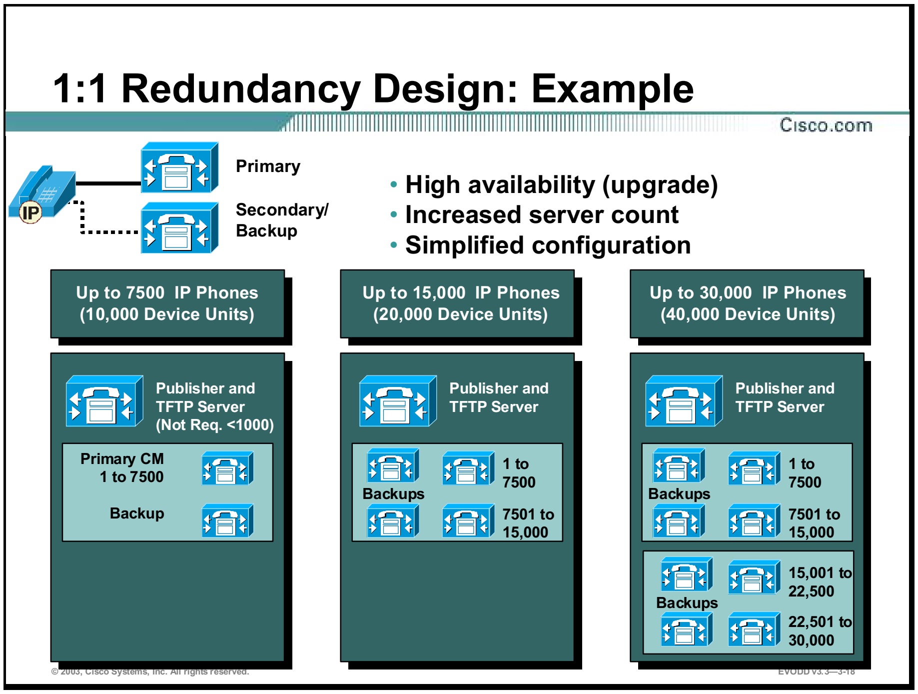


Within a cluster, you can assign a prioritized list of up to three CCMs to each registered IP Phone; it can register to these CCMs for call processing. The assignment of a primary, secondary, and tertiary CCM should ideally distribute failed traffic as widely as possible across other CCMs.

This figure depicts the redundancy scheme using primary, secondary, and tertiary CCM servers. Each IP Phone maintains active Transmission Control Protocol (TCP) sessions with its primary and secondary CCM server, which facilitates an expedited switchover in the event of the failure of the primary. The device reverts to its primary sever when the primary has been restored.

CCM Scalability

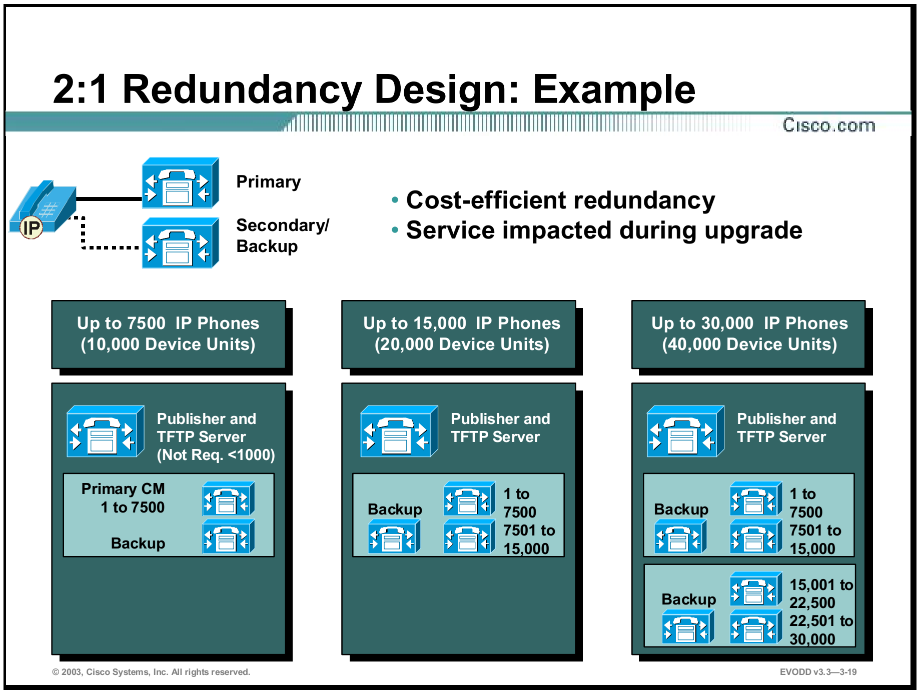
This topic describes Cisco recommendations for cluster configuration based on the amount of users because the CCM is designed to grow with the needs of the client.



When deploying a 1:1 CCM redundancy design, there is a dedicated backup server for each primary server. This design offers the benefit of increased redundancy. If more than one primary server fails, it does not overwhelm the backup servers. Upgrading servers in a 1:1 redundancy design does not affect service, unlike a 2:1 redundancy design.

In the example shown in the figure, you can assume that the MCS 7845 is in use because each CCM server is supporting a maximum of 7500 IP Phones. In the first example (7500 IP Phones), a single CCM is the primary server with a secondary server acting as a dedicated backup.

As the number of IP Phones increases, the number of CCM servers required to support the IP Phones increases in a linear relationship. Some network engineers may consider the 1:1 redundancy design excessive because it is unlikely to lose more than one primary server at a time. Because the possibility of server failure is low and design cost is expensive, some network engineers elect to use a 2:1 redundancy design.

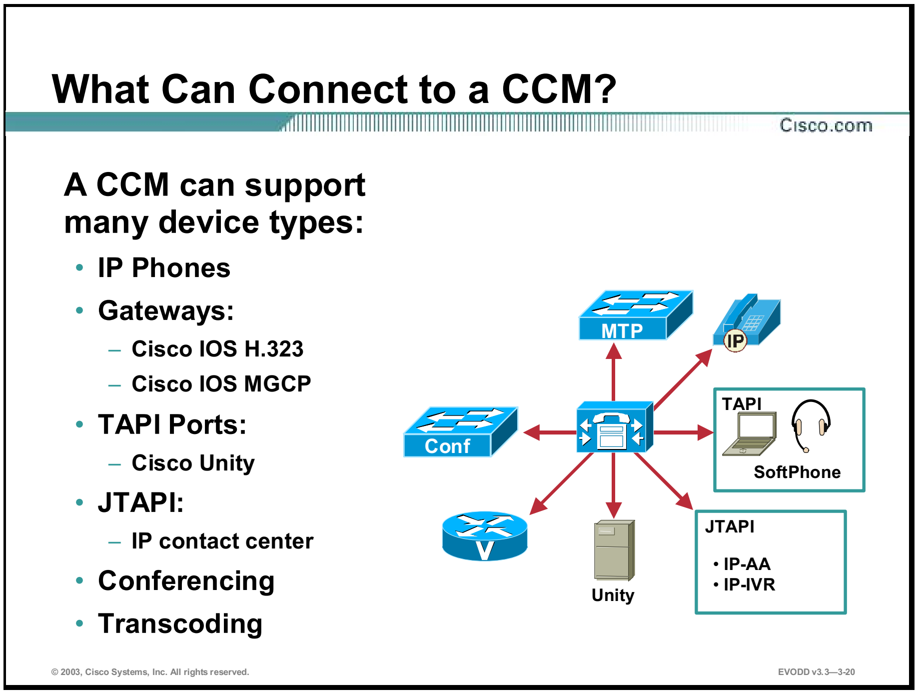


In a 2:1 CCM redundancy design, there is a dedicated backup server for every two primary servers. While this design does offer some redundancy, you risk overwhelming the backup server if multiple primary servers fail. In addition, whenever you perform an upgrade, you must reboot the CCM servers once the upgrade is complete. Consider upgrades and other maintenance when deciding which model to use in your design.

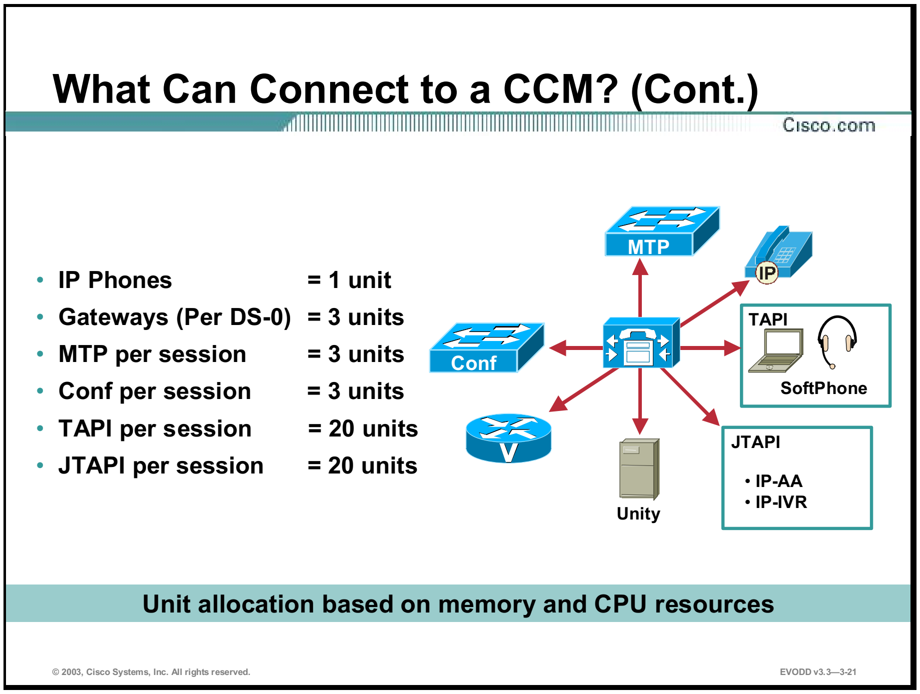
In most IP telephony deployments, network administrators use this 2:1 redundancy model because of the reduced server cost. The MCS 7835 server has redundant hot-swap power supplies and hard drives. It is unlikely that multiple primary servers will fail at the same time if the servers are connected and configured properly, making 2:1 redundancy a viable option for most corporations.

Devices and Weights

This topic discusses the requirements of different devices.



Each device that you add to a CCM environment will affect the performance of that system. Many types of devices can register with a CCM. All of these devices—IP Phones, voice mail ports, TAPI devices, JTAPI devices, gateways, and DSP resources (such as transcoding and conferencing)—consume system resources.



To help calculate the load that a CCM can handle, each CCM platform is rated to by how much weight, or load, it can handle. Each device that the CCM must deal with carries a different weight. Cisco has weight estimates to ensure the proper performance, as seen in Device Weights Table.

Note This link is 3.2 specific, 3.3 has changed considerably and you should get the latest numbers when designing a network. http://www.cisco.com/warp/enterprise/771/srnd/IPtel_srnd.pdf

Table: Base Device Weights

Device Type	Weight per Session on Voice Channel	Session or DSO per Device	Cumulative Device Weight
IP Phone	1	1	1
Analog MGCP ports	3	Varies	3 per DSO
Analog SCCP ports	1	Varies	1 per DSO
CTI route point	2	Varies	Varies[1]
CTI client port	2	1	2
CTI server port	2	1	2
CTI third-party control[2]	3	1	3
CTI agent phone[2]	6	1	6
H.323 client	3	Varies	3 per call
Intercluster trunk gateway	3	Varies	3 per call
H.323 gateway	3	Varies	3 per call
Digital MGCP T1 gateway ports	3	24	72 per T1
Digital MGCP T1 gateway ports	3	30	90 per E1
MoH stream	10	20	200[3]
Transcoding resource	3	Varies	3 per session
MTP software[1]	3	24	72[4]
Conference resource (hardware)	3	Varies	3 per session
Conference resource (software)	3	24	72[4]

[1]Cumulative weight of CTI route point depends on the associated CIT ports used by the application.

[2]Includes the associated IP Phone.

[3]When MoH is installed on the same server as CCM, the maximum number of streams is 20.

[4]When installed on the same server as CCM, the maximum number of sessions is 24.

Note This table is 3.2 specific, 3.3 has changed considerably and you should get the latest numbers when designing a network.

CCM Server Platforms: Device Weights

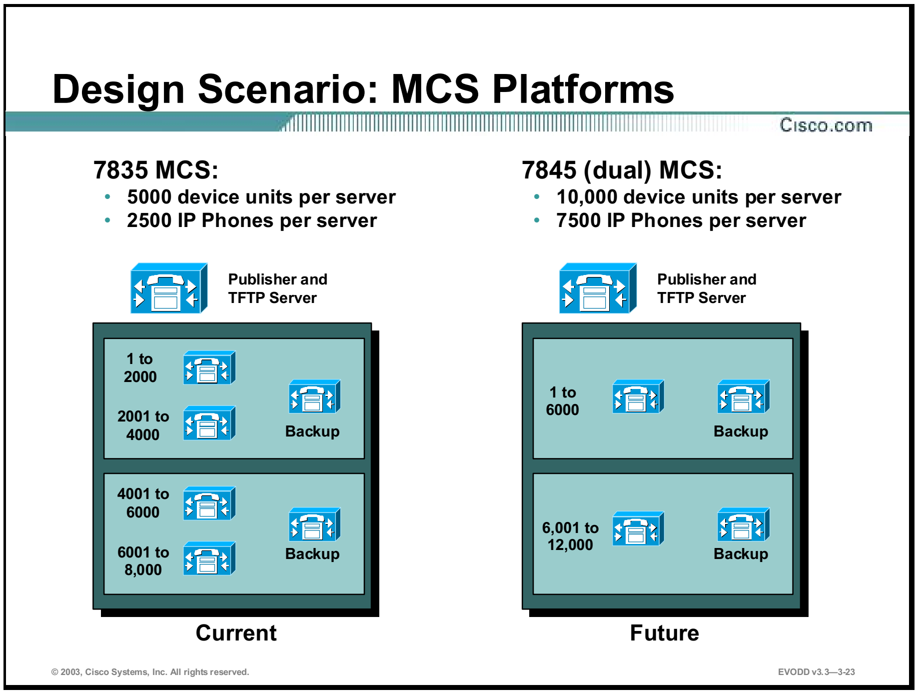
Cisco.com

Platform	Device Units per Server	Maximum IP Phones per Server
MCS 7845-1400 (Dual)	10,000	7500
MCS 7835 All Models	5000	2500
Compaq DL 380	5000	2500
IBM xSeries 342	5000	2500
MCS 7825	2000	1000
SPE 310	2000	1000
Compaq DL 320	2000	1000
IBM xSeries 330	2000	1000
MCS 7815-1000	400	200

EVODD v3.3—3-22

The table in the figure describes the MCS platform capabilities of Cisco IP Phones and device weight units. You can easily calculate the number of Cisco IP Phones that are registered to an MCS platform, but you also need to consider all of the devices that are going to register to an MCS platform. Use the device units from the following table and compare it to the maximum number of device units supported by the MCS platform in the table. These numbers are subject to change. See current MCS documentation on Cisco's web site for up to date specifications. Even though a single IP phone consumes a single device unit, an MCS server does not support as many phones as it has device units, due to the load on the MCS server during peak traffic conditions.

Note These values may change with updated releases of CCM and vary based on the MCS platform. Always check with current Cisco documentation available from Cisco's website for current information.



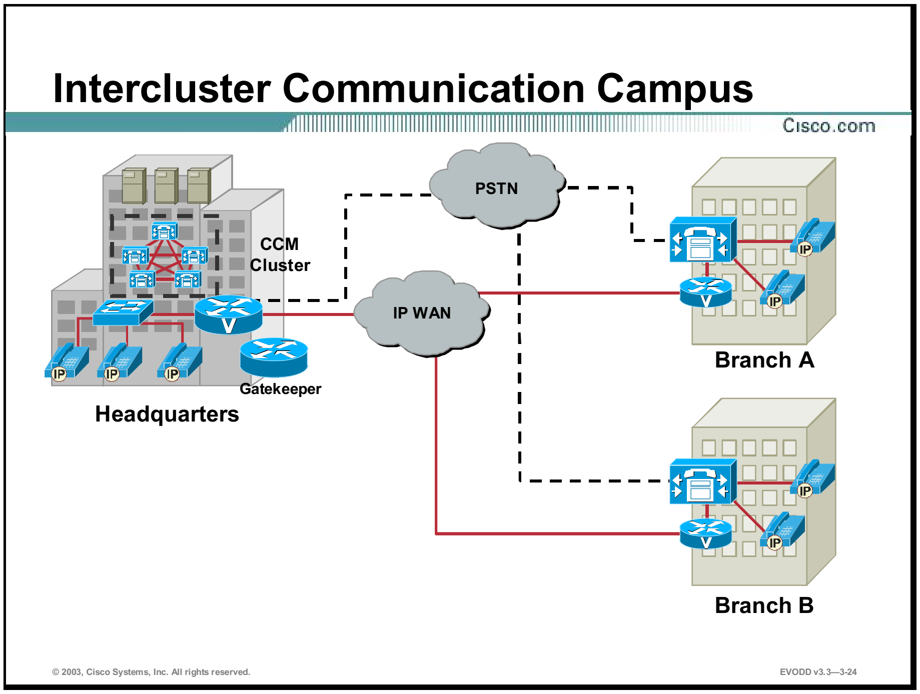
Consider a situation where a customer has an IP telephony solution containing 8000 Cisco IP Phones and a cluster of seven Cisco 7835 MCSs, which are all running CCM 3.1. The customer wants to expand the IP telephony network to include 12,000 IP Phones.

While you can scale CCM 3.1 to support 12,000 IP Phones, using two clusters with an intercluster trunk, migrating the platforms to 7845 dual processor MCS servers, running CCM 3.3, enables a single cluster to support all 12,000 IP Phones.

Specifically, the 7835 MCS platform in the current network supports 2500 IP Phones and 5000 device units per server. The 7845 dual processor MCS supports 7500 IP Phones and 10,000 device units per server. Therefore, without increasing the number of servers in the cluster, you can double the capacity of the cluster by replacing the MCS platforms. Note that this example uses a 2:1 redundancy model. You can obtain additional redundancy by adding two CCM servers to form a 1:1 redundancy model.

Intercluster Communication

This topic describes intercluster communications and addresses issues in provisioning clusters for campus deployments and multisite WAN deployments.



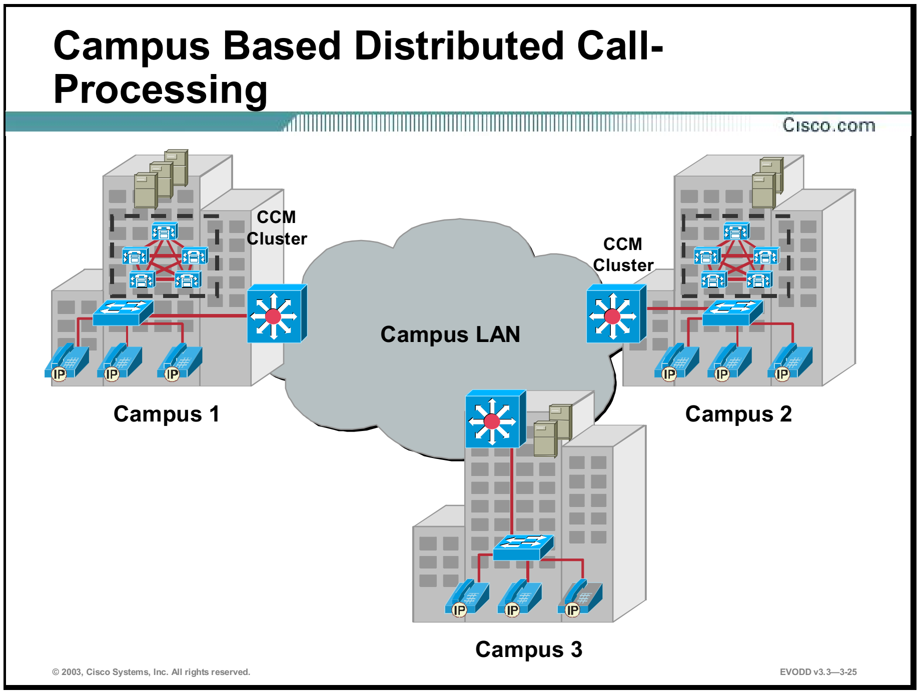
The figure illustrates a typical distributed call processing deployment with CCM and Cisco Voice over IP (VoIP).

CCM supports three trunks for intercluster and H.323 communications, the H.225 gatekeeper controlled trunk, the Intercluster trunk, and Intercluster gatekeeper controlled trunk.

The H.225 gatekeeper controlled trunk allows CCM to communicate with CCMs and H.323 devices registered to an H.323 gatekeeper. The H.225 gatekeeper controlled trunk is not recommended in a pure CCM environment.

- The gatekeeper is configured in each CCM cluster
- The H.225 gatekeeper controlled trunk is configured in the CCM cluster
- Each CCM in a cluster registers an H.225 gatekeeper controlled trunk with the gatekeeper
- Calls are load balanced across the registered trunks in the CCM cluster
- Multiple gatekeepers and trunks are supported in CCM
- The gatekeeper has a gatekeeper/zone configured for each site supporting CCM or voice gateways

- The interzone bandwidth command on the gatekeeper is used to control bandwidth between CCM clusters and H.323 devices registered directly with the gatekeeper
- Up to 100 clusters are supported on a single Cisco Multimedia Conference Manager (MCM) gatekeeper

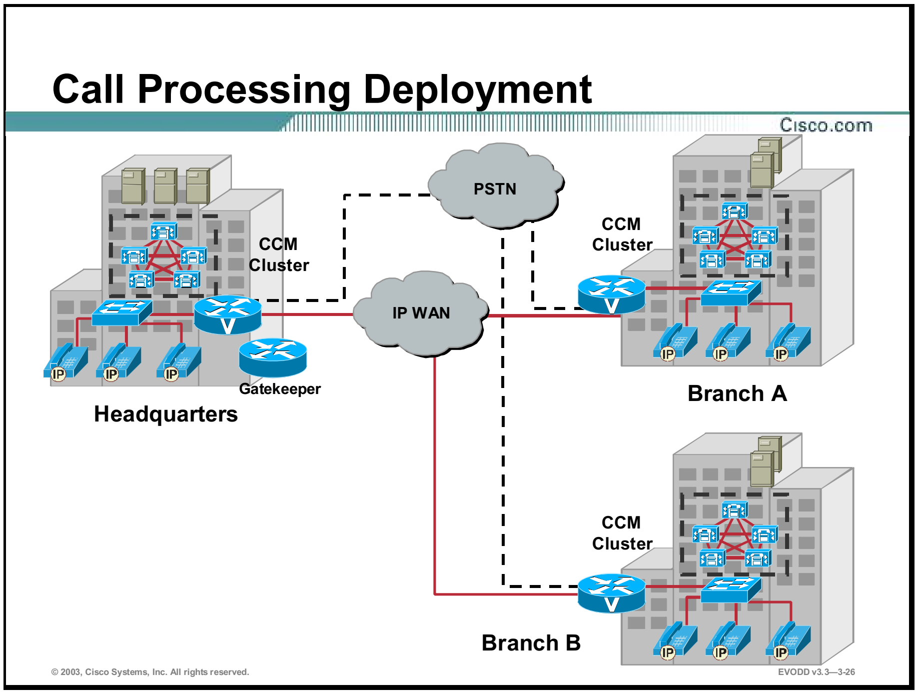


The figure illustrates a campus based distributed call-processing environment. Multiple CCM clusters are configured on the campus and communicate over intercluster trunks.

The Intercluster trunk allows CCM to CCM communication. There is no Call Admission Control available between clusters when using the intercluster trunk and should only be used on high speed LANs or MANs with plenty of available bandwidth

An intercluster trunk is configured on each cluster

The figure illustrates a campus based distributed call-processing environment. Multiple CCM clusters are configured on the campus and communicate over intercluster trunks.



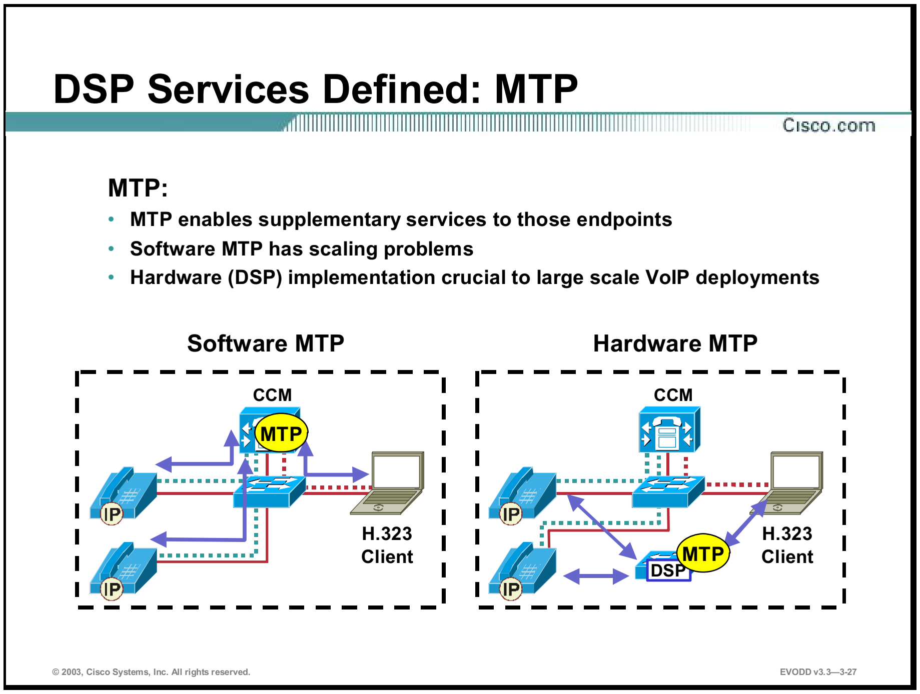
The figure illustrates a typical distributed call processing deployment. Each site contains a CCM cluster registered to a gatekeeper providing CAC between clusters.

The Intercluster gatekeeper controlled trunk allows CCM to communicate with other CCMs registered to an H.323 gatekeeper. The intercluster gatekeeper controlled trunk is recommended in distributed call processing environments.

- The gatekeeper is configured in each CCM cluster
- The intercluster gatekeeper controlled trunk is configured in each CCM cluster
- Each CCM in a cluster is register an intercluster gatekeeper controlled trunk with the gatekeeper
- Calls are load balanced across the registered trunks in the CCM cluster
- Multiple gatekeepers and trunks are supported in CCM
- The gatekeeper has a gatekeeper/zone configured for each cluster
- The interzone bandwidth command on the gatekeeper is used to control bandwidth between CCM clusters and H.323 devices registered directly with the gatekeeper
- Up to 100 clusters are supported on a single Cisco Multimedia Conference Manager (MCM) gatekeeper

DSP Services

This topic describes DSP services.

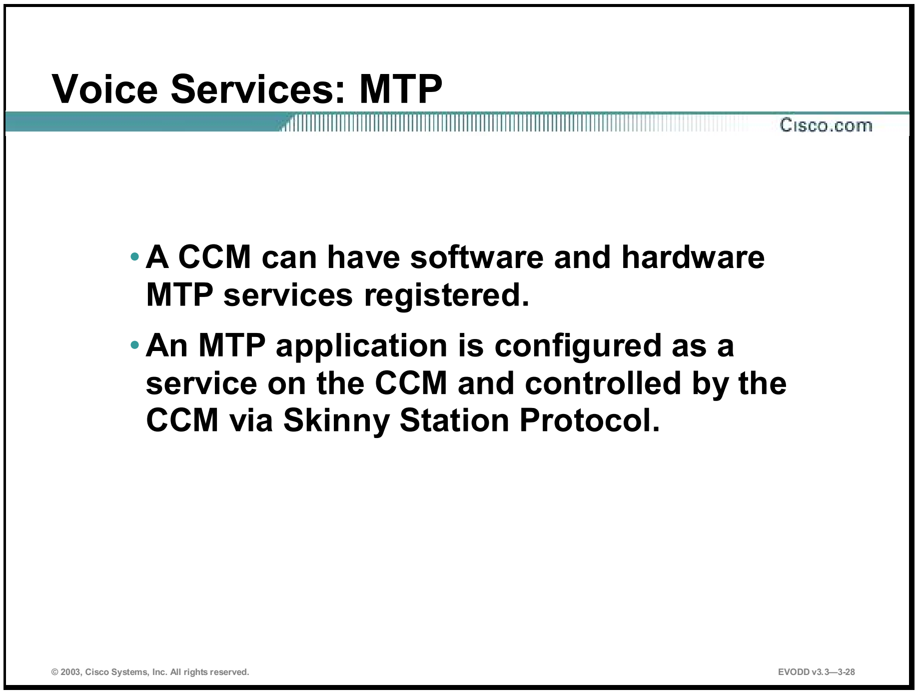


It is important to understand the various uses that clients have for DSPs to ensure that your design meets their needs.

One type of DSP service is an MTP. You can implement MTPs with a CCM through either software or hardware.

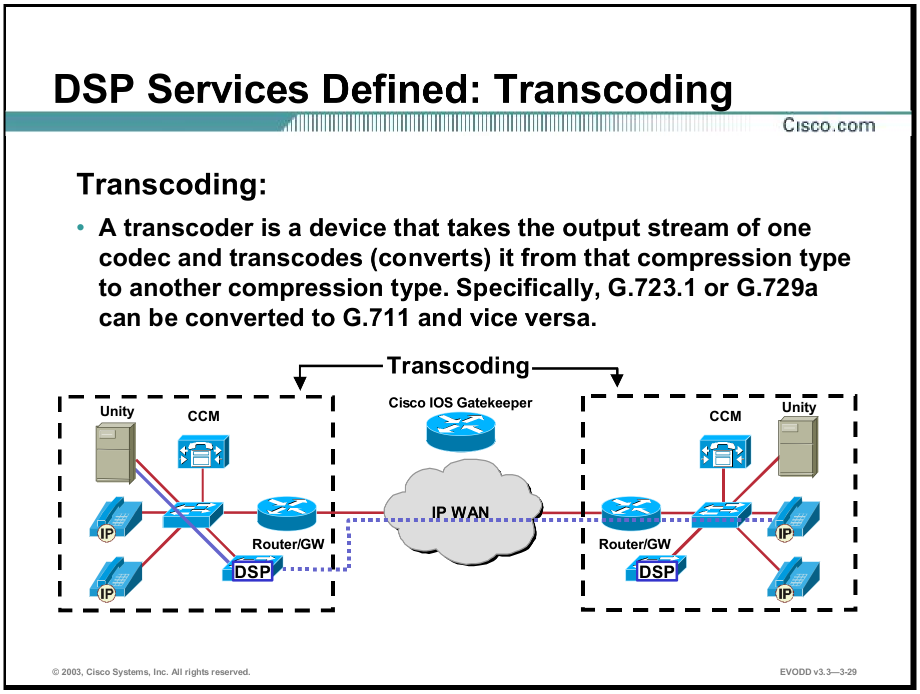
An MTP is invoked on behalf of clients that do not support the H.323v2 features of OpenLogicalChannel and CloseLogicalChannel with the EmptyCapabilitiesSet. MTP provides the call management functions needed for supplementary call services (such as call hold, conference bridge, and so forth) with H.323 clients. With software MTP, the CCM terminates the voice bearer channel for the inbound voice stream and establishes a new voice bearer channel with the endpoint, or eventual endpoint, which results in both endpoints of a call terminating to the CCM MTP process. This solution does not scale well and only works with G.711.

The hardware version uses DSPs on a Catalyst blade that offloads processing from the CCM server.



A CCM can have either the hardware or software MTP associated with it. When you configure software MTP as a kernel mode driver, it can run as a process in CCM.

Note Cisco recommends hardware DSPs.



Transcoding is the ability to convert a voice stream from one encoding format to another in real time through a DSP resource. For example, a transcoder can convert a G.711 stream to G.729a or to G.723.1 and send it across the WAN as a compressed stream to an IP Phone at the far end. At the same time, a transcoder can convert a far end compressed voice stream from the remote branch to G.711 to communicate with a Cisco Unity voice server that supports only G.711.

Note The only supported transcoding is low-bit rate to high-bit rate, and high-bit rate to low-bit rate. Currently, there is no support for transcoding G.723.1 (low-bit rate format) to G.729a (low-bit rate format).

Voice Services: Transcoding

Cisco.com

- Transcoding configured as a service on the CCM and controlled by the CCM via Skinny Station Protocol
- Capacity of hardware transcoding on the DSP resource modules:
  - Catalyst 4000 Transcoding sessions = 24
  - Catalyst 6000 Transcoding Sessions = 32/256
- Transcoding available only with DSPs

EVODD v3.3—3-30

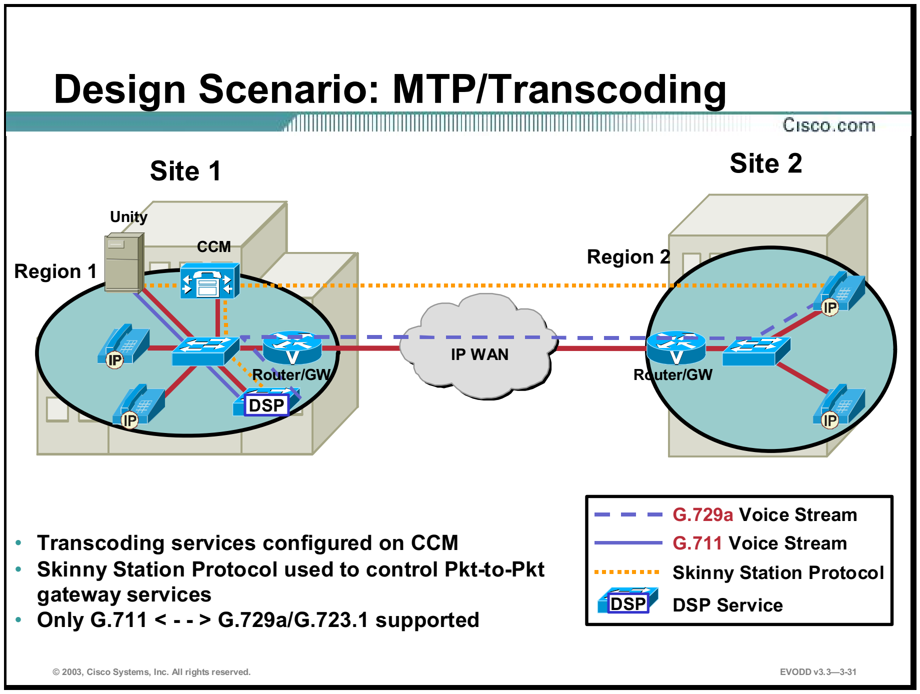
You can configure transcoding as a service on the CCM to implement Skinny Station Protocol for both the Catalyst 4000 and 6500. You then configure the service using the CCM interface.

Transcoding is only available in the DSP resource modules. The Catalyst 4000 provides support for a maximum of 24 sessions per module. The Catalyst 6500 provides support for 32 sessions per port and 256 transcoding sessions per module.

Reference For current scalability limitations, see the following links:

Catalyst 4000: http://cisco.com/en/US/products/hw/switches/ps663/products_data_sheets_list.html

Catalyst 6500: http://cisco.com/en/US/products/hw/switches/ps708/products_data_sheets_list.html



This figure provides an example of transcoding. In the example, an IP Phone at a remote site is transferred to voice mail when the called party does not answer. The calling IP Phone establishes a connection to the centralized CCM, and then the CCM establishes a control channel to the DSP resource modules and the Unity server. The compressed call is connected to the DSP resource and is then transcoded to G.711 to enter the Unity voice-mail system.

The CCM recognizes a compressed voice stream by the call region. The remote IP Phones are in Site 2, and the central IP Phone and Unity are in Site 1. The CCM recognizes that all Unity calls are in G.711 and that all calls from Site 2 are in G.729 and signals the DSP resource to prepare for a transcoding session. The CCM chooses a DSP and sends the G.729 call to it for the conversion to G.711.

There are three call setups: the first call setup with the remote telephone, the second call setup on the Unity server, and the third call setup from the CCM to the DSP farm telling the DSP to establish a transcoding session.

The Catalyst 6500 DSP blade uses MGCP to interact with the CCM and Skinny Station Protocol for DSP resource management.

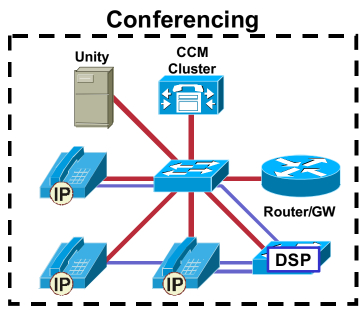
The Catalyst 4000 is an H.323v2 Cisco IOS software-based Gateway; therefore, it uses H.323v2 for all PSTN Gateway functions. For MTP, conferencing, and transcoding functions, it uses MGCP to interact with CCM. The Catalyst 6500 also uses MGCP for DSP resource management.

DSP Services Defined: Conferencing

Cisco.com

Conferencing:

- **Ad-Hoc:**
  - User presses conf button; 1st caller placed on hold; user receives dial tone and dials a second user; presses conf again and all users are now connected on the conference bridge
- **Meet-Me:**
  - Conference controller presses Meet-Me button; receives dial tone and dials conference call number; all conference call attendees call conf call number and are connected on the conference bridge



 EVODD v3.3—3-32

There are two types of conferences: Ad-Hoc and Meet-Me. Ad-Hoc conferencing is used when two people decide to conference in a third call attendee. The person contacting the third person places the second person on hold, calls the third person, and when they are talking, brings the second person back onto the call. This process is a three-person conference.

Meet-Me conferencing refers to all attendees calling into a certain conference bridge number and staying on the call until the end of the meeting.

At any given moment, a conference has a certain number of active attendees—individuals who are speaking and a certain number of inactive attendees—individuals who are listening. The capacity of conference calling usually stipulates the maximum number of active attendees. A conference call can support a large number of inactive attendees; however, if the active attendee number exceeds the capacity, some attendees will not be heard.

Voice Services: Software Conferencing

Cisco.com

Software conferencing:

- **On the CCM server:**
  - G.711 only – MTP/transcoding can be used for compressed sessions
  - Maximum number of conferencing sessions = 24
  - Maximum number of conferences = 8 (with recommended max of 3 attendees each)
  - Maximum conference size = 24 attendees
- **Large conference size (24 attendees vs. max or 6 for DSP-based conferencing) is needed primarily for pre-scheduled Meet-Me calls**
- **Cisco anticipates eco-system partners to develop solutions for large conference size**

 EVODD v3.3—3-33

Software conferencing is only supported on the CCM Windows 2000 server version. It only works with G.711, but MTP transcoding can convert compressed calls to G.711.

The maximum number of sessions using software conferencing on the CCM is 24. When adding these extra services to a CCM, consider the CPU resources. CPU limits are the primary reason why there are hardware conferencing services.

Voice Services: Hardware Conferencing

Cisco.com

Hardware Conferencing on Catalyst DSP Modules:

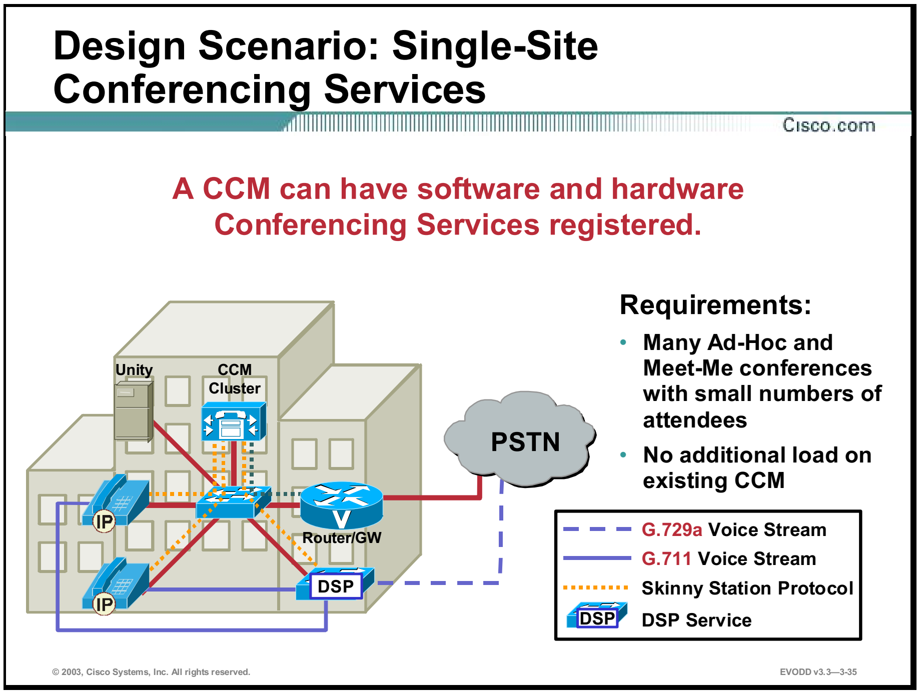
- **On the Catalyst 4000 – WS-X4604-GWY:**
  - G.711 only – MTP/transcoding can be used for compressed sessions
  - Maximum conferencing sessions = 24
- **On the Catalyst 6500 – WS-X6608-x1:**
  - G.711 – 32 concurrent conference
  - G.723 – 32 concurrent conference
  - G.729 – 24 concurrent conference

EVODD v3.3—3-34

The Catalyst 4000 only uses G.711 for conferencing; however, because these blades are MTP transcoding-enabled, that functionality can support compressed voice conferences.

On the Catalyst 4000, 24 is the maximum number of conferencing sessions, or calls. The maximum number of conferences is eight, with up to three attendees on each of those eight conferences. The maximum conference size is six attendees on any particular conference.

The Catalyst 6500 DSP blade supports G.711, G.723, and G.729; and can use transcoding to conference the compressed calls. The maximum number of conferencing calls, or sessions, is 32 per port, if using the G.711/G.723 CODEC or 24 per port if using G7.29.



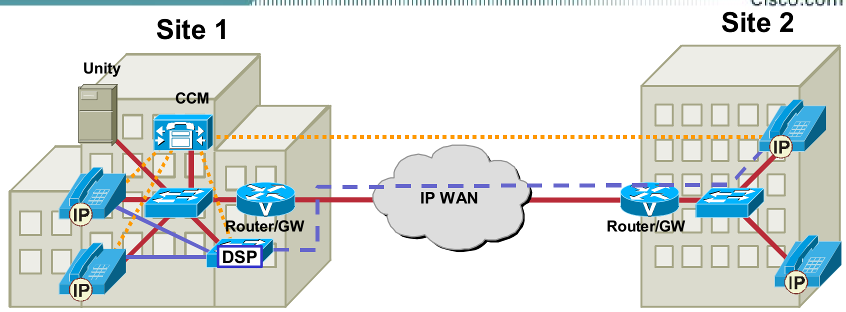
Single-site conferencing is targeted for environments that require a large amount of Ad-Hoc conferencing. Meet-Me conferences tend to have larger numbers of participants than Ad-Hoc conferences.

This design scenario shows an example of the location of the DSP modules from a conferencing point of view—single-site, multiple telephones (but still a small site), and PSTN conferencing. The requirements for this design are many small Ad-Hoc and Meet-Me conferences, small numbers of attendees per conference, no additional NT/2000 servers, and no additional load on existing CCM.

This solution also supports PSTN callers. The PSTN caller is connected directly to the conferencing service (which runs on the DSP farm) and each of the IP Phones selected for conferencing.

With a Catalyst 6500, the call comes in to the WS-6608-X1 gateway from the IP network. The gateway sends the call over to the conferencing port; it converts all of the G.711 streams to analog voice, which are summed together. The resulting stream is then converted back to G.711 and streamed out to the IP network and to the PSTN.

Design Scenario: Multisite Conferencing Services



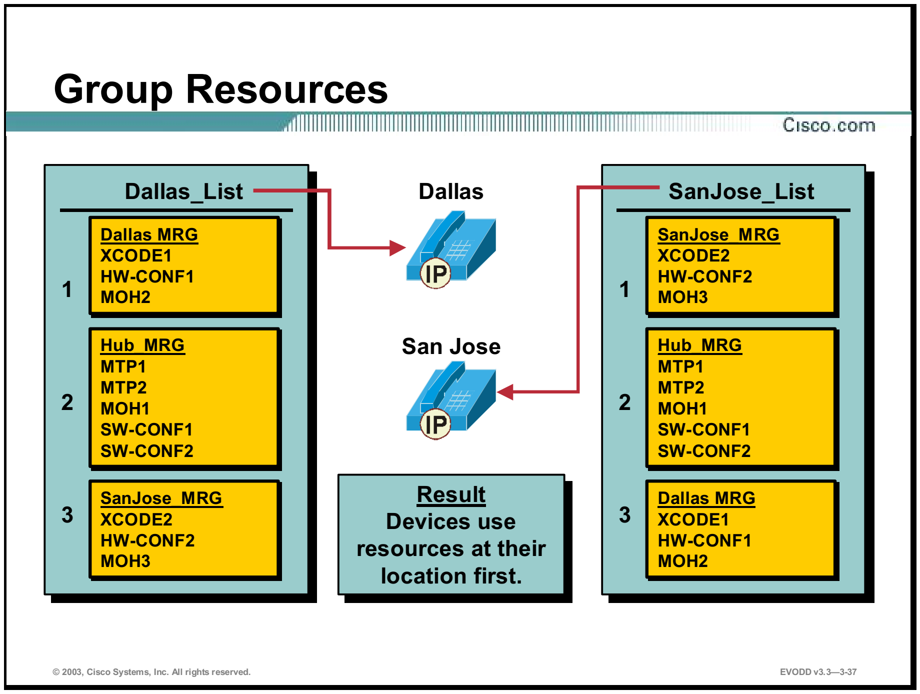
- Callers at central site conference user at remote site
- Suitable solution for many Ad-Hoc and Meet-Me conferences with small numbers of attendees
- Places no additional load on existing CCM
- Multisite attendees (using voice compression service)

G.729a Voice Stream
G.711 Voice Stream
Skinny Station Protocol
DSP Service

 EVODD v3.3—3-36

This figure shows an example of multisite conferencing where transcoding and conferencing services are used together. Many small Ad-Hoc and Meet-Me conferences, no additional NT servers, no additional CCMs, and no additional load on the CCM are requirements for this design. This design requires a DSP hardware resource. This scenario requires voice compression because the attendees are located at multiple sites.

In this case, Site 1 signals the CCM at Site 2, sets up the call, and the call is forwarded to each of the appropriate resources—two IP Phones and the DSP conferencing resource. The conference call is set up using compressed voice from the remote site to the DSP resource, and uncompressed voice from the local site. Because the conferencing hardware in this scenario only supports G.711, a G.729a voice stream must be transcoded to G.711 before that call can participate in the conference.



With the advent of media resource groups, resources may be shared among devices that are registered to different CCMs. Media resource groups are a major advancement because they allow you to assign resources at the device level, instead of the CCM level, which provides a pool of resources.

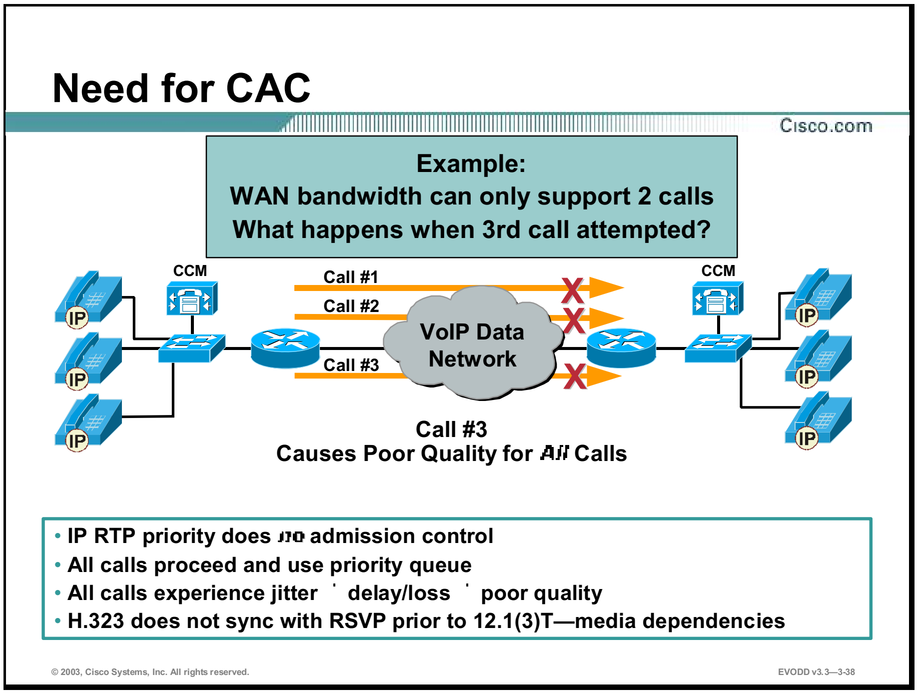
Note On the Catalyst 6500, conferencing and MTP transcoding services cannot extend over DSPs or ports.

You should learn the conference call patterns. You can learn the patterns by talking to the telecom manager, recording the number of calls, and tracking their patterns across the WAN. You should also note provisions, maximum conference call size, and other characteristics.

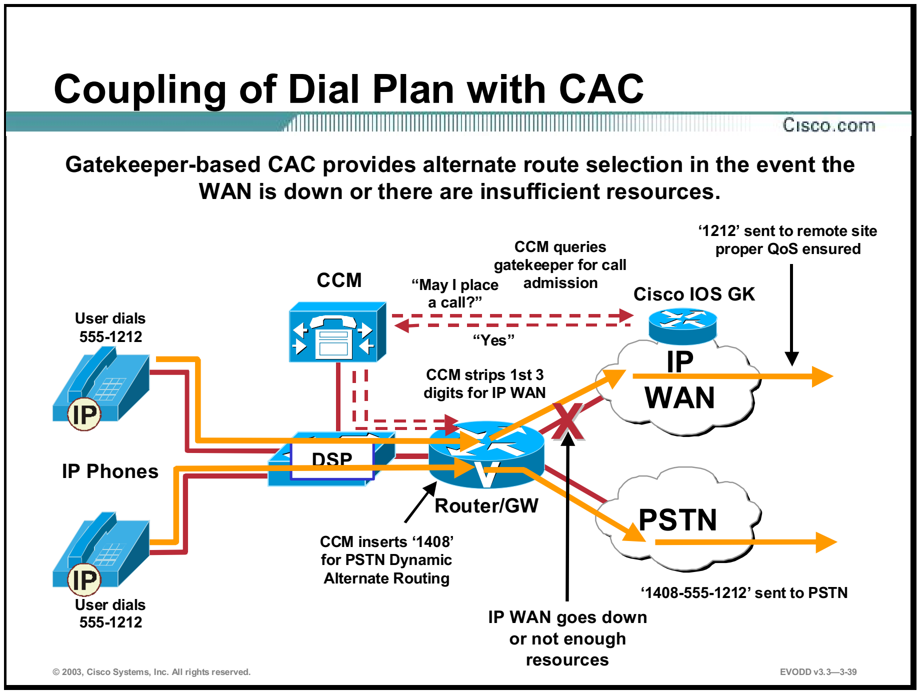
Note The central site must have the MTP and transcoding resources with the CCM cluster because you cannot configure them in a location for CAC.

CAC

This topic discusses CAC. You can implement CAC to ensure that you do not oversubscribe bandwidth and to establish voice quality.

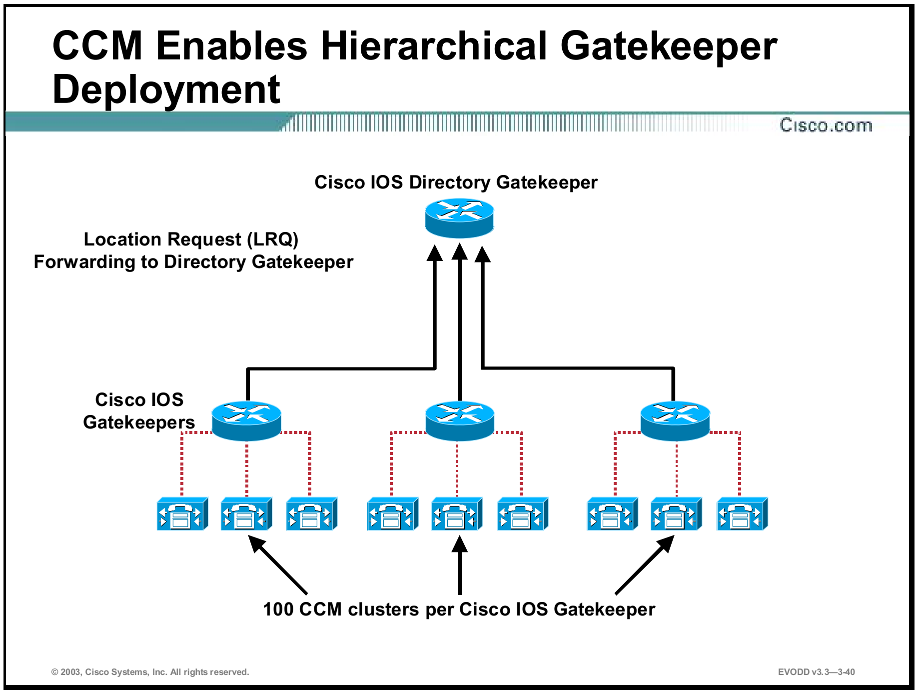


QoS tools ensure voice quality in two ways: by giving voice priority over data, and by preventing voice from oversubscribing a given WAN link. You accomplish the second task by using CAC mechanisms. The need for CAC in Cisco IP telephony networks is amplified by the fact that all IP Phones have an open IP path to the WAN. Toll bypass networks, in contrast, can limit the number of physical trunks eligible to initiate calls across the WAN.



As shown in the figure here, the dial plan is now coupled with CAC. When a remote site requests a call, the CCM queries the gatekeeper for permission to send the call over the data network (call admission). If sufficient bandwidth is available, the Cisco IOS gatekeeper gives permission to establish the call. CCM adds or removes digits, as required, to route the call over IP WAN.

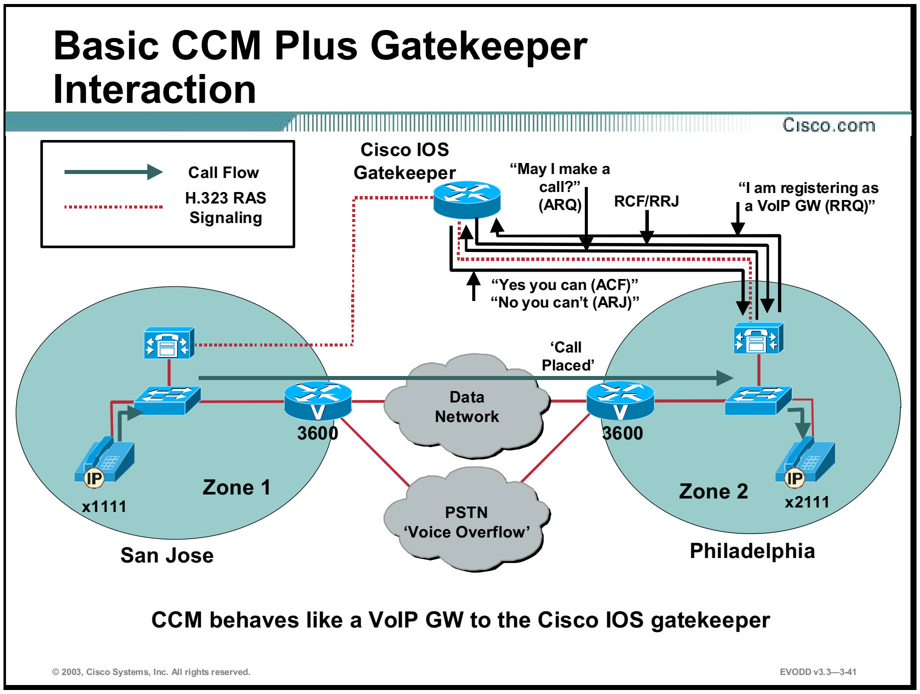
If there is insufficient WAN bandwidth, then the gatekeeper does not grant permission for the call, and CCM adds digits to send the call over the PSTN.



CCM can use a hierarchical deployment of gatekeepers with 100+ CCMs per gatekeeper.

Remote sites can have 100 CCMs controlled by a single gatekeeper. A directory function at the next hierarchical level stores information about all of the other gatekeepers.

Note	This deployment provides significant improvements in scalability.



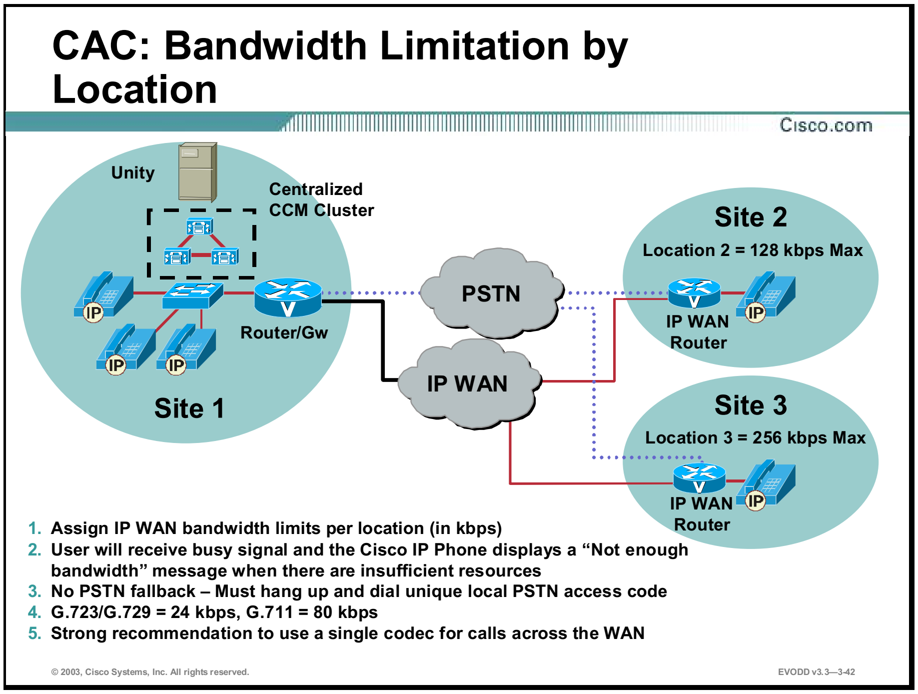
These four steps describe basic gatekeeper interaction:

Step 1 CCM registers as a gateway with gatekeeper.

Step 2 The gatekeeper responds with a registration confirmation (RCF) or a registration rejection (RRJ).

Step 3 CCM sends an admission request to the gatekeeper when it wishes to initiate a call.

Step 4 The gatekeeper responds with an admission confirmation (confirm) if bandwidth is available, or an admission rejection (reject) if no bandwidth is available.



Each device connected to the network is associated with a known location called Site 1 or Site 2 (as in the slide) or named by zip code, city, or other designation.

Each location has an associated bandwidth bucket allocated to it. When this bandwidth resource pool is depleted, no device from that location can be admitted to the network. In the slide, Site 2 has 128 kbps assigned to it, and Site 3 has 256 kbps assigned to it.

Note This example uses arbitrary numbers for demonstration. You can make the assignments based on need and availability.

From the point of view of IP Phone mobility, this topology-dependent scheme has major ramifications. If an IP Phone is connected at a different location from where it is registered when it places a call, it causes bandwidth to be decremented from its home location resource pool even though it is not using this bandwidth. The system omits decrementing bandwidth from the resource pool that the IP Phone is using, which defeats the resource allocation scheme.

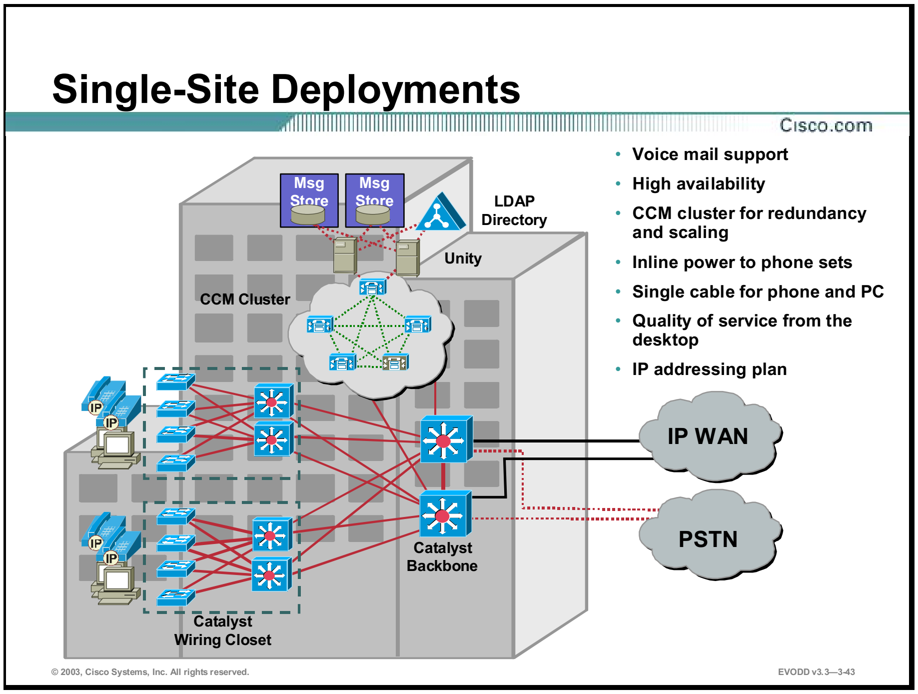
Alternative routing over the PSTN is not available with location-based bandwidth control. If users try to place a call and bandwidth is not available, they receive a busy signal and have to hang up and dial the PSTN access code. CCM can display a configurable message on the IP Phone, as well as returning the busy tone.

There are reasons to select a single codec for all wide-area calls. A single codec makes the bandwidth requirement per call predictable so that CAC can safeguard the system from oversubscription. If calls in both G.711 and G.729 formats are allowed across the IP WAN, it is possible to oversubscribe the WAN. For example, if 28 kbps remains in the bit bucket, a G.711 IP Phone can be granted admission and exceed the budget. In a large WAN, with many IP

Phones and the possibility of applying voice activity detection (VAD), oversubscription may not cause trouble, but in smaller environments, it can lead to serious call problems.

CCM Deployments

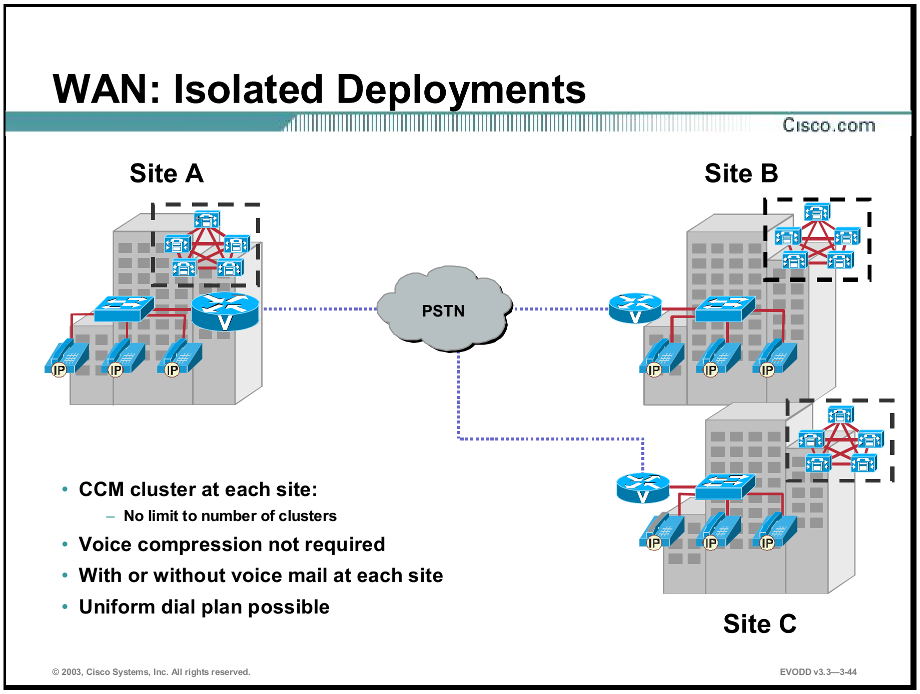
This topic discusses CCM deployments.



Depending on the needs of the client, you can deploy CCM in one of four ways: single-site, WAN isolated, WAN distributed call processing, and WAN centralized call processing. The number of users and network and geographical topology of the client will determine the proper deployment type.

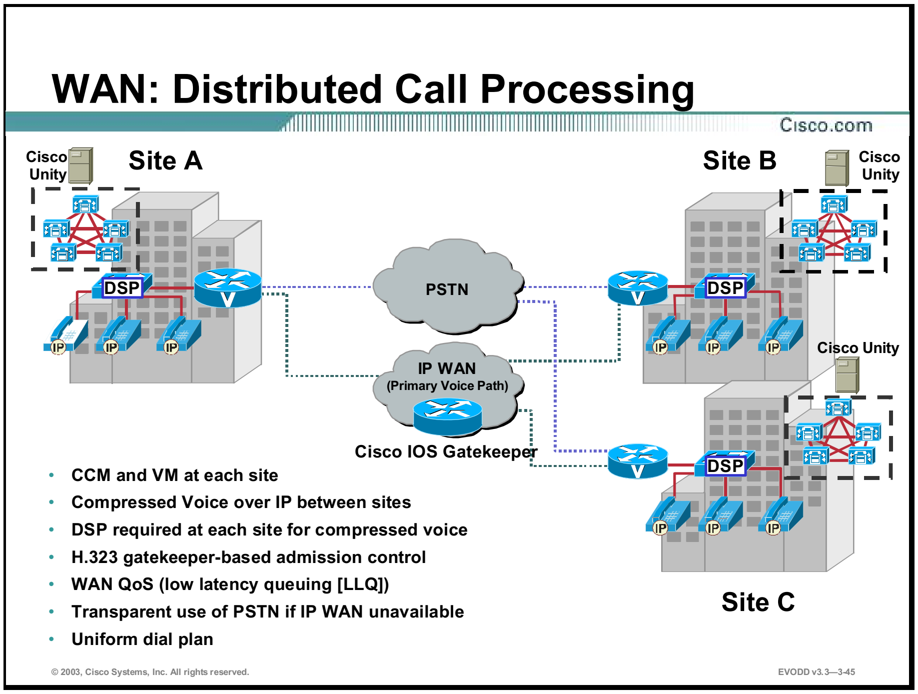
You use a single-site deployment shown here if the client has one location and does not have IP Phones across the WAN connection. This deployment includes the following:

- Voice-mail support from Cisco or a third party
- Switched network design for high availability
- CCM cluster for redundancy and scaling
- Inline power to IP Phone sets
- Single cable for IP Phone and PC
- QoS from the desktop
- IP addressing plan



A WAN isolated deployment is similar to the single-site deployment, as each site is a single site deployment. This deployment includes the following:

- CCM cluster at each site
- No limit to the number of clusters
- Voice compression not required
- With or without voice mail at each site
- Possible uniform dial plan

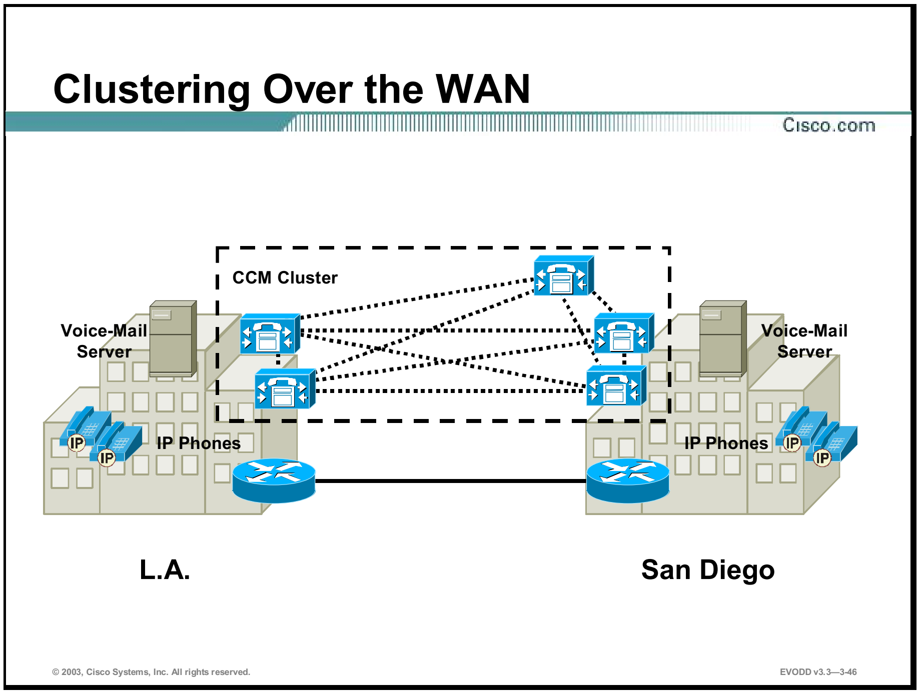


You use the WAN distributed call processing deployment when the client has multiple locations, wishes to install a CCM at each location, and wishes to use the WAN to carry inter-site calls. This deployment includes the following:

- CCM and voice mail at each site
- Compressed VoIP between sites
- DSP required at each site for compressed voice
- H.323 gatekeeper-based CAC
- WAN QoS (low latency queuing [LLQ])
- Transparent use of PSTN if IP WAN unavailable
- Uniform dial plan

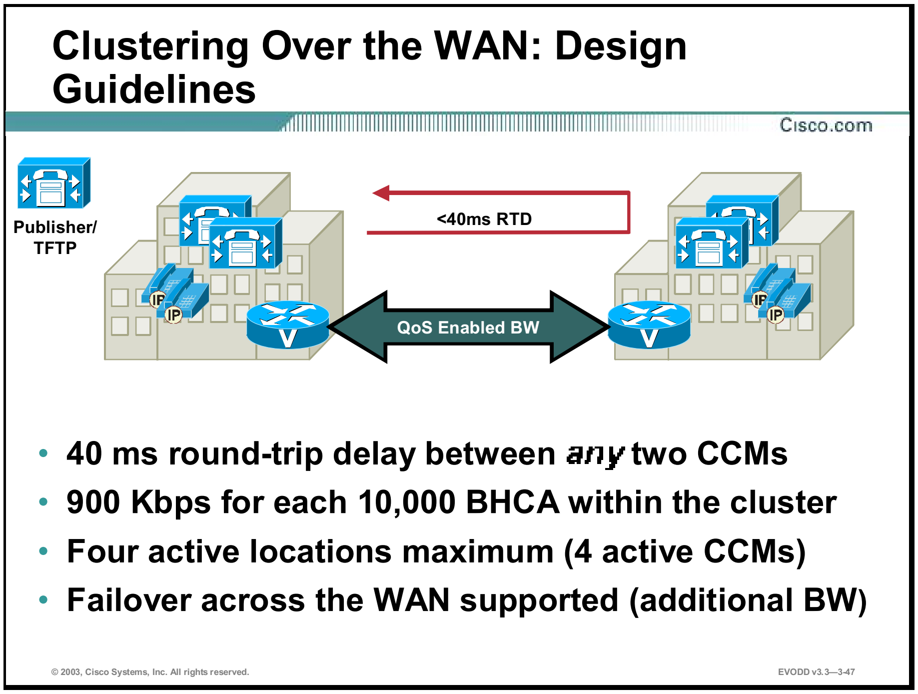
Voice calls between sites use the IP WAN as the primary path and the PSTN as the secondary path, in the event the IP WAN is down or has insufficient resources to handle additional calls.

The primary advantage of this deployment model is that using local call processing provides the same level of features and capabilities, whether the IP WAN is available or not.



Cisco now supports CCM clusters over the WAN. While there are stringent requirements for this design, it offers the advantage of a unified dial plan and an extension of all features to all offices in the IP telephony network. Clustering over the WAN is useful for customers that require further functionality than the limited feature set offered by Survivable Remote Site Telephony (SRST).

This network design also allows the remote offices to support more IP Phones than SRST if the connection to the primary CCM is lost. Because all CCMs are part of the same cluster, you now have the benefit of a single point of administration, which offers an advantage over the multicluster distributed design, where each location maintains its own CCM database and configuration.



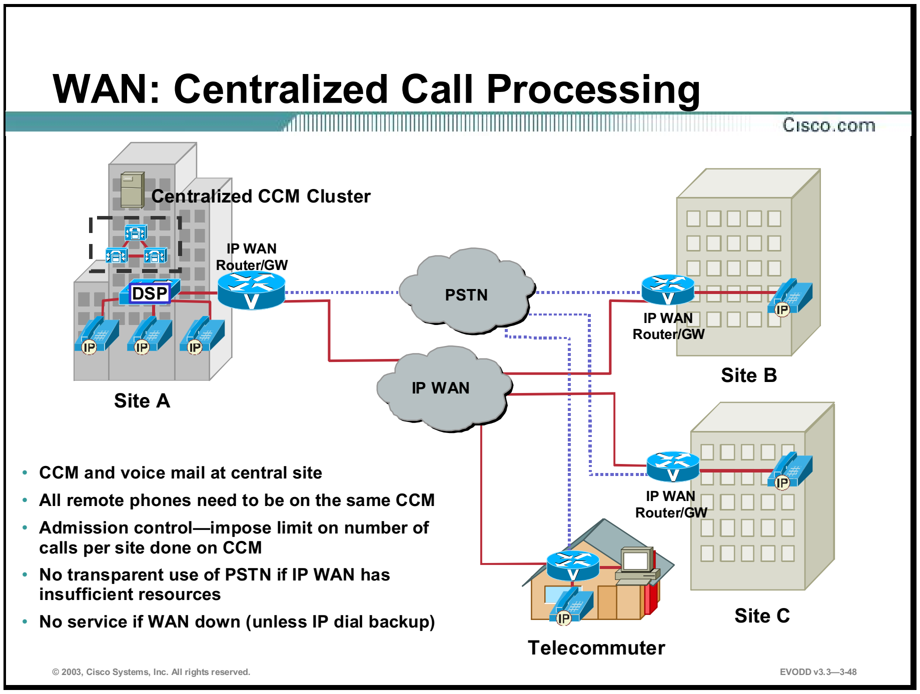
While the distributed single-cluster call processing model offers some significant advantages, it must adhere to the following design guidelines:

- The maximum round-trip delay is 40ms between any two CCMs in the cluster. The maximum one-way delay between any two CCMs is 20 ms. In comparison, high quality voice guidelines dictate that one-way delay should not exceed 150 ms (300 ms for a round trip). Because of this strict guideline, you can only use this design between high-speed, closely connected locations.

- For each 10,000 busy hour call attempts (BHCA) within the cluster, you need to support an additional 900 kbps of WAN bandwidth for intracluster run-time communication.

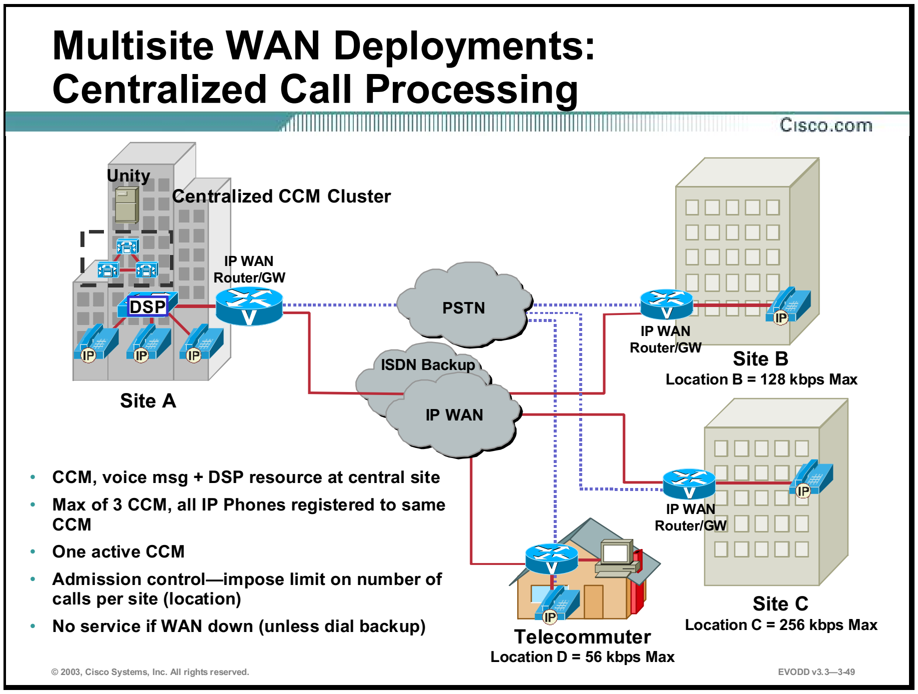
- The distributed single-cluster design supports a maximum of four primary CCMs and correlates directly to the maximum number of supported locations.

- SRST design is not necessary in this model because the IP Phones can failover across the WAN to other CCM servers. Failover can cause significant additional bandwidth requirements, depending on the number of IP Phones at each location.



You use the WAN centralized call processing deployment when the client has multiple locations but wishes to have CCM at one location. This deployment includes the following:

- CCM and voice mail at central site
- All remote IP Phones do not have to reside on the same CCM
- CAC
- No transparent use of PSTN if IP WAN has insufficient resources
- No service if WAN is down (unless there is an IP dial backup; for example, an ISDN backup or SRST)



The model for centralized call processing is shown here. All CCMs are located at a hub or aggregation site, without call processing at the branch locations. You can also centralize voice-messaging equipment, along with DSP resources. It reduces equipment requirements for the branch locations and makes it possible to provision these services at small branch offices. It also centralizes system administration, which reduces cost.

All phones in a single remote location should register on the same CCM, referred to as the active CCM, because of CAC. CCM is configured with an allocation of bandwidth that it decreases each time an additional call is connected. CCM, therefore, has to recognize the IP Phones that can consume its allocated bandwidth. This CAC approach uses the concept of locations.

Before SRST, call processing ended if the IP WAN went down, unless there was a dial backup, such as the backup ISDN WAN shown here. If the IP WAN went down, CCM was unaware of the change. However, that also meant that the backup must have the same bandwidth as the primary CCM, because CCM would continue to send the same number of bits through the pipe.

Using SRST and sending all of the calls out on the local PSTN until the WAN comes back up is a better solution; however, it has limitations, because it has a smaller feature set than CCM.

Reference	For the specific number of telephones supported on various Cisco routers running SRST, see the following link: http://www.cisco.com/univercd/cc/td/doc/product/access/ip_ph/srs/fallbak2.htm#57115

Summary

This topic summarizes the key points you learned in this lesson.

Summary

Cisco.com

- Cost efficiency, flexibility, and better technology are the main driving forces behind the demand for the Cisco IP telephony solution.
- Creating a well-designed CCM cluster is a key element to consider when you initially design an IP telephony network.
- CCM scalability grows with the needs of the client.
- Each device that you add to a CCM environment will affect the performance of that system.
- Communication between clusters is achieved by using standards-based, H.323 signaling.
- It is important to understand the various uses that clients have for DSPs to ensure that your design meets their needs.
- You can implement CAC to ensure that you do not oversubscribe bandwidth and to establish voice quality.
- Depending on the needs of the client, you can deploy CCM in one of four ways: single-site, WAN isolated, WAN distributed call processing, and WAN centralized call processing.

 EVODD v3.3—3-50

Next Steps

After completing this lesson, go to:

- Gateway Types lesson

References

For additional information, refer to these resources:

- CCM Clusters:
  http://www.cisco.com/univercd/cc/td/doc/product/voice/ip_tele/network/dgclustr.htm

- Catalyst DSP Resources for Transcoding and Conferencing:
  http://www.cisco.com/univercd/cc/td/doc/product/voice/c_callmg/3_1/sys_ad/adm_sys/ccmsys/a05dsp.htm

- CCM Products and Technologies:
  http://www.cisco.com/warp/public/cc/pd/nemnsw/callmn/index.shtml

Laboratory Exercise: CCM

The laboratory exercises are designed to reinforce concepts discussed throughout the course. This laboratory exercise introduces CCM design considerations.

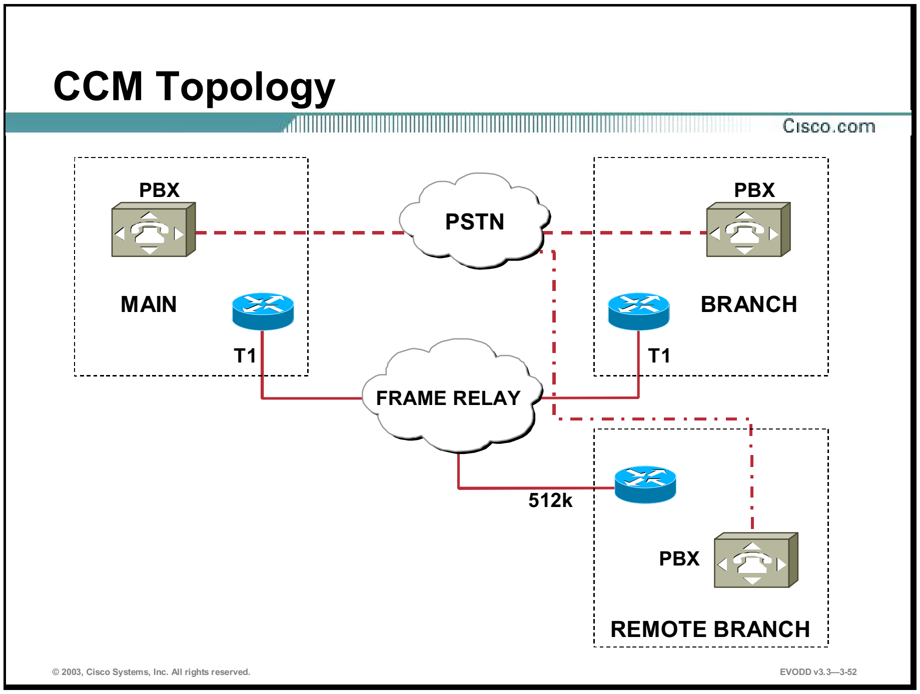
Exercise Objective

In this exercise, you will design a solution for a client that would allow them to use their existing data network for voice.

After completing this exercise, you will be able to:

- Analyze client requirements for voice
- Design a voice and data network using CCM

Exercise Procedure



This topology depicts the current voice and data network of DJB, Inc. The company currently has a separate infrastructure for voice and data. You have been contracted to design a solution for the client that would allow them to use their existing data network for voice. The client wants to ensure redundancy but also needs to keep cost within budget.

Practice

Use the table provided on the following page to document your solution. Make sure that you take the following issues into consideration and include them in your solution.

- General issues:
  - Admission control
  - Codec used
  - Additional hardware needed
  - Redundancy
  - Bandwidth requirements
  - Sever placement
  - Gateways

The details of each site are as follows:

- Main:
  - 4200 employees
  - Three buildings connected with fiber
  - T1 bandwidth utilization is at 40 percent
- Branch:
  - 1040 employees
  - One building
  - T1 bandwidth utilization is at 40 percent
- Remote Branch:
  - 20 employees
  - One building
  - 512 bandwidth utilization is at 60 percent
- Additional details:
  - Allows for 20 simultaneous calls over the WAN between the main office and the branch office
  - Allow for five simultaneous calls over the WAN between the remote branch and the branch office

Table: CCM Chart

Questions	Main	Branch	Remote
How many servers?			
What type of gateways?			
Bandwidth Requirement?			
What type of call admission control?			
What code is used?			

Lesson Review

This practice exercise reviews what you have learned in this lesson.

Q1) Which three of the following components are “Enterprise Cisco IP telephony Building Blocks”? (Choose three.)

A) PBX migration strategies

B) Properly provisioned QoS-enabled infrastructure

C) Unified messaging

D) Fax relay

Q2) What database stores a Cisco IP telephony network configuration database?

A) IMAP

B) SQL

C) LDAP

D) DC Directory

Q3) What number of IP Phones on a network indicates that you should add a separate CCM to act as a SQL publisher and TFTP server?

A) 250

B) 500

C) 1000

D) 2500

Q4) What is the weight (per session per voice channel) for an H.323 client?

A) 1

B) 3

C) 5

D) 20

Q5) What is the protocol for signaling between CCM clusters?

A) H.323

B) T.120

C) RTCP

D) QSIG

Q6) Which of the following statements applies to transcoding via DSPs?

A) Transcoding is only supported between a high bit rate codec and a high bit rate codec.

B) Transcoding is only supported between a low bit rate codec and a low bit rate codec.

C) Transcoding is only supported between a high bit rate codec and a low bit rate codec.

D) All of the above are valid transcoding actions.

Q7) Which two of the following statements apply to location-based CAC? (Choose two.)

A) IP bandwidth limits are assigned per location (in kbps).

B) If insufficient bandwidth is available, the call attempt returns to the PSTN.

C) Cisco recommends using a single codec for all calls crossing the WAN.

D) Probes are sent to determine conditions within the network.

Q8) Which of the following statements best describes an isolated CCM deployment?

A) CCMs are located at a single site.

B) CCMs are located at multiple sites and are interconnected via the IP WAN.

C) CCMs are located at multiple sites but are not interconnected.

D) None of the above

Gateway Types

Overview

Cisco IP telephony gateways enable Cisco CallManager (CCM) to communicate with non-IP telecommunications devices. This lesson teaches you about the types of gateways and gateway technologies supported by CCM.

Importance

Communication outside of a local IP telephony network is a critical migration component. This lesson addresses gateways that you can use to achieve this "off-net" communication.

Objectives

Upon completing this lesson, you will be able to:

- Identify critical gateway features for Cisco IP telephony solutions
- Describe the functionality and features of H.323 and Media Gateway Control Protocol (MGCP) gateways
- Describe the functionality and features of VG200 and VG248 gateways
- Describe the functionality and features of Catalyst 6500 gateways
- Describe the functionality and features of Catalyst 4000 gateways
- Describe the features and functionality of Cisco IOS Telephony Service and Survivable Remote Site Telephony (SRST)
- Determine the proper gateway for a given environment

Learner Skills and Knowledge

To fully benefit from this lesson, you must have these prerequisite skills and knowledge:

- A fundamental understanding of IP telephony, as presented in the preceding modules
- A familiarity with the Cisco router and switch products

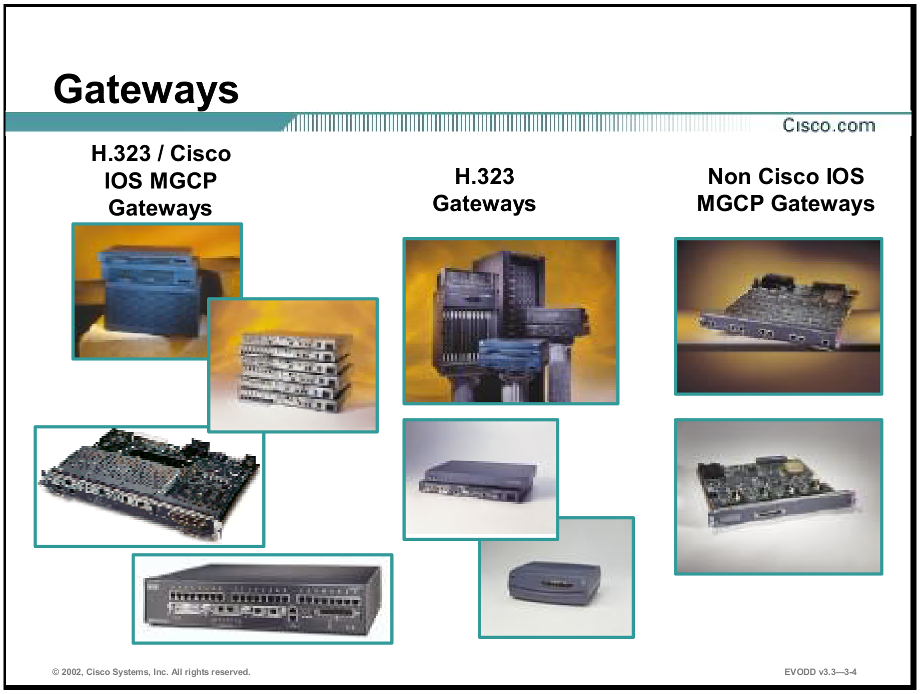
Outline

This lesson includes these topics:

- Overview
- Gateway Features
- H.323 and MGCP Gateways
- VG200 and VG248 Gateways
- Catalyst 6500 Gateways
- Catalyst 4000 Gateways
- Cisco IOS Telephony Services and SRST
- Gateway Reference Charts
- Summary
- Lesson Review

Gateway Features

This topic describes the features of gateways.



Gateways provide protocol conversion between terminals running different types of protocols. IP telephony gateways allow CCM to communicate with non-IP telecommunications devices. The figure shown here describes the H.323 and MGCP gateways. There are two types of Cisco access gateways: analog and digital.

- Cisco access analog gateways have two categories:
  - Access analog station gateways connect CCM to Plain Old Telephone Service (POTS) analog telephones, interactive voice response (IVR) systems, fax machines, and voice-mail systems. Station gateways provide Foreign Exchange Station (FXS) ports.
  - Access analog trunk gateways connect CCM to a Public Switched Telephone Network (PSTN) central office (CO) or PBX trunks. Trunk gateways provide Foreign Exchange Office (FXO) ports for PSTN or PBX access and recEive and transMit (E&M) ports for the analog trunk connection to a legacy PBX.
- Cisco access digital trunk gateways connect CCM to the PSTN or to a PBX via digital trunks, such as PRI or T1 channel associated signaling (CAS). Digital T1 PRI trunks may connect to some legacy voice-mail systems. Whenever possible, use digital gateways to minimize any answer and disconnect supervision issues. Analog Direct Inward Dialing (DID) is also available for PSTN connectivity.

Critical Gateway Features for IP Telephony Designs

Cisco.com

- DTMF relay capabilities
- Supplementary services support
- CCM redundancy support

 EVODD v3.3—3-5

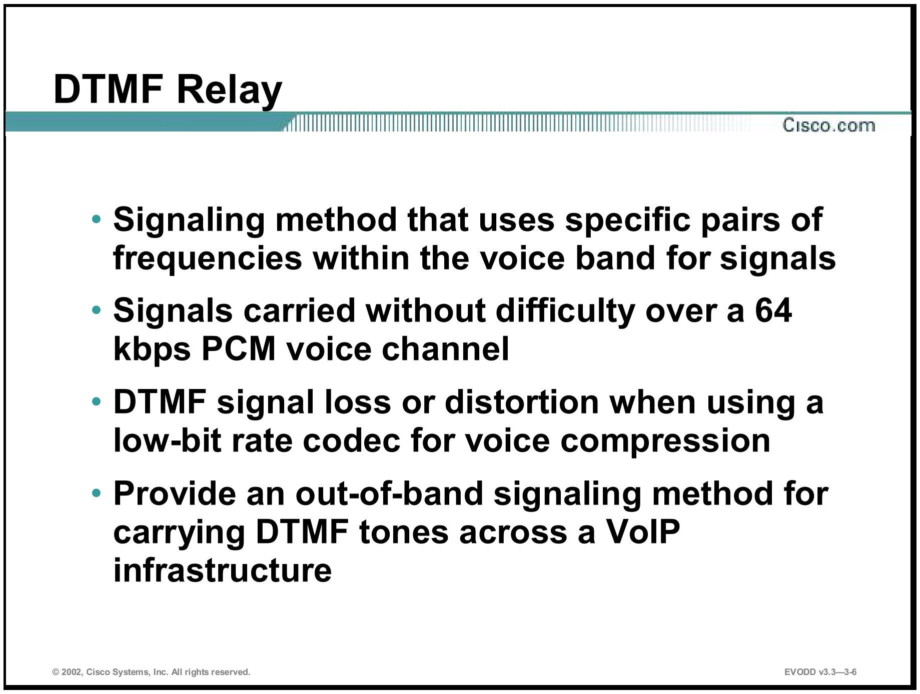
When selecting an IP telephony gateway, consider the common or core requirements and the site and implementation-specific features. IP telephony gateways must meet the following core requirements:

- **Dual tone multifrequency (DTMF) relay capabilities:** DTMF relay capability, specifically out-of-band DTMF, separates DTMF digits from the voice stream and sends them as signaling indications through the gateway protocol (H.323 or MGCP) signaling channel instead of as part of the voice stream or bearer traffic. Out-of-band DTMF is needed when using a low bit rate coder-decoder (codec) for voice compression because the potential exists for DTMF signal loss or distortion.

- **Supplementary services support:** Supplementary services are typically basic telephony functions such as hold, transfer, and conferencing.

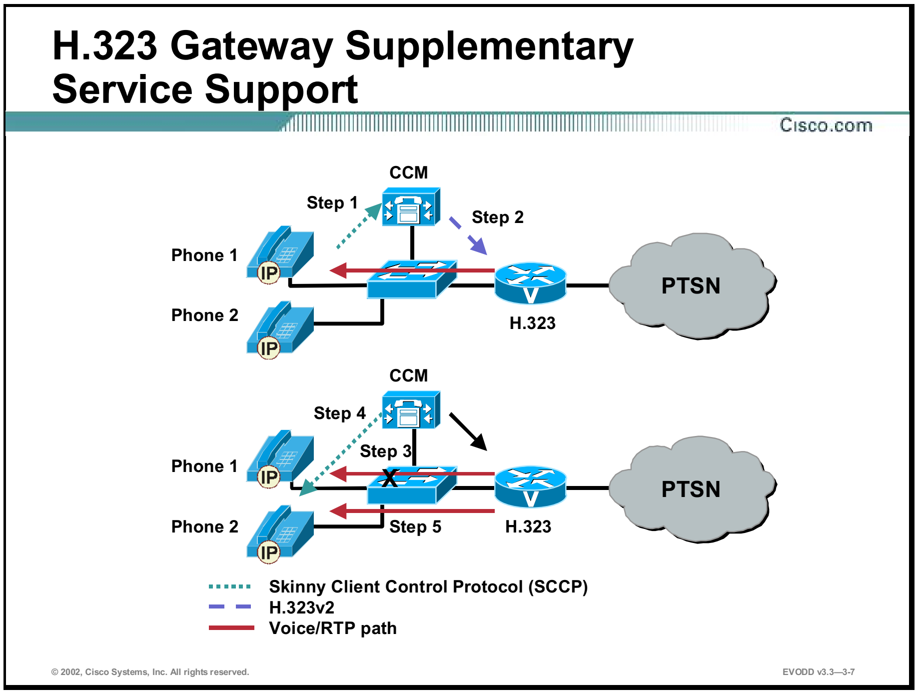
- **CCM redundancy support:** CCM clusters provide CCM redundancy. The gateways must support the ability to "re-home" to a secondary CCM in the event that a primary CCM fails. This differs from call survivability in the event of a CCM or network failure, when an SRST-enabled router takes over the role of the CCM.

- In addition to meeting the core requirements, every IP telephony implementation has its own site-specific feature requirements, such as analog or digital access, DID, and capacity requirements.



DTMF is a signaling method that uses specific pairs of frequencies within the voice band for signals. The signal, in this case, is the digits. The actual representation of the two frequencies is a number or digit. These signals travel over a 64 kbps pulse code modulation (PCM) voice channel without distortion. However, when using a low-bite rate codec for voice compression, the potential exists for DTMF signal loss or distortion. An efficient solution for these problems is to provide an out-of-band signaling method for carrying DTMF tones across a Voice over IP (VoIP) infrastructure.

The Skinny Gateway Protocol can carry DTMF signals out-of-band using TCP port 2002. H.323 gateways can transport DTMF signals as plain text, and MGCP gateways communicate DTMF signals to the CCM using “symbols” sent over the User Datagram Protocol (UDP) control channel.



Supplementary services provide user functions such as hold, transfer, and conferencing. These services are fundamental requirements for any voice installation. Each gateway that you evaluate for use in an IP telephony network should provide support for supplementary services natively, without the use of a software media termination point (MTP). A description of the gateways follows:

- **SCCP Gateways:** The SCCP gateways use the Gateway-to-CCM signaling channel and SCCP to exchange call control parameters.

- **H.323 Gateways:** H.323v2 implements Open/Close LogicalChannel and the emptyCapabilitySet features. CCM Release 3.1 and later only allocates a transcoder dynamically if it is required during a call to provide access to G.711-only devices, while still maintaining a G.729 stream across the WAN. Full support for H.323v2 is available in Cisco IOS Release 12.1.1T.

After an H.323v2 call is set up between a Cisco IOS gateway and an IP Phone, using CCM as a H.323 proxy, the IP Phone can request to modify the bearer connection. Because the Real-Time Transport Protocol (RTP) stream is directly connected to the IP Phone from the Cisco IOS gateway, a supported voice codec can be negotiated between the IP Phone and the gateway. The figure above and the following steps illustrate a call transfer between two IP Phones:

Step 1 In order to transfer the call from the Cisco IOS gateway to Phone 2, Phone 1 issues a transfer request to CCM using SCCP.

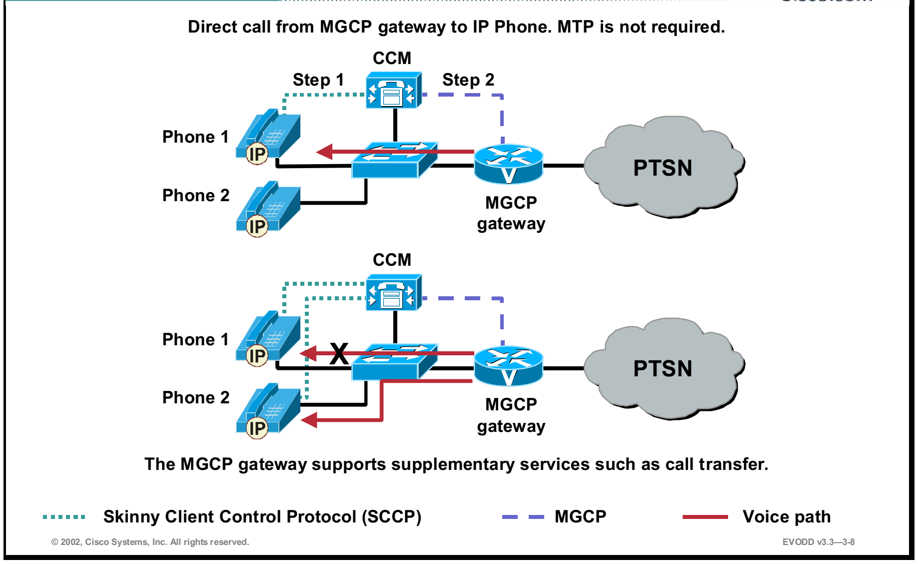
Step 2 CCM translates this request into an H.323v2 CloseLogicalChannel request to the Cisco IOS gateway for the appropriate session ID.

Step 3 The Cisco IOS gateway closes the RTP channel to Phone 1.

Step 4 CCM issues a request, using SCCP, to Phone 2 to set up an RTP connection to the Cisco IOS gateway. At the same time, CCM issues an OpenLogicalChannel request to the Cisco IOS gateway with the new destination parameters, but it uses the same session ID.

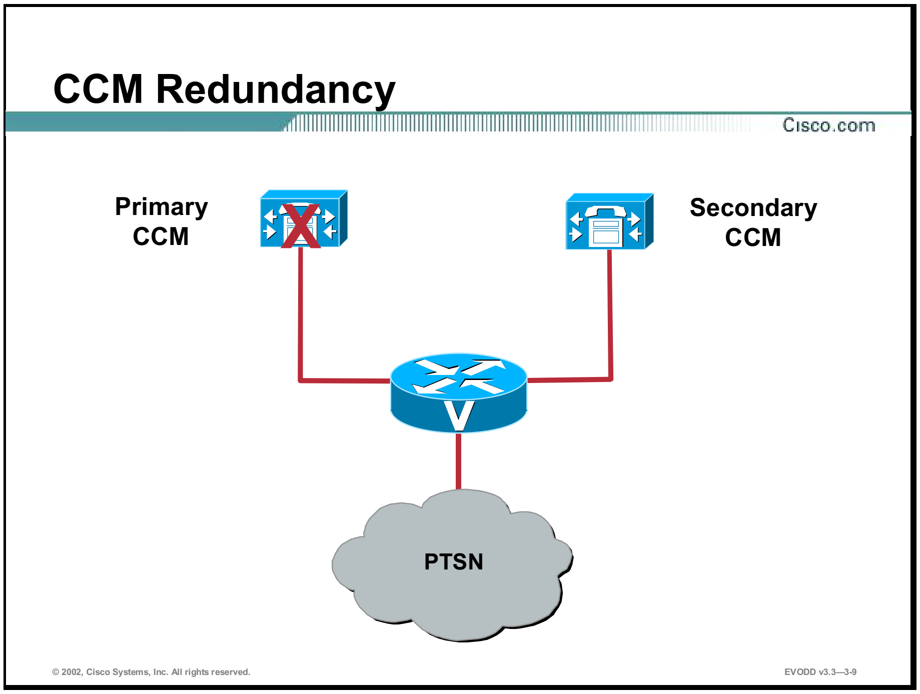
Step 5 After the Cisco IOS gateway acknowledges the request, an RTP voice bearer channel is established between Phone 2 and the Cisco IOS gateway.

MGCP Gateway Supplementary Service Support



The MGCP gateways provide full support for the hold, transfer, and conference features through the MGCP protocol. Because MGCP is a master and/or slave protocol and CCM controls all session intelligence, CCM can easily manipulate the MGCP gateway voice connections.

If an IP telephony endpoint (for example, an IP Phone) needs to modify the session (for example, transfer the call to another endpoint), the endpoint notifies CCM using SCCP. CCM then sends this information to the MGCP gateway, using the MGCP UDP control connection to terminate the current RTP stream associated with the session ID and to start a new media session with the new endpoint information. The figure illustrates the protocols exchanged among the MGCP gateway, endpoints, and CCM.



CCM redundancy is the third gateway requirement. When SCCP gateways initialize, a list of CCMs downloads to the gateways. This list is prioritized into a primary CCM and secondary CCM section. If the gateway cannot reach the primary CCM, it registers with the secondary CCM.

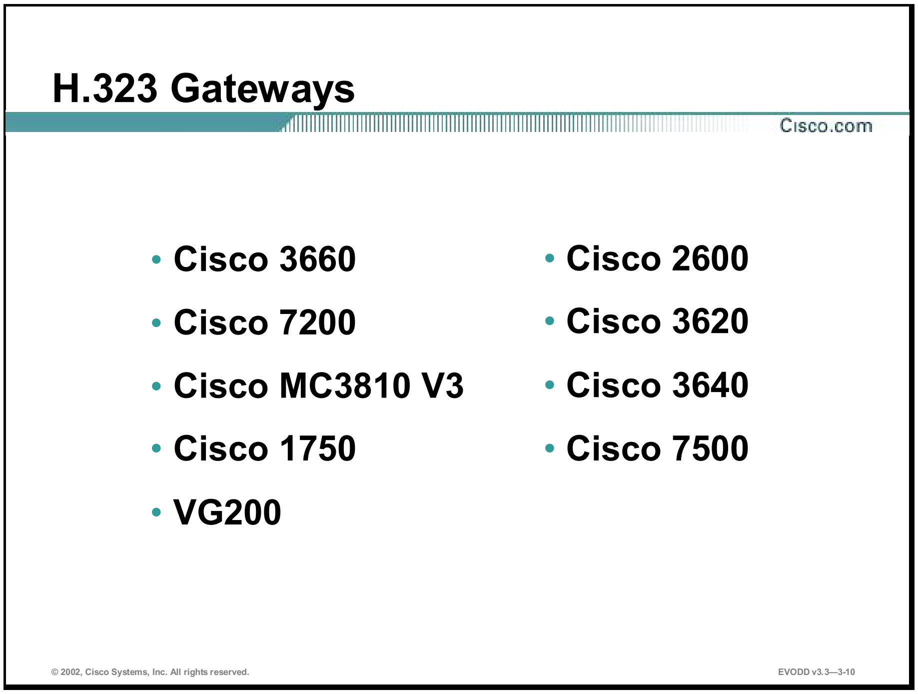
Cisco H.323 gateways support redundant CCMs using several enhancements to the **dial-peer** and **voice class** command sets in Cisco IOS Release 12.1(2)T. This release has a new command, **H.225 tcp timeout <seconds>**, that tracks the time it takes for the H.323 gateway to establish an H.225 control connection for H.323 call setup. If the H.323 gateway cannot establish an H.225 connection to the primary CCM, it tries a second CCM from another **dial-peer** statement. The H.323 gateway shifts to the **dial-peer** statement with next highest preference setting.

MGCP gateways also have the ability to failover to a secondary CCM in the event of communication loss with the primary CCM. When the failover occurs, active calls are preserved.

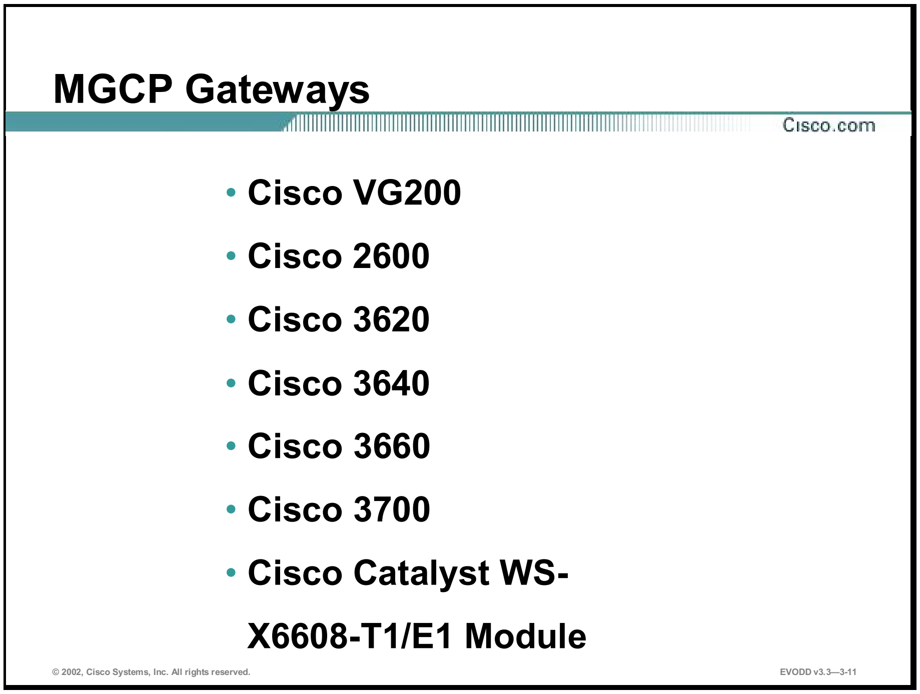
Within the MGCP gateway configuration file, the **call-agent <hostname>** command identifies the primary CCM, and the **ccm-manager redundant-host** command adds a list of the secondary CCMs. Keepalives with the primary CCM are through the MGCP application-level keepalive mechanism, whereby the MGCP gateway sends an empty MGCP notify message to CCM and waits for an acknowledgement.

H.323 and MGCP Gateways

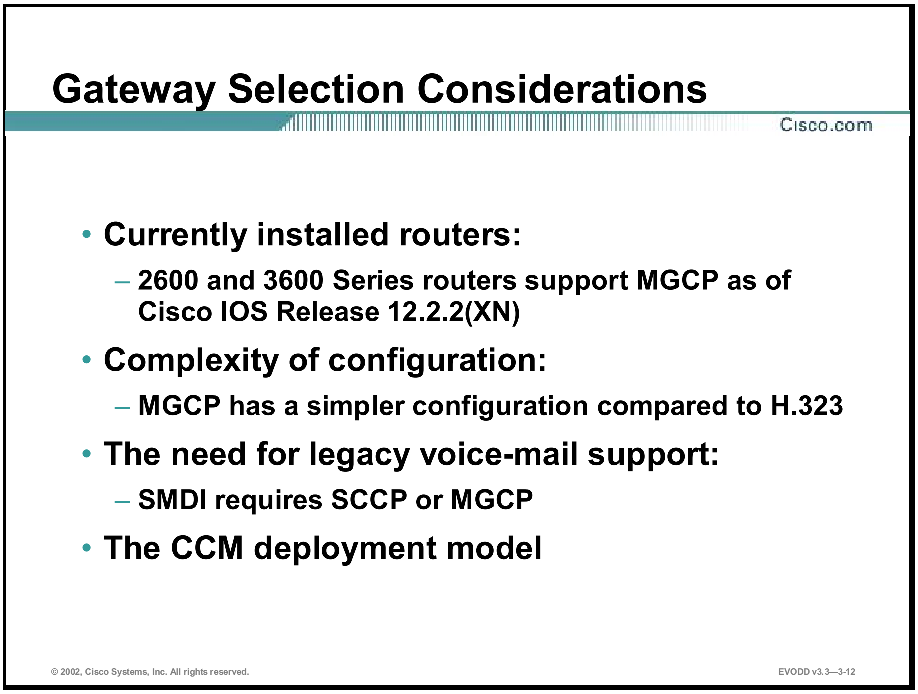
This topic describes the H.323 and MGCP gateways, but it does not cover specifics of each gateway listed.



This figure shows a list of the H.323 gateways that were available from Cisco at the time of this printing. Please see References at the end of this lesson for specific information on gateways.



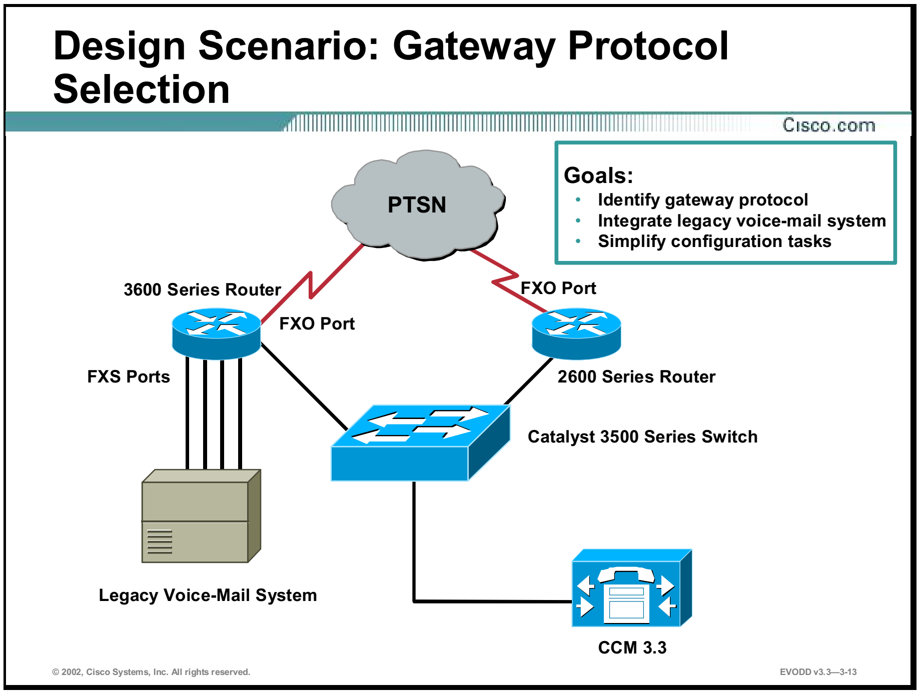
This figure shows a list of the MGCP gateways that were available from Cisco at the time of this printing. The main distinction between H.323 gateways and MGCP gateways is the location of the call routing intelligence (dial plan information). With an H.323 gateway, you configure dial plans via dial peers in the Cisco IOS. Conversely, MGCP gateways rely on the CCM to specify the destination for a VoIP call.



Protocol selection depends on site-specific requirements and the installed base of equipment. For example, most remote branch locations have Cisco 2600 or 3600 series routers installed. These routers support H.323 and MGCP with Cisco IOS Release 12.2.2(XN) and CCM Release 3.1 or later. For gateway configuration, some customers may prefer MGCP to H.323 due to simpler configuration or support for call survivability during a CCM switchover from a primary to a secondary CCM; however, some customers may prefer H.323 to MGCP because of the robustness of the interfaces that it supports.

Simplified Message Desk Interface (SMDI) is a standard for integrating voice-mail systems to PBXs or Centrex systems. Connecting to a voice-mail system via SMDI and using either analog FXS or digital T1 PRI requires either the SCCP or MGCP protocol because H.323 devices do not identify the specific line that is used from a group of ports. Use of H.323 gateways for this purpose means the Cisco message interface cannot correctly correlate the SMDI information with the actual port or channel used for an incoming call.

Additionally, the CCM deployment model can influence gateway selection. For example, a centralized deployment model requires gateways at remote sites that support a feature such as SRST.



Consider the design scenario in the figure. The customer has two routers that connect to the PSTN, via FXO ports, inside Cisco 2600 and 3600 series routers. While the customer has migrated from a PBX to a CCM for call processing, a legacy voice-mail system remains. Part of your design must integrate this voice-mail system into the new IP telephony network.

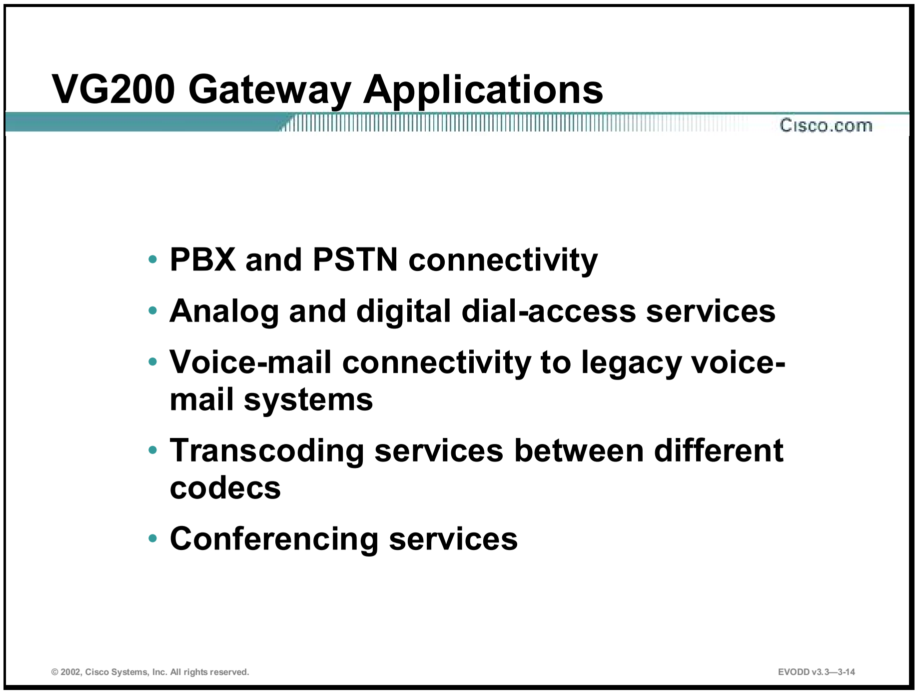
Additionally, the customer wants to minimize configuration tasks. Therefore, the goal for this design scenario is to identify an appropriate gateway protocol that will integrate the legacy voice-mail system, while minimizing configuration.

As of Cisco IOS Release 12.2.2(XN), the Cisco 2600 and 3600 series routers support the MGCP gateway protocol. Therefore, it is possible to choose between the H.323 and MGCP gateway protocols. First, consider how to integrate the legacy voice-mail system into the existing network. CCM can access ports on the legacy voice-mail system using the SMDI protocol. However, although the H.323 gateway protocol does not fully support an SMDI implementation while MGCP does.

Because MGCP relies on CCM for dial-plan information, it is easier to configure an MGCP gateway. So, based on the customer requirements and existing platforms, MGCP is an appropriate gateway protocol selection.

VG200 and VG248 Gateways

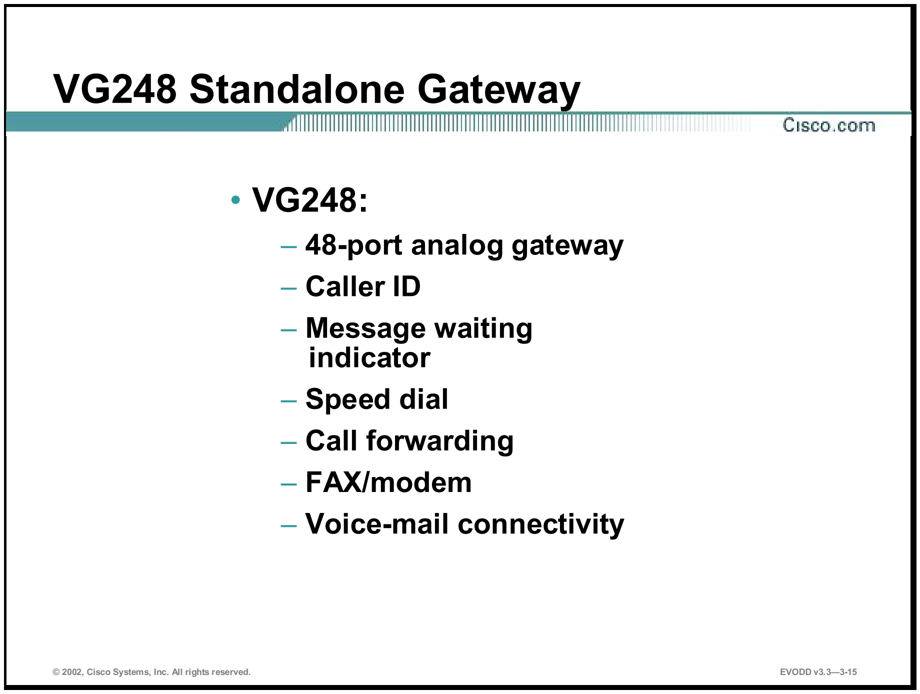
This topic describes the functions and features of the Cisco VG200 and VG248 gateways.



The Cisco VG200 gateway allows you to connect POTS devices to the Cisco IP telephony solution. The Cisco VG248 is ideal for implementations with analog telephones because it provides a high level of functionality at these locations.

The Cisco VG200 connects a Cisco IP telephony network to traditional telephone trunks or analog devices. These telephone trunks may connect to the PSTN or existing PBX systems. Analog devices include legacy telephones, fax machines, and voice conference units. On the data network side, the Cisco VG200 provides an autosensing 10/100 Ethernet port. Internally, the Cisco VG200 is equipped with digital signal processors (DSPs) that convert analog and digital voice into IP packets for transport through the IP network using standard codecs, including G.711, G.723.1, G.729a, and others.

Note	Cisco designed and validated the VG200 for CCM environments only. It is not intended, tested, or supported for applications, such as H.323 toll bypass, or as a session initiation protocol (SIP) gateway.



The Cisco VG248 analog telephone gateway is a mixed-environment solution, which allows organizations to support their legacy analog devices while taking advantage of the new opportunities afforded by IP telephony. The Cisco VG248 is a high-density gateway for using analog telephones, fax machines, modems, voice-mail systems, and speakerphones within an enterprise voice system that is based on CCM. Integrating these devices with the IP-based telephone system provides increased manageability, scalability, and cost-effectiveness.

The key features and benefits of using the Cisco VG248 include the following:

- **Analog telephones:** Full-feature analog telephone connectivity is needed when the infrastructure (wiring) or application does not support or require IP Phones. Full-featured analog telephone lines allow organizations to deploy IP telephony without having to purchase IP Phones for all of the users.

- **Fax and modem:** Fax machine and modem connectivity are required for business needs in many locations, and the Cisco VG248 is the ideal device to provide the lines for them. Fax machine and modem connectivity allows organizations to support these legacy devices with New World IP Phones.

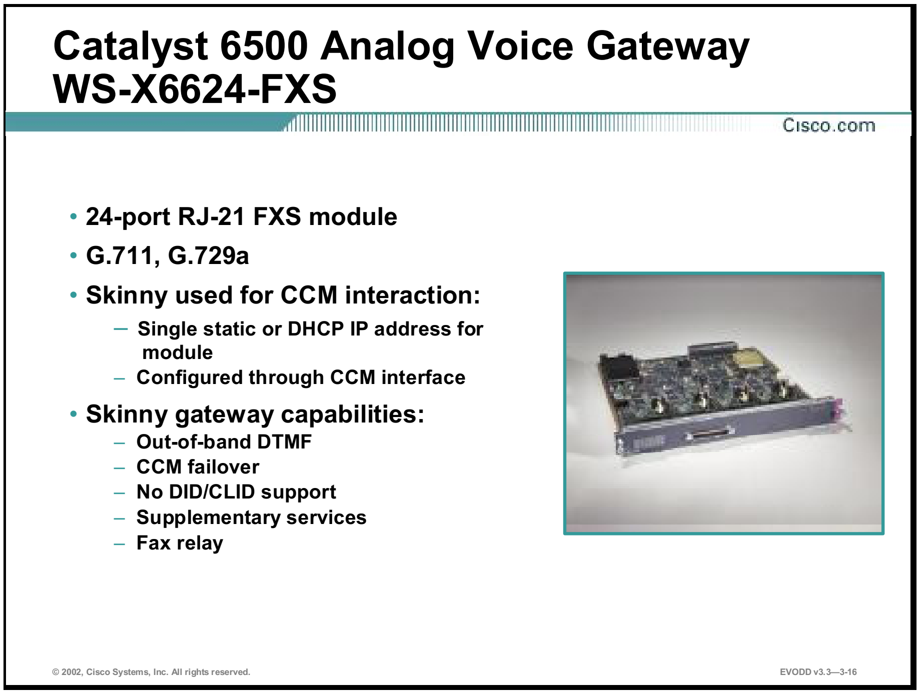
- **Voice mail:** Many legacy voice-mail systems require analog and SMDI connectivity. The Cisco VG248 provides support in more configurations for legacy voice-mail systems with CCM, and with higher reliability than was previously possible.

- **Investment protection:** Customers can continue to use existing telephones, fax machines, and voice-mail systems while taking advantage of IP telephony.

- **Reduced barrier to entry:** With a low-cost alternative for low-end analog telephones, organizations can take advantage of IP telephony with a lower overall investment in this new technology.

Catalyst 6500 Gateways

This topic describes the functionality and features of the Cisco Catalyst 6500 gateway family.



Several telephony modules are available for the Cisco Catalyst 6500 family. These modules act as telephony gateways that enable the implementation of IP telephony on the network using existing Cisco Catalyst 6500 family devices. The Catalyst 6500 gateways are line cards that you can install in any 6500 series switch.

The figure shows an analog voice gateway for the Catalyst 6500. A 24-port FXS gateway supports G.711 and G.729a. You use the CCM interface for configuring and the MGCP for all gateway interaction with CCM. However, because it is a Catalyst blade, it is also possible to configure a static IP address. You must also configure the TFTP server, so that the module recognizes where to obtain the configuration information to download.

Note	This module is very similar to the analog station gateway, with the same basic functionality, except it has 24 ports.

The main capabilities of the Catalyst 6500 gateway are full out-of-band DTMF and CCM failover. It does not provide DID support or calling line identification (CLID) support.

The Catalyst 6500 gateway supports supplementary services and Cisco fax relay, which is an ideal blade for connecting several fax machines or Polycom (speakerphone) devices within a building. You can attach Polycom analog conferencing telephones or fax machines and use the Catalyst 6500 for voice-mail integration when the WS-X6624-FXS module is in a Catalyst 6000.

Catalyst 6500 Digital Voice Gateway WS-X6608-T1/WS-X6608-E1

Cisco.com

- 8 ports T1/E1 PRI—User/network side
- G.711, G.729a
- Skinny used for CCM interaction:
  - Static or DHCP IP addresses for each port
  - Configured through CCM interface
- MGCP gateway capabilities:
  - Out-of-band DTMF
  - CCM failover
  - DID/CLID support
  - Supplementary services
  - Fax relay

EVODD v3.3—3-17

The component in the figure is the digital counterpart of the Catalyst 6500 analog voice gateway, WS-X6624-FXS. The WS-X6608-T1/E1 module allows larger enterprises to connect the PSTN and legacy PBXs directly to the campus multiservice network. Telephony signaling types supported include:

- **Common channel signaling (CCS):** In this mode, there are 23 digital signal level 0 (DS-0) channels for T1 and 29 for E1 for voice traffic; the 24th T1 DS-0 or 16th E1 channel is for signaling. It is possible to configure any channel for CCS.

- **ISDN PRI signaling:** Each interface supports 23 channels for T1 and 30 channels for E1. The default mode is for the 24th T1 channel or 16th E1 channel reserved for signaling. Both network side and user side operation modes are supported.

- **CAS:** This gateway supports T1 CAS if you are using it with CCM 3.1 or greater.

- **MGCP:** This gateway supports MGCP if you are using it with CCM 3.1 or greater. The MGCP protocol provides gateway failover support.

Catalyst 6500 Voice T1 and Services Module

Cisco.com

Catalyst 6000 8-port voice T1 or E1 and services module:

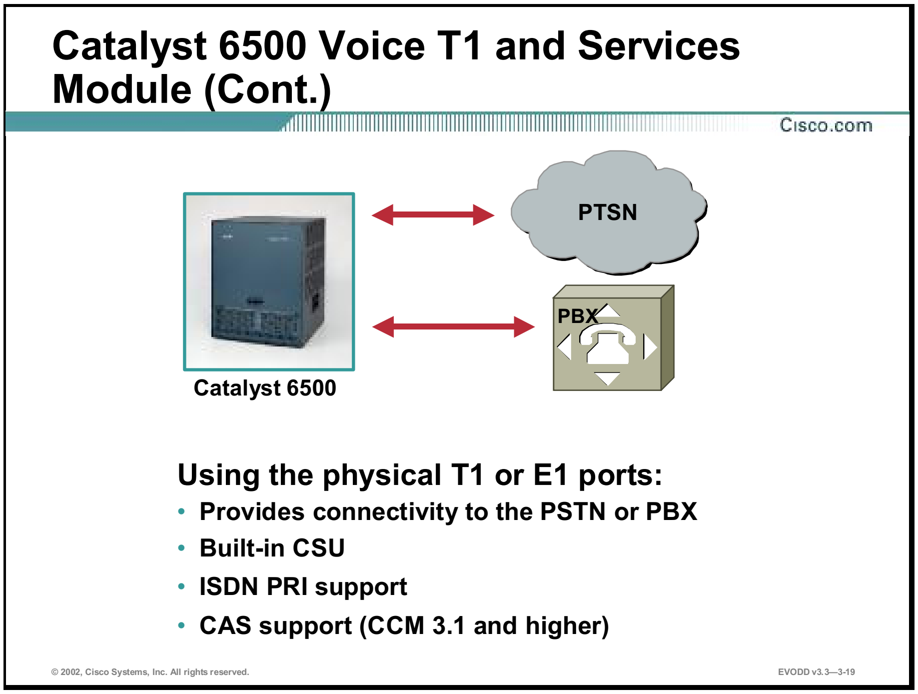
- **Provides 8 physical or logical ports:**
  - User configurable on a port-by-port basis
- **Configured as a physical port:**
  - T1 or E1 port for connections to the PSTN or PBX
- **Configured as a logical port:**
  - Provides voice services, such as conferencing or transcoding
  - Shared network resource

 EVODD v3.3—3-18

The eight-port voice T1 module for the Catalyst 6500 provides eight physical T1 ports for connecting to the PSTN or to the PBX. There is also a separate module with eight physical E1 ports.

Note	Cisco has shipped both modules since the release of CCM 3.0.

If users do not want to use all eight ports, it is possible to configure the unused ports as logical ports. If a port is configured as a logical port, its digital signal processors (DSPs) can provide network voice services, such as compression or conferencing. Transcoding converts between a compressed codec (for example, G.729) and G.711. You configure these transcoding and conferencing features via the CCM. Each of the eight ports of the module has its own MAC address, so the CCM can address each port by its individual address.



In the figure shown here, the WS-X6608-T1 module acts as a physical T1 connection to the PBX or the PSTN. These ports connect to a time-division multiplexing (TDM)-based infrastructure. The T1 ports have a built-in CSU. This module supports ISDN PRI connections.

The CCM can use DSPs for processor-intensive tasks, such as transcoding or supporting a conference call, if all of the eight T1 ports are not used for physical connections to a PBX or PSTN.

Catalyst 4000 Gateways

This topic describes the functionality and features of the Catalyst 4000 gateways.

Catalyst 4000 Voice Gateway WS-X4604-GWY

Cisco.com

- 6 ports of analog (FXS, FXO, or E&M) or 4 ports of analog and 2 T1/E1 ports (T1 CAS/PRI, E1 PRI)
- G.711 and G.729a
- H.323v2 used for CCM interaction:
  - Single IP address for gateway
  - Configured through Cisco IOS CLI and H.323 GW added via CCM interface
- H.323v2 gateway capabilities:
  - Out-of-band DTMF—12.0(7)T
  - Digital DID only
  - Analog CLID—12.1(2)x
  - Cisco fax relay
  - CCM failover—12.1(2)T
  - Digital CLID support via PRI & FGD only
  - Supplementary services 12.0(7)T

 EVODD v3.3—3-20

The Catalyst 4000 access gateway module combines Cisco IOS routing with the Catalyst 4000 family of switches to provide infrastructure consolidation and Cisco IP telephony support for the branch office. This figure summarizes the features for the Catalyst 4000 voice gateway.

The WS-X4604-GWY supports six ports of analog (FXS, FXO, or E&M), or four ports of analog and two ports of T1/E1 voice. It also supports both the G.711 and G.729a codecs.

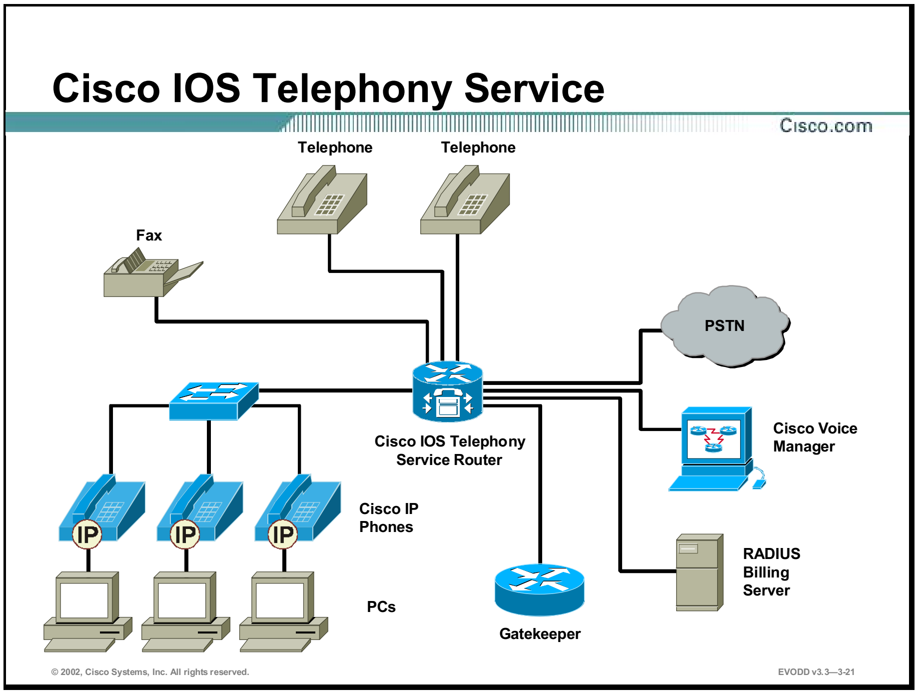
All of the CCM interactions for Catalyst 4000 gateways use H.323v2, unlike the Catalyst 6500 voice gateway modules that use the Skinny Gateway Protocol.

This gateway supports out-of-band DTMF and CCM failover under Cisco IOS Release 12.1(2)T, DID, digital DID only, digital CLID only (Feature Group D only), or PRI.

Note	This module also supports Cisco fax relay.

Cisco IOS Telephony Service and SRST

This topic describes the functionality and features of the Cisco IOS Telephony Service and SRST, which are available for selected gateways that are based on the same technology.



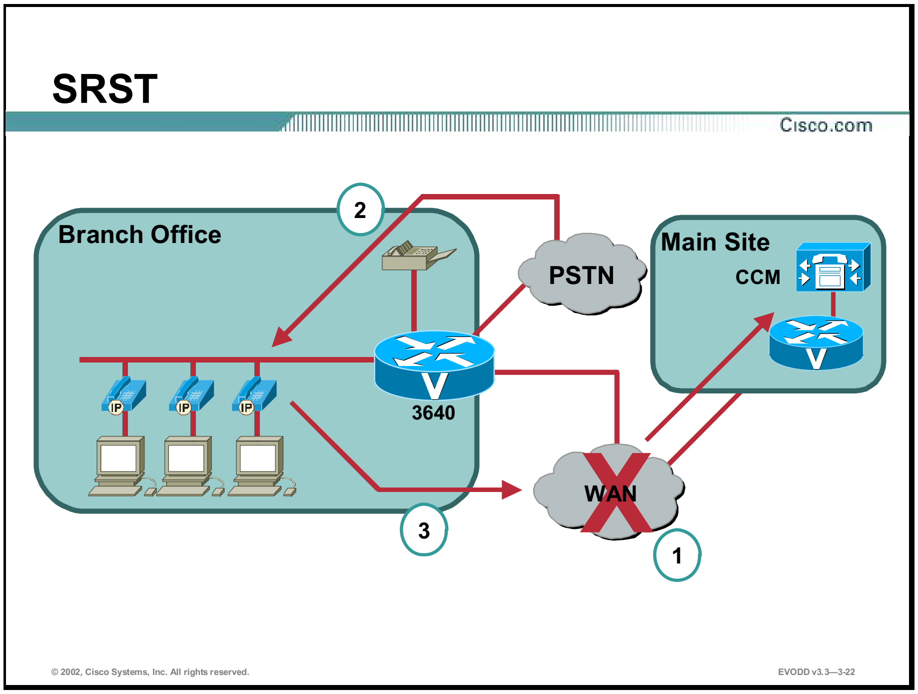
The Cisco IOS Telephony Service feature provides basic Cisco IP Phone call-handling capabilities in a LAN environment on the Cisco routers. Cisco IOS Telephony Service allows you to install IP Phones in a small office without the need of a CCM. The telephones register to the router that is running Cisco IOS Telephony Service. The RADIUS server in the diagram can control who has the ability to call out.

Note Many of the features that CCM offers are not supported in this type of environment; however, it has sufficient features for most small offices.

The following features are supported on the Cisco IP Phones in a Cisco IOS Telephony Service environment:

- Function keys
- Dial-plan class of restriction
- Call hold and retrieve
- Call pickup of on-hold calls
- Multiple lines per Cisco IP Phone

- Multiple line appearance across telephones
- Call forwarding functions, such as *all*, *busy*, and *no answer*
- Call transferring
- Speed-dialing
- Cisco IP Phones derive the date and time from the router through Network Time Protocol (NTP)
- Interworking with Cisco gatekeeper
- Distinctive ringing: External ringing versus internal ringing
- Caller identification display and blocking
- Analog FXS and FXO ports
- On-net calls using VoIP H.323, Voice over Frame Relay (VoFR), and Voice over Asynchronous Transfer Mode (VoATM)
- Supports 1 to 48 telephones depending on the platforms
  - Maximum of 24 Cisco IP Phones for the Cisco 2600 series, Cisco 3620, and the Cisco IAD2420 IADs
  - Maximum of 48 Cisco IP Phones for the Cisco 3640 and Cisco 3660 routers
- Supports 1 to 96 directory numbers depending on the platforms
  - Maximum of 48 directory numbers or virtual voice ports for the Cisco 2600 series, Cisco 3620, and Cisco IAD2420 Integrated Access Devices (IADs)
  - Maximum of 96 directory numbers or virtual voice ports for the Cisco 3640 multiservice routers



The second feature is SRST. In the past, when the WAN link failed in a centralized deployment, the remote site telephones would fail because the telephones were unable to communicate with CCM. SRST allows the telephones to failover to the router that is running SRST. The router then routes all calls out of the PSTN.

As enterprises extend their IP telephony and high-value application deployments from central sites to remote offices, one of the vital deployment factors is the ability to cost-effectively provide backup redundancy functions at the remote branch office. However, the size and number of these small office sites precludes most enterprises from deploying dedicated call-processing servers, unified messaging servers, or multiple WAN links to each site to achieve the high availability required.

The CCM IP telephony solution, combined with the SRST feature in Cisco IOS, allows enterprises to extend high-availability IP telephony to their small branch offices with a cost-effective solution that is extremely simple to deploy, administer, and maintain.

When remotely placing IP Phones from a CCM cluster, provide call-processing redundancy in case there is a WAN failure. Call-processing redundancy is especially critical when users need to place emergency calls, such as 911 calls in the United States, during a WAN outage.

Cisco has developed SRST technology for Cisco 2600, Cisco 3600 series, and Cisco 7200 series branch office access routers, which include network intelligence integrated into Cisco IOS software. This service can act as the call-processing engine for IP Phones located in the branch office during the WAN outage.

SRST is a capability that is embedded in Cisco IOS software and runs on the local branch office access router. SRST automatically detects a failure in the network and uses Cisco Simple Network Automated Provisioning capabilities to initiate a process to autoconfigure the router. This process provides call-processing backup redundancy for the IP Phones in that office. The router provides call processing for the duration of the failure, ensuring that the telephone capabilities are operational.

Upon restoration of the WAN and connectivity to the network, the system automatically shifts call-processing functions to the primary CCM cluster. You can configure this capability in the central CCM, which simplifies deployment, administration, and maintenance.

Note IT staff is not needed at the remote sites to enable and disable this functionality because of the intelligence and simplicity of the SRST feature.

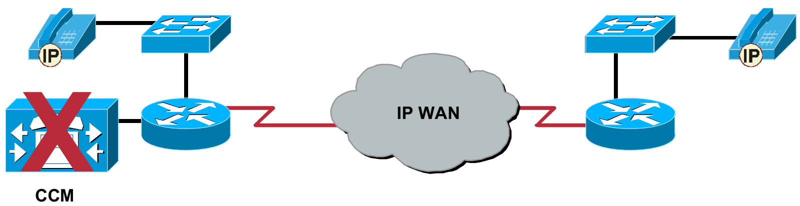
In addition to needing the ability to failover in the case of a WAN outage, a complete CCM solution is not cost justifiable in all cases. In these instances, you can offer a Cisco IOS Telephony Service solution. The features of SRST and Cisco IOS Telephony Service are very similar.

The following are features of SRST:

- Support for re-homing of IP Phones to use call processing on a local router upon CCM fallback
- Support for IP and POTS telephones on the router
- Extension to extension dialing
- Extension to PSTN dialing
- Support for on-net calling
- Primary line on the telephone
- DID
- Direct outward dialing (DOD)
- Calling party identification display
- Calling party name display
- Last number redial
- Call transfer without consultation
- Call hold and retrieve on a shared line

Endpoint Rules for Gateway Call Survivability

Cisco.com



- 2 non-survivable endpoints = call fails
- 1 survivable endpoint, 1 non-survivable endpoint = call fails if non-survivable endpoint's CCM fails

 EVODD v3.3—3-23

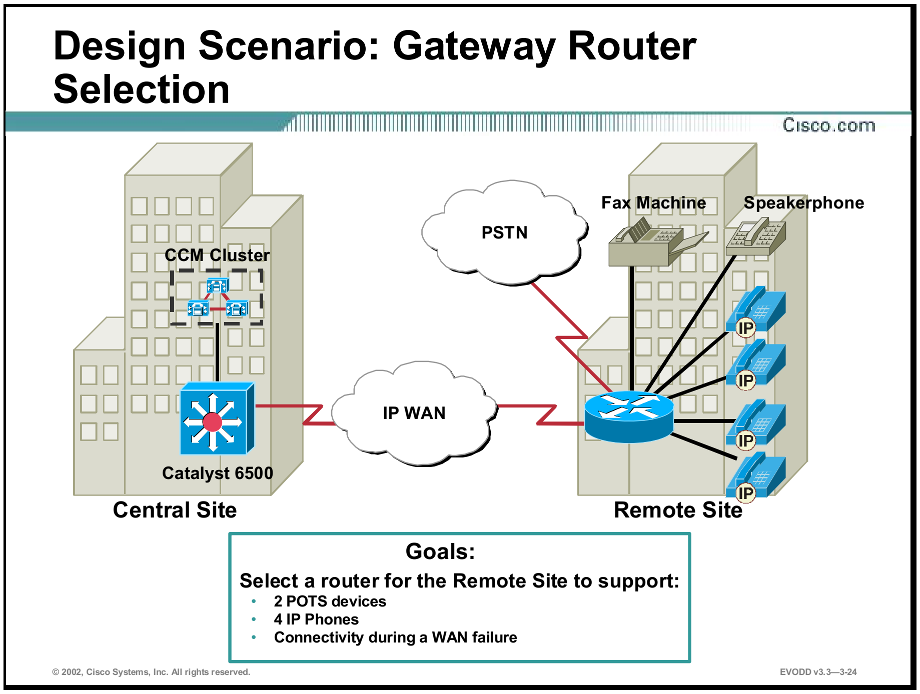
The following rules apply for various endpoint and Cisco Call Manager failure scenarios.

- If the call involves only non-survivable endpoints, the call fails if there is a failure in any of the CCMs to which one of the endpoints is registered. This is true regardless of whether a conference bridge, an MTP, or a transcoder is involved in the call and regardless of which CCM (when there are more than one) fails.

- If the call involves one non-survivable endpoint and one survivable endpoint, the call fails only if the CCM associated with the non-survivable endpoint fails.

- If the call involves only survivable endpoints and one or more CCMs fail, the streaming connection between the endpoints is maintained. However, the endpoints do not have call processing services available to them after the failure. For example, the unavailable services would include transfer, conference, hold, park, pickup, and resume.

In general, MGCP gateways provide the highest degree of call survivability. In CCM Release 3.1 and later, the Catalyst 6000 T1/E1 gateway modules use MGCP supporting PRI and T1 CAS, thus enhancing call survivability when an SCCP-based IP phone and an MGCP gateway are the two endpoints.



Consider the design scenario in the figure. The customer's central site has converted to IP telephony, and they are ready to deploy IP telephony at one of their remote sites. The customer wants to support four IP Phones at this remote site, while still having support for two legacy POTS devices, specifically a fax machine and a speakerphone. The design challenge is to select a gateway router to meet the customer's need.

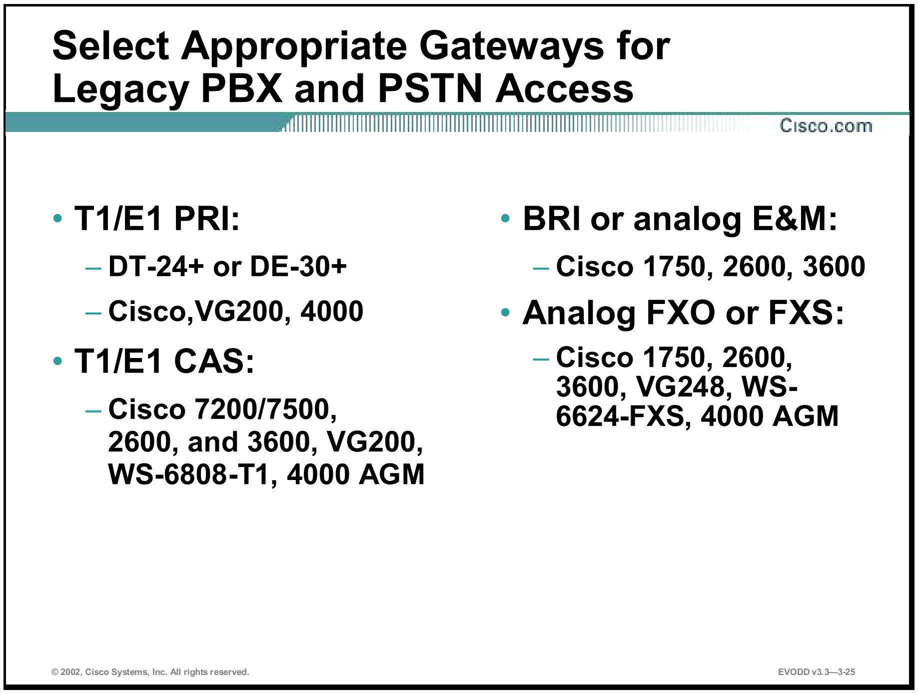
After examining Cisco's voice-enabled router selection, the decision was made to select the Cisco 3725 router. The Cisco 3725 supports a 36-port Ethernet module, which supports in-line power for the IP Phones. Additionally, a VWIC (Voice-WAN Interface Cards) can be installed to support the required two FXS ports, which the legacy speakerphone and fax machine use for connectivity. Additionally, the 3745 can support a T1 interface in its network module for connectivity back to the central site.

While this solution allows the IP Phones to register with the centralized CCM cluster, all the telephony devices at the remote site still need connectivity with the PSTN and with each other, in the event of a WAN failure. Fortunately, the 3725 also supports SRST, and a VWIC containing two FXO ports can be added to connect directly to analog PSTN lines. Therefore, if the WAN link to the central site goes down, both analog and IP Phones at the remote site maintain connectivity within the site, and there is connectivity out to the PSTN.

Note	For more information on the Cisco 3700 Series of routers see the following link: http://www.cisco.com/en/US/products/hw/routers/ps282/products_data_sheet09186a008009203f.html

Gateway Reference Charts

This topic contains gateway reference charts to help you determine the proper gateway for a given network environment.



As with any innovative technology, the number of supported gateways is increasing rapidly.

Table: CPE and Branch Office Multiservice Platform Choices

	Egress Choices	Analog Voice Ports	Digital Voice Ports (T1/E1) DS-Os	Voice Compression	Voice Packets	
					VoIP	VoFR
Cisco 1750	10/100 serial	4/6	4	G.711, G.729	Yes	Yes
Cisco 3810	10BaseT serial Up to 2 T1/E1 T1 ATM	6	24/24	G.711, G.729, G.729a, G.723.1, G.726	Yes	Yes
Cisco 261x 262x	FE TR serial T1/E1 ATM through OC-3	4	48/60	G.711, G.729, G.729a, G.723.1, G.726, G.728	Yes	Yes
Cisco 3620/40	FE TR serial T1/E1 ATM through OC-3	12	96/120	G.711, G.729, G.729a, G.723.1, G.726, G.728	Yes	Yes
Cisco 3660	FE TR serial T1/E1 ATM through OC-3	24	288/300	G.711, G.729, G.729a, G.723.1, G.726, G.728	Yes	Yes

Table: Head-End Multiservice Platform Choices

	Egress Choices	Analog Voice Ports	Digital Voice Ports (T1/E1) DS-Os	Voice Compression	Voice Packets	
					VoIP	VoFR
Cisco 3660	FE TR serial T1/E1 ATM through OC-3	24	288/300	G.711, G.729, G.729a, G.723.1, G.726, G.728	Yes	Yes
Cisco 7200	FE TR serial T1/E1 ATM through OC-3	0	288/360	G.711, G.729, G.729a, G.723.1, G.726	Yes	Yes
Cisco 7500	FE TR serial T1/E1 ATM through OC-3	0	288/360	G.711, G.729, G.729a, G.723.1, G.726, G.728	Yes	Yes

Table: Summary: Cisco IP Telephony Gateways and Supported Protocols

Gateway	Skinny Gateway Protocol	H.323	MGCP
VG200	No	Yes	Yes
Catalyst 4000 WS-X4604-GWY gateway module	Yes, for conferencing and MTP transcoding services	Yes, for PSTN interfaces	Future
Catalyst 6000 WS-X6608-T1 and WS-X6608-E1 gateway modules	Yes	No	Future
Cisco 1750	No	Yes	No
Cisco 3810 V3	No	Yes	Future
Cisco 2600	No	Yes	Future
Cisco 3600	No	Yes	Future
Cisco 7200	No	Yes	No
Cisco 7500	No	Future	No
Cisco AS5300	No	Yes	No

Table: Summary: Analog Gateways by Site-Specific Features

Gateway	FXS	FXO	E & M	Analog DID/CLID
VG200	Yes	Yes	In H.323v2 mode	Future
Cisco 1750	Yes	Yes	Yes	Future
Cisco 2600	Yes	Yes	Yes	12.1(3)T/12.1(2)XH
Cisco 3600	Yes	Yes	Yes	12.1(3)T/12.1(2)XH
Cisco 7200	No	No	No	—
Cisco 7500	No	No	No	—
Cisco AS5300	No	No	No	—
Catalyst 4000 WS-X4604-GWY gateway module	Yes	Yes	Yes	12.1(5)T/12.1(5)T
Catalyst 6000 WS-X6608-T1 and WS-X6608-E1 gateway modules	Yes	No	No	No/Yes

Table: Summary: Digital Gateways by Site-Specific Features

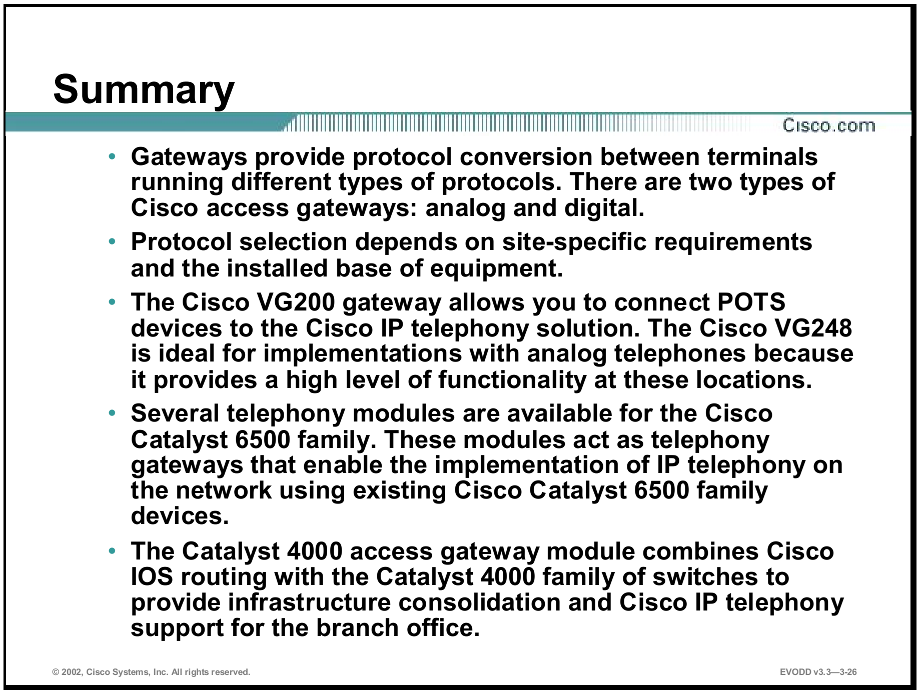
Gateway	T1 CAS	E1/R2	E1 CAS	User Side PRI	Network Side PRI	User Side BRI	Network Side BRI	Digital DID/CLID
VG200	In H.323v 2 mode	No	In H.323v2 mode	No	No	No	No	—
Cisco 1750	No	No	No	No	No	Future	Future	—
Cisco 2600	Yes	12.1(3)T	12.1(3)T	12.1(3)T	12.1(3)T	Yes	12.2(1)T	Yes/Yes
Cisco 3600	Yes	12.1(3)T	12.1(3)T	12.1(3)T	12.1(3)T	Yes	12.2(1)T	Yes/Yes
Cisco 7200	Yes	12.1(3)T	12.1(3)T	12.1(3)T	12.1(3)T	No	No	Yes/Yes
Cisco 7500	Yes	12.1(3)T	12.1(3)T	12.1(3)T	12.1(3)T	No	No	Yes/Yes
Cisco AS5300	Yes	Yes	Yes	Yes	12.0.7T	No	No	Yes/Yes
Catalyst 4000 WS-X4604-GWY gateway module	Yes	Yes	Yes	Yes	Yes	Future	Future	Yes/Yes
Catalyst 6000 WS-X6608-T1 and WS-X6608-E1 gateway modules	No	No	No	Yes	Yes	No	No	Yes/Yes

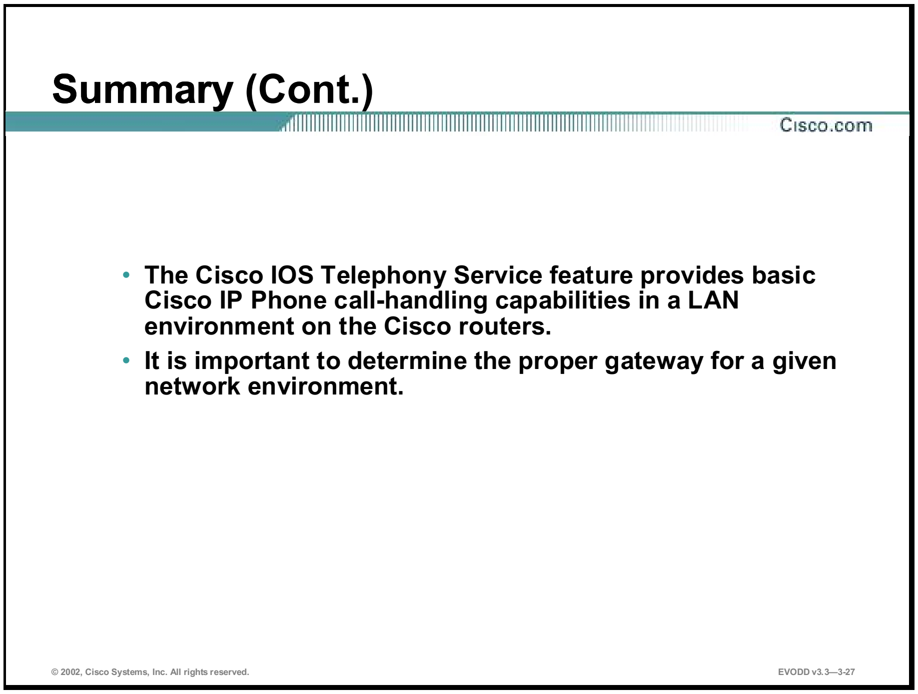
Table: Summary: Gateways with Supported Interfaces and Compression Types

Gateway Type	Gateway	Data Interfaces	Analog PSTN Interfaces	Digital PSTN Interfaces in DS-0s	Voice Compression
Skinny Gateway Protocol	Catalyst 6000 WS-X6624-FXS	10/100/1000 Ethernet	24	0	G.711, G.729a
	Catalyst 6000 WS-X6608-T1	10/100/1000 Ethernet, POS/FlexWAN	0	192	G.711, G.729a
	Catalyst 6000 WS-X6608-E1	10/100/1000 Ethernet, POS/FlexWAN	0	240	G.711, G.729a
MGCP	VG200	100BaseT	4	0	G.711, G.729a, G.723.1
H.323	Cisco 1750	10BaseT, T1/E1 serial	4	0	G.711, G.729
	VG200	100BaseT	4	48/60	G.711, G.729a, G.723.1
	Cisco 2600	10/100BaseT, Token Ring, T1/E1 serial	4	48/60	G.711, G.729a, G.723.1
	Cisco 3620	10/100BaseT, Token Ring, T1/E1 serial, T1-OC3 ATM	4	48/60	G.711, G.729a, G.723.1
	Cisco 3640	10/100BaseT, Token Ring, T1/E1 serial, T1-OC3 ATM	12	136/180	G.711, G.729a, G.723.1
	Catalyst 4000	10/100/1000 Ethernet	6 at FCS	48/60	G.711, G.729a, G.723.1
	Cisco 3660	10/100BaseT, Token Ring, T1/E1 serial, T1-OC3 ATM, HSSI	24	288/360	G.711, G.729a, G.723.1
	Cisco 7200	10/100BaseT, Token Ring, T1/E1 serial, T1-OC12 ATM	0	288/360	G.711, G.729a, G.723.1

Summary

This topic summarizes the key points you learned in this lesson.





Next Steps

After completing this lesson, go to:

- Dial Plans and Voice Mail Considerations module

References

For additional information, refer to these resources:

- VG200 Specifications:
  http://www.cisco.com/warp/public/cc/pd/ga/prodlit/vg200_ds.htm

- Cisco IOS Telephony Service:
  http://www.cisco.com/univercd/cc/td/doc/product/access/ip_ph/ip_ks/ipkey2.htm

- WS-X6608 Gateway Specifications:
  http://www.cisco.com/warp/public/788/products/6608_t1_e1.html

- Gateway Selection:
  http://www.cisco.com/univercd/cc/td/doc/product/voice/ip_tele/network/dggatewy.htm

Lesson Review

This practice exercise reviews what you have learned in this lesson.

Q1) What is the function of DTMF relay?

A) DTMF relay converts pulse-dialed digits into tone-dialed digits.

B) DTMF relay preserves dialed digit information with low-compression codecs.

C) DTMF relay preserves dialed digit information with high-compression codecs.

D) DTMF relay uses TCP, instead of UDP, for tone signaling.

Q2) Which three features belong to Cisco H.323v2 gateways? (Choose three.)

A) DTMF relay

B) Supplementary services

C) CCM failover

D) Analog and digital DID

Q3) Which statement best describes a Cisco VG248?

A) A digital gateway with two T1 interfaces

B) A 48-port analog gateway

C) An ISDN gateway with two PRIs

D) A SIP gateway module for the Cisco Catalyst 6500

Q4) Which gateway allows you to connect analog telephones to a Catalyst 6500 series switch?

A) WS-X6624-FXO

B) WS-X6608-T1

C) WS-X6624-FXS

D) WS-X6607-E1

Q5) What technology is designed for installing IP Phones in a small office without a CCM?

A) Cisco IOS Telephony Service

B) SRST

C) Cisco IVR

D) Cisco Unity

Q6) Which gateway does not support FXS interfaces?

A) Cisco Catalyst 6500

B) Cisco 1750

C) Cisco AS5300

D) Cisco 2600

Q7) You need to attach a remote location to the IP telephony network in the main office. You will NOT convert the telephones in the remote office from analog to IP Phones. This site must have the largest feature set possible. Which device would be best suited for this situation?

A) VG200

B) VG248

C) Catalyst 5500

D) Catalyst 6500

4

Dial Plans and Voice-Mail Considerations

Overview

Most Cisco IP telephony deployments are migrations from existing PBX systems. To make the Voice over Data network as transparent to the user as possible during the migration phase, proprietary PBX signaling must be transported over the Cisco network infrastructure. As the migration progresses, dial plans migrate from the PBX to the Cisco CallManager (CCM).

Upon completing this module, you will be able to:

- Compare standard and nonstandard signaling and describe how PBXs are networked and PBX signaling is transported
- Describe dial plan functionality, features, and considerations

Outline

The module contains these lessons:

- Signaling, Networking, and Transporting
- Dial Plans

Signaling, Networking, and Transporting

Overview

PBX handling requires background information about signaling methods. The signaling type (standard or nonstandard) determines which protocols to use, and the ISDN implementation determines the Transport or Translate model.

Importance

In many cases, customers want to implement Voice over Data and at the same time maintain an existing PBX. To accomplish this, you need to understand the current PBX signaling and transport model for each customer.

Objectives

Upon completing this lesson, you will be able to:

- Describe the differences between standard and nonstandard signaling
- Identify methods of transporting proprietary PBX signaling across a WAN
- Describe networking PBXs over a WAN
- Describe methods for transporting PBX signaling across a WAN
- Identify the specific challenges in leveraging the assets of the Lucent Definity ECS and Nortel Meridian PBXs when deploying VoIP

Learner Skills and Knowledge

To fully benefit from this lesson, you must have these prerequisite skills and knowledge:

- Fundamental understanding of legacy PBX signaling
- Fundamental understanding of IP network characteristics

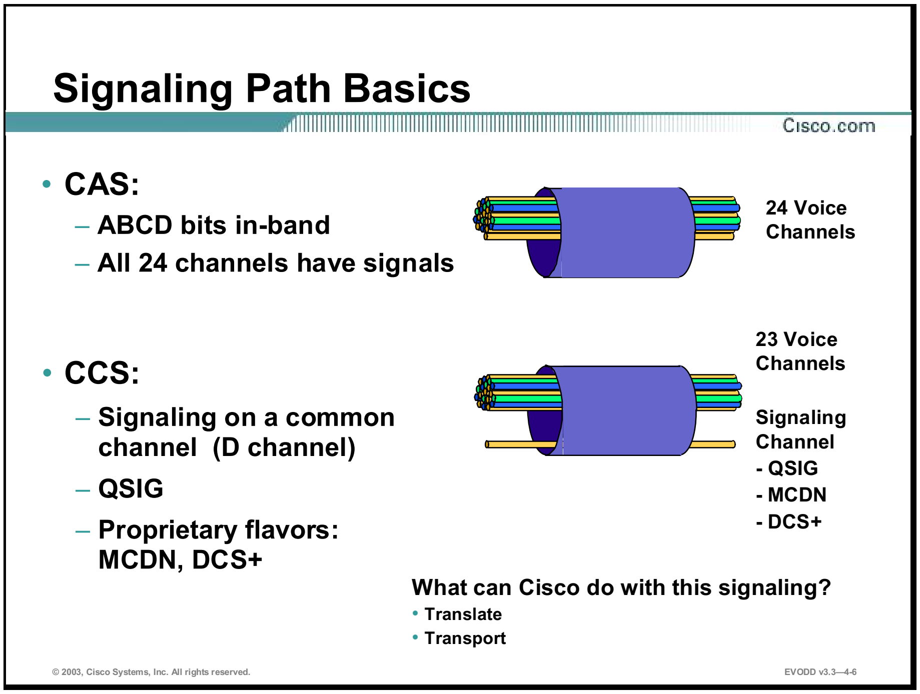
Outline

This lesson includes these topics:

- Overview
- Standard or Nonstandard Signaling
- Signaling Channels
- Networking
- Translate versus Transport Decision Trees
- Traditional PBXs
- Summary
- Lesson Review

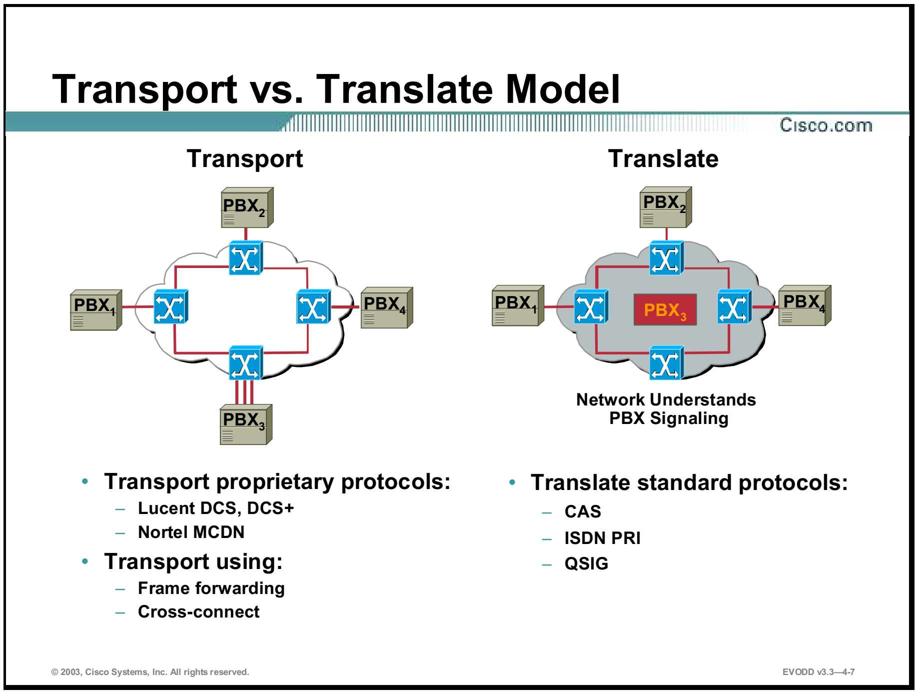
Standard or Nonstandard Signaling

This topic describes how you map PBX signaling to a data network and how the PBX handles PBX-to-PBX or PBX-to-central office (CO) signaling.



A digital circuit (for example, a T1) typically carries PBX signaling in one of two modes: channel associated signaling (CAS) or common channel signaling (CCS). In the CAS scenario, each channel carries signaling information on the same channel that carries the voice or data traffic. However, the CCS approach uses one or more channels in the digital circuit solely for the transport of signaling information, and the other channels carry voice or data.

The Cisco router can interpret some of the signaling types. These standard signaling types include ISDN PRI and Q Signaling (QSIG). Because the Cisco router understands the signaling language of these protocols, the router sends these protocols across the IP network using the “Translate” model. The Cisco router does not understand nonstandard protocols, such as MCDN and DCS+; therefore, the network transparently sends MCDN and DCS+ using the “Transport” model.



Signaling protocols understood by the Cisco router can use the Translate model to carry signaling information across an IP network. Standard protocols include ISDN PRI and QSIG. However, the primary focus of this topic is how Cisco routers support nonstandard PBX protocols.

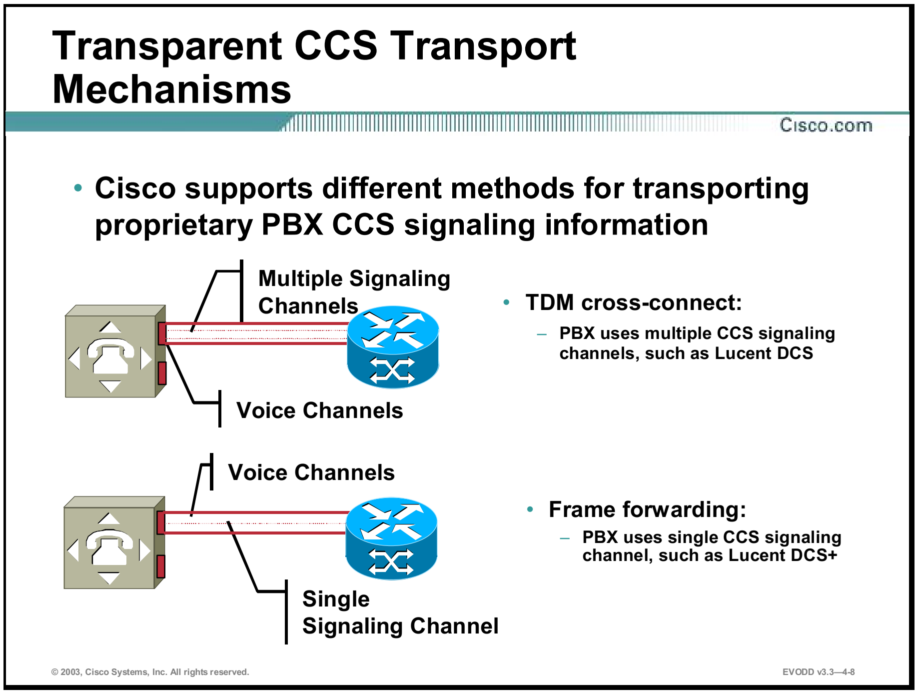
The network uses the Transport model for non-ISDN and nonstandard implementations. The signaling channel(s) must be carried transparently across any intervening network equipment to preserve end-to-end feature operation. For example, the Transport model includes networks using one or more of the following:

- Lucent DCS (traditional or DCS+)
- A Non-Facility Associated Signaling (NFAS) (nB + D) configuration
- Nortel MCDN

The transport mode of operation is the only method for implementing a packet-based network between PBXs when:

- The PBXs use an out-of-band, message-based (data) signaling protocol for controlling and monitoring calls (such as ISDN PRI D channel) that is separated out onto one of the T1/E1 channels, and the protocol stack being used is not supported by Cisco equipment.
- The PBXs use an out-of-band, message-based (data) signaling protocol for controlling and monitoring calls (such as ISDN PRI or QSIG) that is separated out onto one of the T1/E1 channels, and the PBXs use additional proprietary messaging in the signaling channel to implement vendor-specific PBX private network features (such as Nortel Networks ACD

features on top of ISDN PRI). These proprietary messages cannot be interpreted by Cisco equipment.



PBXs traditionally interconnect via tie lines. However, as customers migrate from a PBX environment to an IP telephony environment, they replace the dedicated leased line between their PBXs with connections to local routers, which interconnect over the WAN. These routers must preserve the inter-PBX communication (signaling). Cisco equipment cannot interpret nonstandard or proprietary signaling sent between PBXs. Consequently, it must separate the proprietary signaling channels and transport them over Frame Relay, ATM, or High-Level Data Link Control (HDLC).

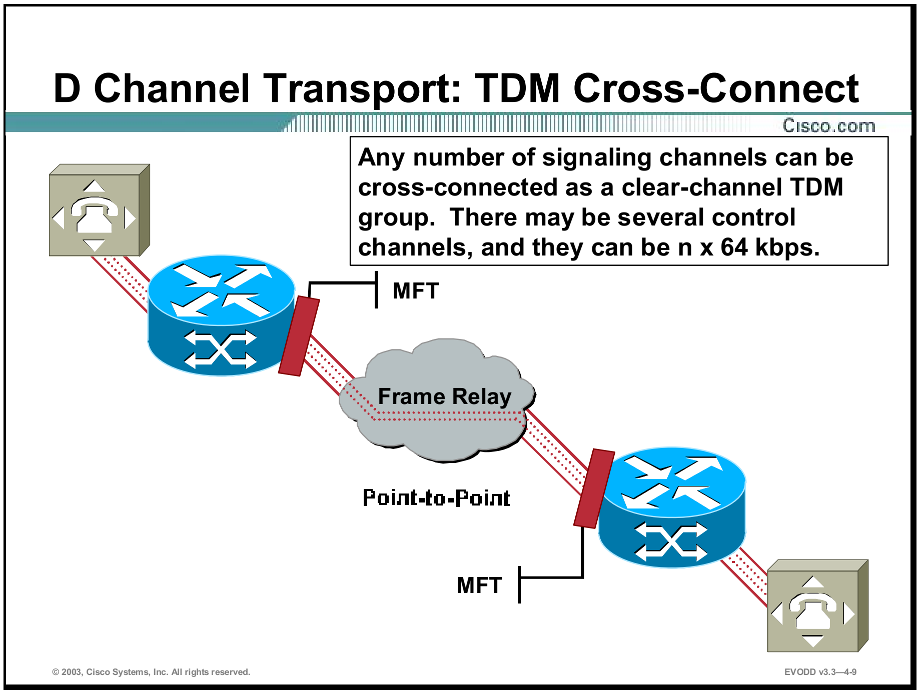
If the signaling is limited to a single D channel, you can use frame forwarding and transport the signaling over Frame Relay, ATM, or HDLC. If the signaling occurs on multiple channels through a single interface, you must use the cross-connect method.

Cross-connect requires use of the multiflex trunk (MFT) interface. Other methods include D channel to ATM circuit emulation service (CES) and D channel to Serial Tunneling (STUN).

Note	Both of these methods require the ability to receive a separate, out-of-band signaling channel from the PBX.

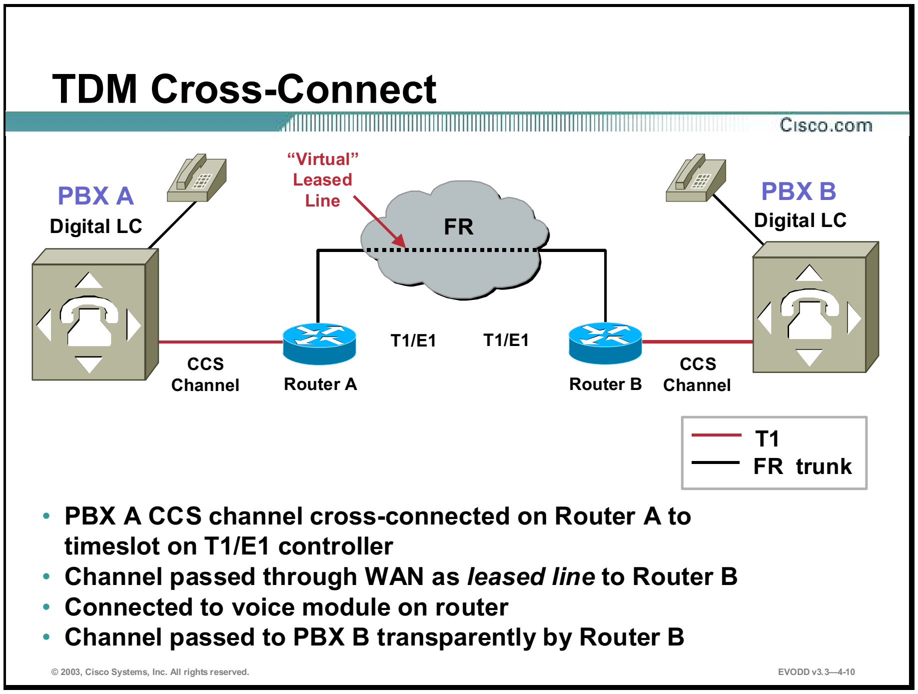
Signaling Channels

This topic examines the cross-connect and frame-forwarding approaches for sending proprietary PBX signaling across a WAN.



Cross-connect with T1/E1 trunks considerations include:

- The connection from the network to the router (for example, 2600/3600 Series) must be via the MFT port.
- This method of signaling is thought of as a point-to-point topology because the method requires a T1/E1 trunk and the access device only supports one such trunk.
- Any channel can be cross-connected as a D channel with this method. There can be n x 64 kbps.
- There is no compression on the D channel; the D channel is bit transparent, not frame mode.
- There is no restriction on which timeslots or how many timeslots can be cross-connected in T1 environments. For E1, the D channel must be the 16th timeslot.

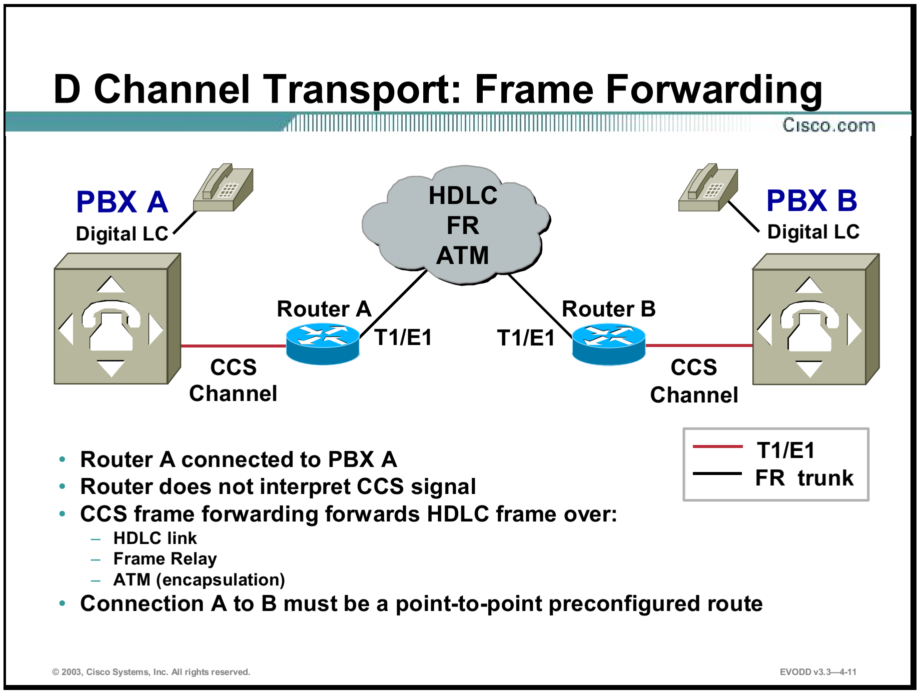


In this figure, the CCS channel from the PBX is cross-connected on the first router, Router A, to a timeslot on the T1/E1 controller. The channel passes through the WAN as a leased line to the second router, Router B. Router B is cross-connected to the digital voice module (DVM) signaling timeslot (timeslot 24 for T1, or timeslot 16 for E1). The channel then passes to the second PBX. The Cisco router then passes through the CCS signal byte stream transparently.

Note	The wide area link must be a point-to-point full or fractional T1 or E1.

TDM cross-connect Cisco routers act transparently for the signaling channel, and is a bit-in, bit-out situation. In other words, no bits are appended, so the router allows the PBX to use the non-HDLC signaling channel that connects the PBX to the router DVM.

You can configure multiple channels for a cross-connect. There are PBXs that use multiple signal channels, and they may require more than one channel for signaling.



Routers, such as the 2600/3600/3700 Series, support CCS frame forwarding, which allows the router voice network module to be connected to a PBX without the router having to interpret CCS signals for call processing. CCS frame forwarding forwards D channel HDLC frames over preconfigured interfaces running HDLC, Frame Relay, or ATM encapsulation. With CCS frame forwarding, the connection between PBXs over the network must be preconfigured. Calls are not routed but follow a preconfigured route to their destination.

Frame forwarding considerations include:

- The D channel transport using frame forwarding is supported over T1/E1, serial, ATM, Frame Relay, or HDLC trunks.

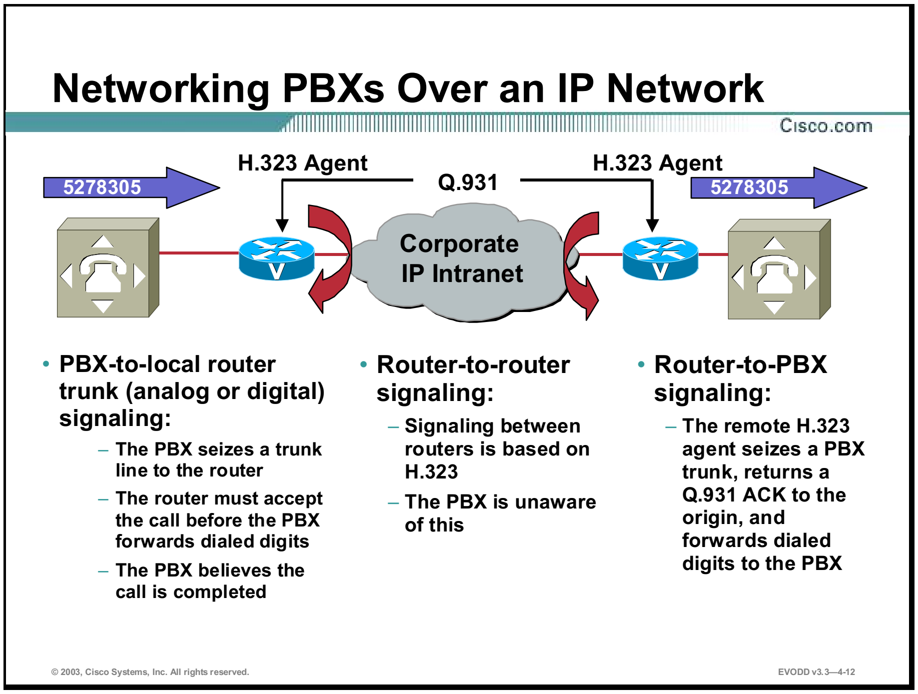
- This method of D channel transport only supports a single D channel per T1/E1. This method does not support multiple D channels per controller, or n x 64 kbps D channels.

- CAS signaling cannot be used on the router's digital circuit for any timeslot using this method.

- Frames on the D channel are not fragmented for transport over the trunk.

Networking

This topic describes methods for networking PBXs over an IP network, such as PBX-to-local router trunk signaling, router-to-router signaling, and router-to-PBX signaling.



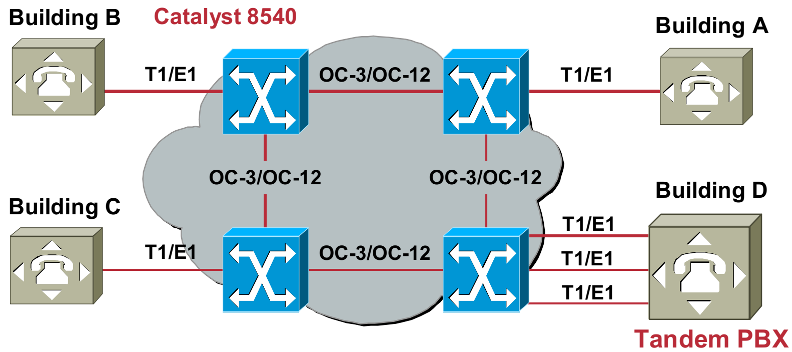
When using PBX-to-local router trunk signaling, the PBX seizes a trunk line to the router. The router must accept the call before the PBX forwards the dialed digits, and then the PBX believes the call is completed.

Note	The local router trunk can be either analog or digital.

With router-to-router signaling, the signaling between routers is based on H.323. However, the PBX is unaware of which protocol is being used.

When using router-to-PBX signaling, the remote H.323 agent seizes a PBX trunk (on the destination PBX), returns a Q.931 ACK to the original router, and forwards the dialed digits to the PBX. If the remote PBX refuses the connection, the originating router (if so configured) falls back to an alternate connection (for example, the PSTN).

Networking PBXs Over ATM CES

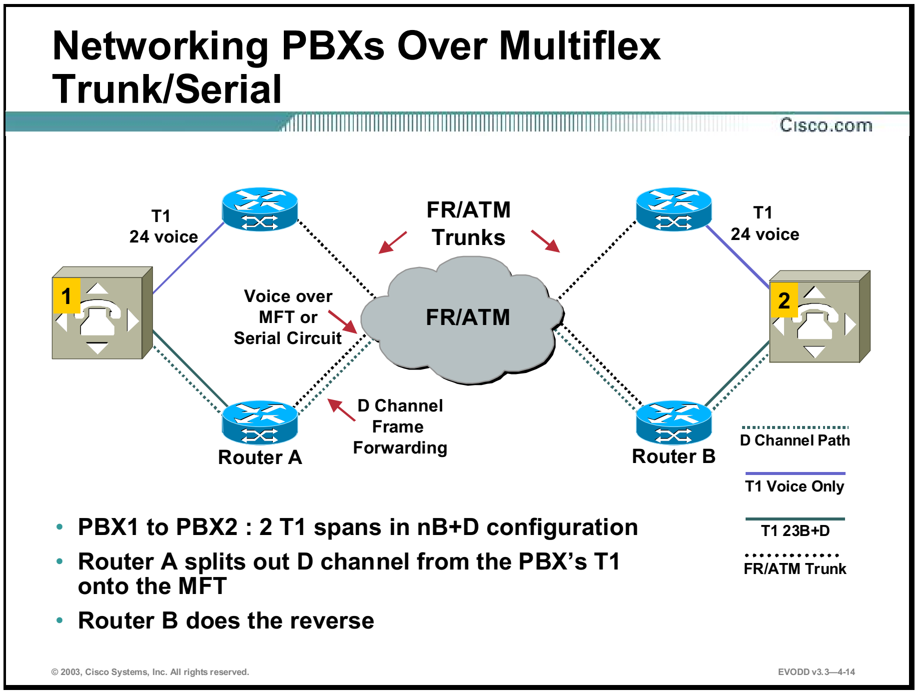


- **Transport is provided via clear-channel T1/E1 links:**
  - **Basically providing TDM transport**
- **All inter-PBX calls are required to traverse the tandem switch**

 EVODD v3.3—4-13

One networking solution incorporates networking PBXs over ATM Circuit Emulation Services (CES). The figure shown depicts a solution where PBXs are interconnected through an ATM cloud, using CES. Notice that the emulated circuits connect back to Building D, which serves as the Tandem PBX. This solution may be the most cost effective in terms of the number of PVCs required, due to the small number of PBXs.

As the size of a campus or MAN grows, the size of the required tandem PBX required also grows, as does the number of PRIs required. Such growth can cause significant expense, and at the same time the tandem PBX adds questionable value and poses a single point of failure.

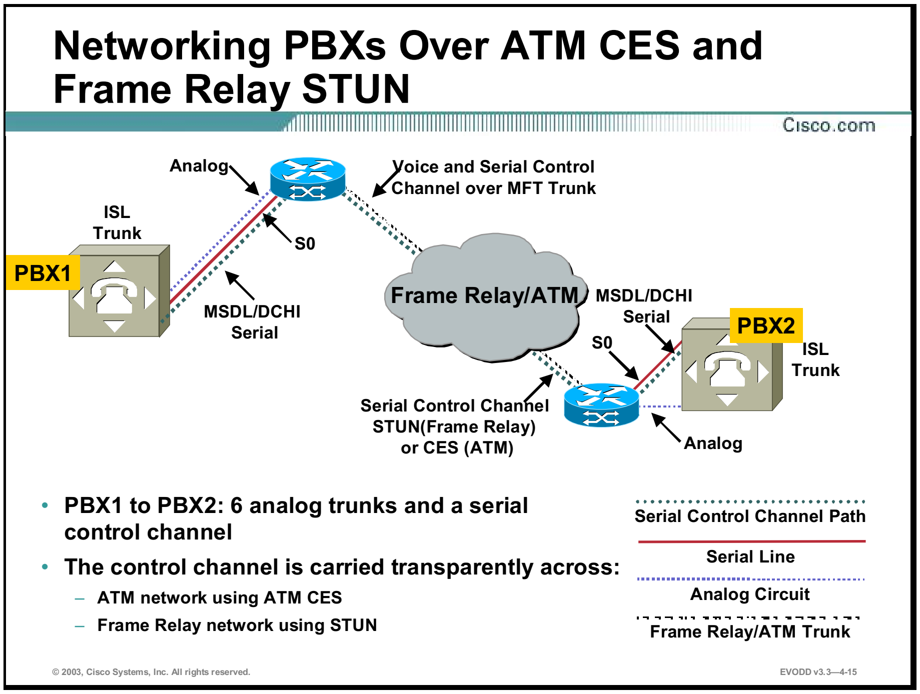


The connections between PBX1 and PBX2 are two T1 spans in an nB+D configuration, or a single D channel, as opposed to having a D channel for each T1. One T1 carries 24 channels for voice, and at the same time the other carries 23 voice channels and the D channel. The D channel can be carried transparently across a Frame Relay or ATM transport network using the D channel frame-forwarding feature of the router (2600/3600/3700/7200 series router).

The source router splits out the D channel from the PBX T1 and puts it on the MFT or a serial connection, in the case of Frame Relay or HDLC only. The receiving router performs this operation in reverse. D channel frame forwarding can use either MFT or serial connections. The MFT provides ATM, Frame Relay, and HDLC connectivity at the same time that the serial options provide Frame Relay and HDLC connectivity.

In addition to the frame forwarding feature, the D channel can be carried by two additional methods: TDM cross-connect and serial hairpin.

TDM cross-connect is expensive and inefficient, but it can be used if the backbone network is Frame Relay or HDLC, and if the connection from the router to the network is via the MFT. In this configuration, the voice bearer channels are carried over Frame Relay or HDLC at the same time that the D channel is carried across a TDM network, or public carrier.

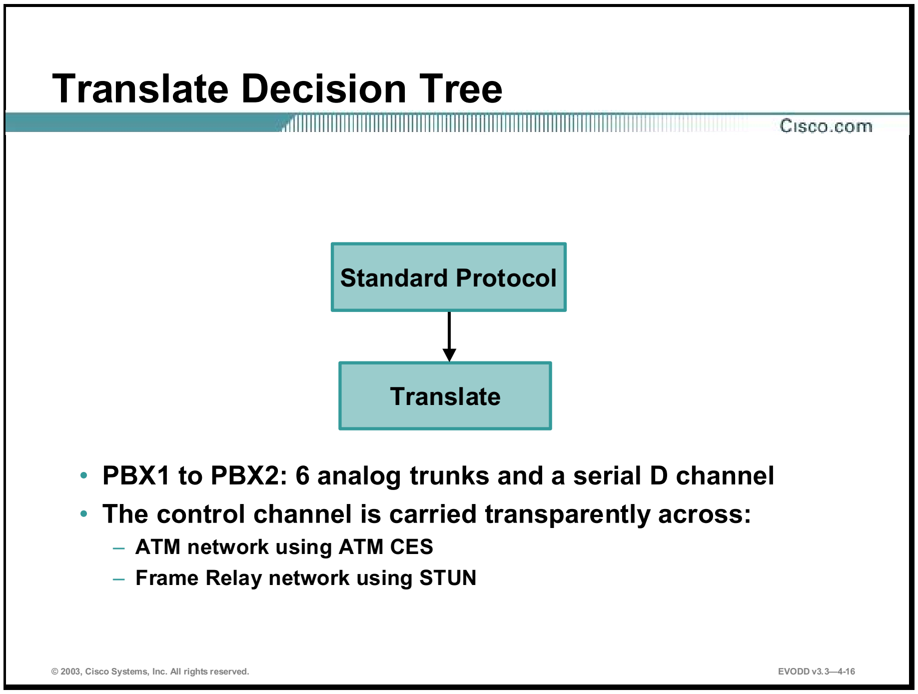


In the figure shown, the connections between PBX1 and PBX2 are six analog trunks and a serial control channel. The control channel can be carried transparently across an ATM network using the ATM CES feature of the router, or a Frame Relay network using the STUN feature.

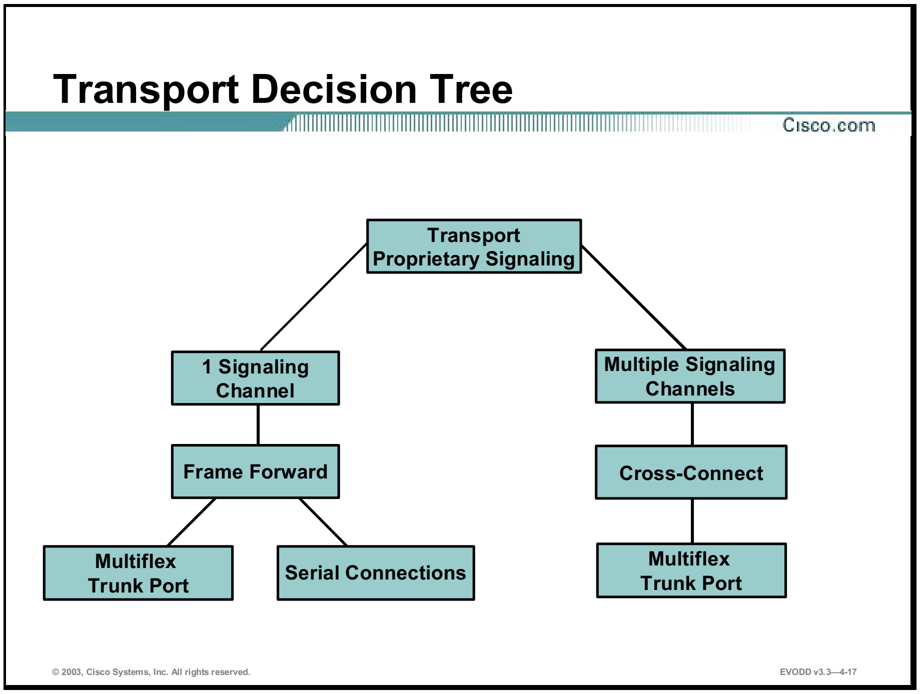
Note The ATM CES feature applies to 2600/3600/3700/7200 Series devices.

Translate versus Transport Decision Trees

This topic describes criteria for selecting a Translate versus a Transport model. If you choose the Transport model, this topic provides guidelines for choosing between a cross-connect and frame forwarding design.



When a standard such as ISDN PRI or ISDN QSIG is used, all networking equipment understands how to interpret the signals. Some PBXs may have a separate serial interface to carry signaling information. To accommodate this design, ATM CES or FR using STUN can be used to carry the signaling information.



When using proprietary signaling protocols, you must establish how many signaling channels are being used and what other protocols are in use. After determining the required number of channels for signaling, you can choose to use the frame forwarding or the cross-connect method.

You use the frame forwarding method if there is only one proprietary signaling protocol. For implementation, you must choose either the MFT or a serial connection.

Note If a backup D channel is required, you need to use the MFT port because the serial ports will already be in use.

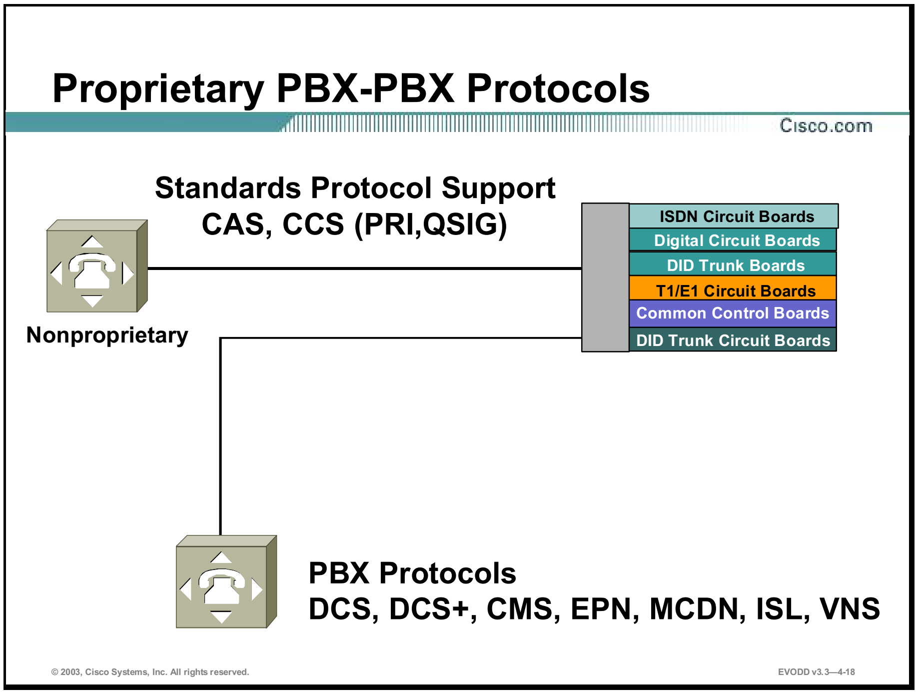
Use cross-connect if there is a need for more than one signaling channel. Multiple cross-connects require an MFT port.

Mapping PBX signaling to a data network requires an understanding of how the PBX handles PBX-to-PBX or PBX-to-CO signaling. Some designs force you to use either the Translate or the Transport model. The translation is a standard ISDN PRI or QSIG configuration that provides interfaces to the network that can be interpreted by standard network equipment. Transport is used for non-ISDN and nonstandard implementations. The signaling channel must be carried transparently across any intervening network equipment to preserve end-to-end feature operation.

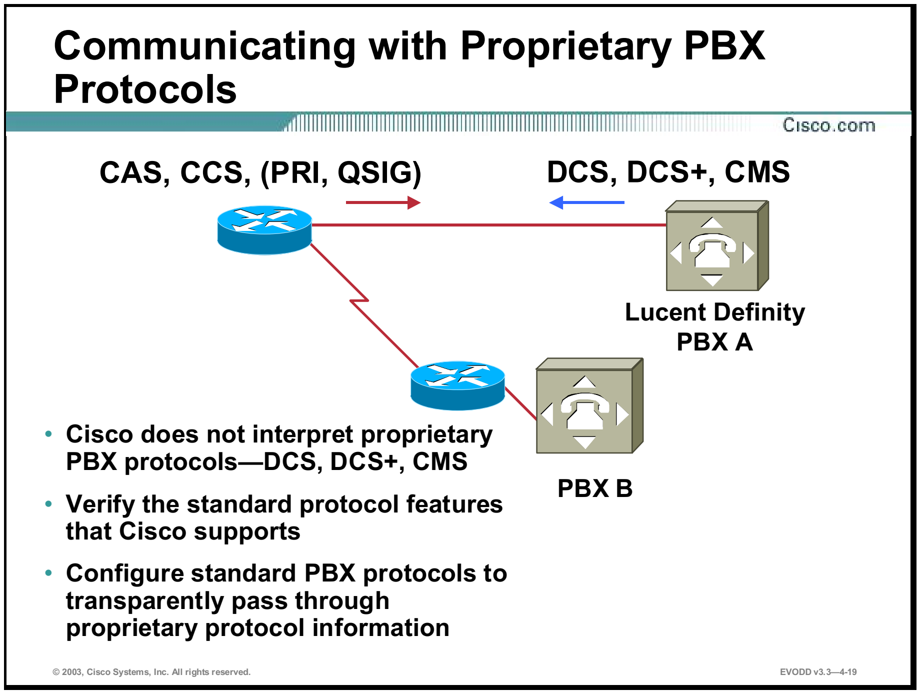
The cross-connect method of transporting D channel signaling over a WAN trunk can only be used in Frame Relay or HDLC mode with a T1/F-T1 or E1 trunk.

Traditional PBXs

This topic discusses design approaches for accommodating PBX-proprietary protocols, including Lucent and Nortel-specific protocols.



PBXs communicate with other PBXs in two ways: using international or industry-standard protocols, such as QSIG, or using PBX-proprietary protocols.



Cisco does not support proprietary PBX protocols, such as DCS, DCS+, and CMS, and some user-visible features for a customer may only be supported by using Lucent protocols.

When considering integration with the PBX, verify the standard protocol features that Cisco supports, and determine if that protocol will be satisfactory to the customer. If not, configure the Cisco router to transparently pass through the proprietary PBX protocols.

Lucent Summary

Cisco.com

Intersystem Protocol	Integrated Transport Method	Integration Platform
DCS	CCS cross-connect TDM cross-connect Circuit emulation Transparent signaling	2600/3600/7200 * 2600/3600/7200 * LS1010 IGX™
DCS+	CCS cross-connect CCS frame forwarding TDM cross-connect Circuit emulation Transparent signaling	2600/3600/7200 * 2600/3600/7200 * 2600/3600/7200 * LS1010 IGX
EPN	Circuit emulation	LS1010

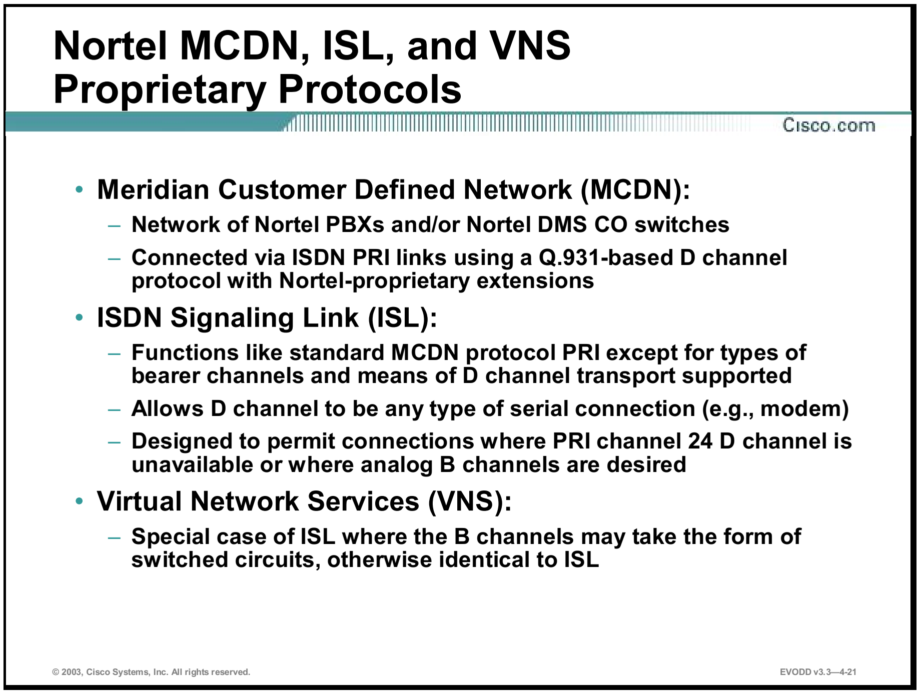
*** Supported on 2600/3600/7200 with Cisco IOS Rel 12.0(6)T or later**

 EVODD v3.3—4-20

Methods for transporting Lucent proprietary protocols across a Cisco IP network are shown in this figure. Notice that frame forwarding is not an option for the Lucent DCS signaling protocol because DCS uses two signaling channels. Therefore, the Cisco cross-connect approach is required.

Lucent uses:

- **DCS**: Traditional DCS is non-ISDN-based. It uses two signaling channels on a T1/E1 circuit.
- **DCS+**: DCS+ is ISDN PRI-based and uses a single channel. Both versions use HDLC-framed data signaling links to communicate information between PBXs to build feature transparency across the private network.
- **EPN**: Expansion Port Network (EPN), which is a circuit emulation “protocol,” is typically used to connect PBXs among buildings. This is accomplished by extending the cabinet, which creates a logical single PBX from two physical locations. The Lucent EPN uses ATM for physical connectivity. Therefore, it could be sent over a Cisco LS1010.



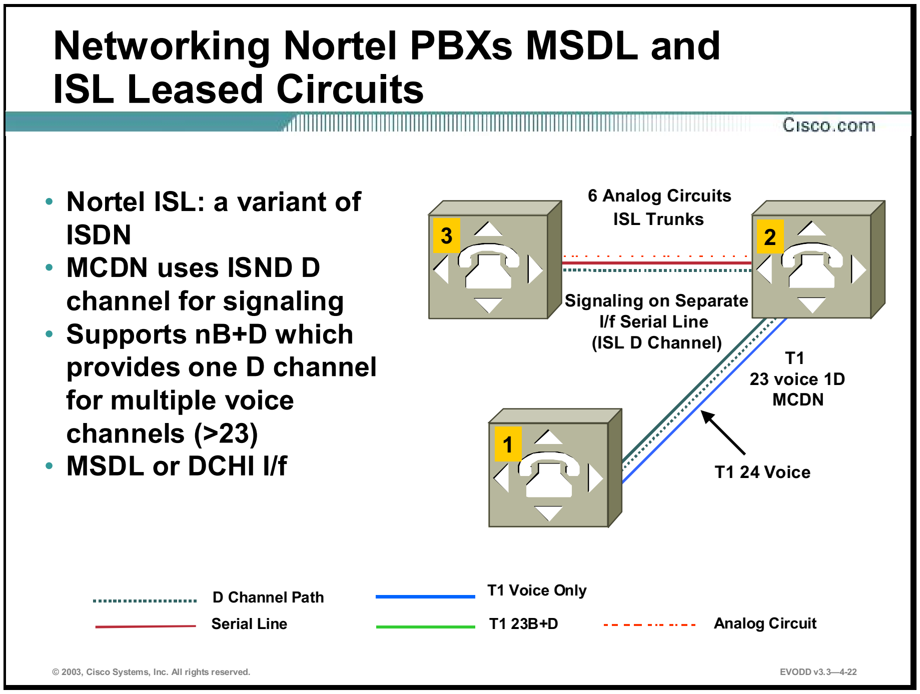
These protocols are Nortel-proprietary extensions to international and national standard ISDN protocols.

Other Nortel protocols include:

- DMS-100 Custom
  - Nortel standard for PRI—Nortel PBX to Nortel LEC COs
  - Provides MCDN protocol to a CO switch
- DMS-250
  - Nortel standard for PRI—Nortel PBXs to Nortel inter-exchange carrier (IXC) COs
  - Provides MCDN protocol to an IXC CO switch
- SL-100
  - Nortel standard for PRI—Nortel PBX to Nortel CO-scale PBX
  - Provides MCDN protocol to an institutional-sized PBX
- National ISDN-2
  - Standardized Bellcore-defined ISDN protocol

- Typically for PBX-to-CO switch in North America

- QSIG
  - European Telecommunication Standards Institute (ETSI) and International Organization for Standardization (ISO) versions of the international QSIG standard
  - Both versions with or without Generic Function (GF) support
  - Typically for brand A PBX to brand B PBX globally



This figure shows the current PRI connections between the PBXs. The connection uses ISL analog circuits and a Multirate Subscriber Digital Line (MSDL) as the D channel.

MSDL is both the functionality and the name of the interface card. The MSDL interface is a card on the Nortel PBX. MSDL functionality is also supported by other Nortel interfaces, such as D-Channel Handler Interface (DCHI), found on the Option 11. The MSDL card has four serial ports, with D channel information for a PRI carried on one of these ports. Each PRI channel in the Nortel PBX requires an MSDL port to support the D channel. The MSDL card, or another interface such as a DCHI card, may provide the port.

The D channel connection must be nailed up (operational). If modems or other on-demand devices are used, they need to link or redial to maintain the circuit because the PBX does not support initiation and restoration of the connection. If the serial D channel fails, established calls continue, but the teardown will be abnormal. The B channels then use analog signaling to restore. MSDL supports Nortel nB+D signaling, which is like the Lucent Non-Facility Associated Signaling (NFAS), in which one D channel can support multiple T1s, allowing all circuits after the first to have a 24th channel of voice.

Note	Design Issue: Look for MSDL and DCHI cards in the customer PBX network. Because they can give you signaling on a single interface out of the PBX, you can run this into a Cisco 2600/3600/3700/7200 series router and use frame forwarding.

Nortel

Cisco.com

Intersystem Protocol	Integrated Transport Method	Integration Platform
MCDN (ISL &VNS)	CCS cross-connect CCS frame forwarding TDM cross-connect Circuit emulation Transparent signaling	2600/3600/7200 * 2600/3600/7200 * 2600/3600/7200 * LightSpeed® 1010 IGX

*** Supported on 2600/3600/7200 with Cisco IOS Rel 12.0(6)T or later**

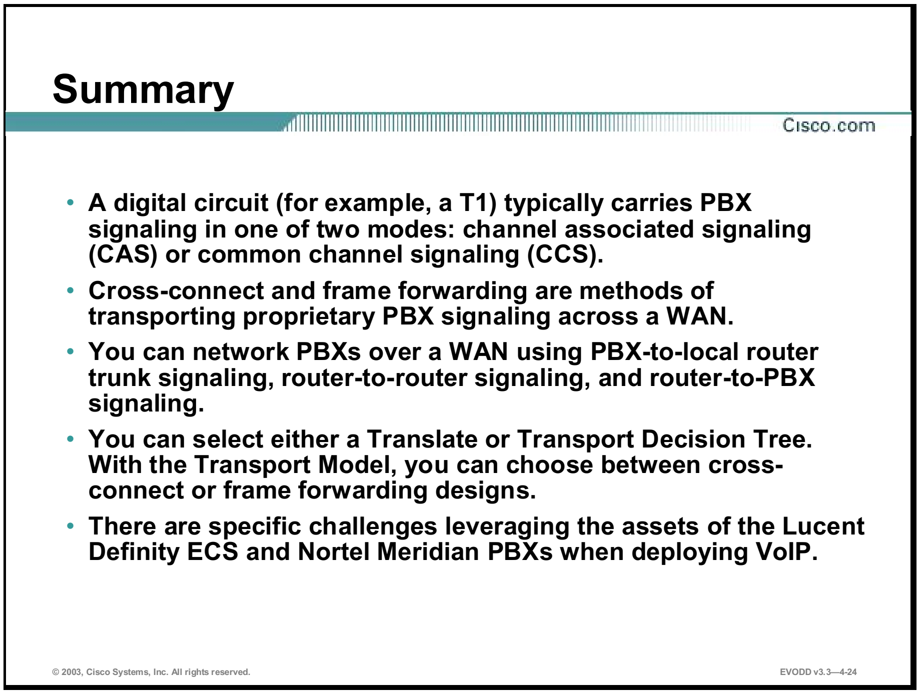
EVODD v3.3—4-23

This figure shows methods for transparently sending Nortel protocols across a Cisco IP network.

Note CCS frame forwarding can be used because Nortel uses a single channel for signaling, unlike the Lucent DCS signaling protocol.

Summary

This topic summarizes the key points discussed in this lesson.



Next Steps

After completing this lesson, go to:

- Dial Plans lesson

References

For additional information, refer to these resources:

- Configuring Support for PBX Signaling Formats: http://www.cisco.com/univercd/cc/td/doc/product/access/multicon/3810soft/swcfg/pbxcfg.htm

- Nortel Meridian PBX and Cisco CallManager (CCM) Integration: http://www.cisco.com/warp/public/788/AVVID/meridianpbx_cm.html

- Configuring Support for PBX Signaling Protocols: http://www.cisco.com/univercd/cc/td/doc/product/software/ios121/121cgcr/multi_c/mcprt1/mcdpbxsg.htm

Lesson Review

This practice exercise reviews what you have learned in this lesson.

Q1) Which of the following best describes the Cisco Transport Model for forwarding PBX signaling?

A) The Transport model is used when the Cisco router can interpret the PBX signaling, such as QSIG.

B) The Transport model is used when the Cisco router cannot interpret the PBX signaling, such as DCS+.

C) The Transport model is only used when the PBX uses a single CCS channel.

D) The Transport model is only used when the PBX uses multiple CCS channels.

Q2) Which of the following is required to transport proprietary PBX signaling using multiple signaling channels?

A) Cross-connect

B) Frame forwarding

C) CES

D) STUN

Q3) Which of the following is a PBX interconnection requiring a call to pass through one or more intermediate PBXs on the way from the source PBX to the destination PBX?

A) Tunnel PBX

B) CES PBX

C) Tandem PBX

D) CCS PBX

Q4) Which of the following signaling protocols can use the Cisco Translate model to forward PBX signaling?

A) DCS

B) DCS+

C) MCDN

D) QSIG

Q5) Which proprietary protocol CANNOT use frame forwarding?

A) DCS

B) DCS+

C) MCDN

D) ISL

Dial Plans

Overview

Dial plans serve as the logic for system configurations, and describe the handling of calls that do not reach their intended target. The enhanced Cisco CallManager (CCM) dial plan allows for greater scalability, flexibility, and ease of use when compared to traditional dial plans. It is important to understand dial-plan tables, class of service groups, and different approaches for migrating to a dial plan for Voice over IP (VoIP). A major dial-plan consideration is the support of 911 services, including the PBX and Cisco IP telephony approaches to E911 support. To make the VoIP migration transparent to the end user, you must also maintain services, such as messaging, using the Cisco messaging solution.

Importance

Dial plans are the core component in a telephony solution. It is essential that you understand what they are and how to design them. You must also understand how traditional dial plans work because you will need to integrate your telephony solution with the existing PBX.

Objectives

Upon completing this lesson, you will be able to:

- Describe the functionality and benefits of a dial plan
- Describe how partitions and calling search spaces allow CCM to restrict calls
- Describe how translation patterns enable CCM to accommodate overlapping dial plans
- Identify the dial plan issues associated with the migration from a legacy PBX system to an IP telephony solution
- Describe the history of basic 911 and E911 services
- Identify the PBX and Cisco IP telephony approaches to E911 support
- Describe a typical voice-messaging system

- Describe the Cisco solution for integrating a legacy voice-mail system into a unified messaging system

Learner Skills and Knowledge

To fully benefit from this lesson, you must have these prerequisite skills and knowledge:

- Fundamental understanding of legacy PBX signaling

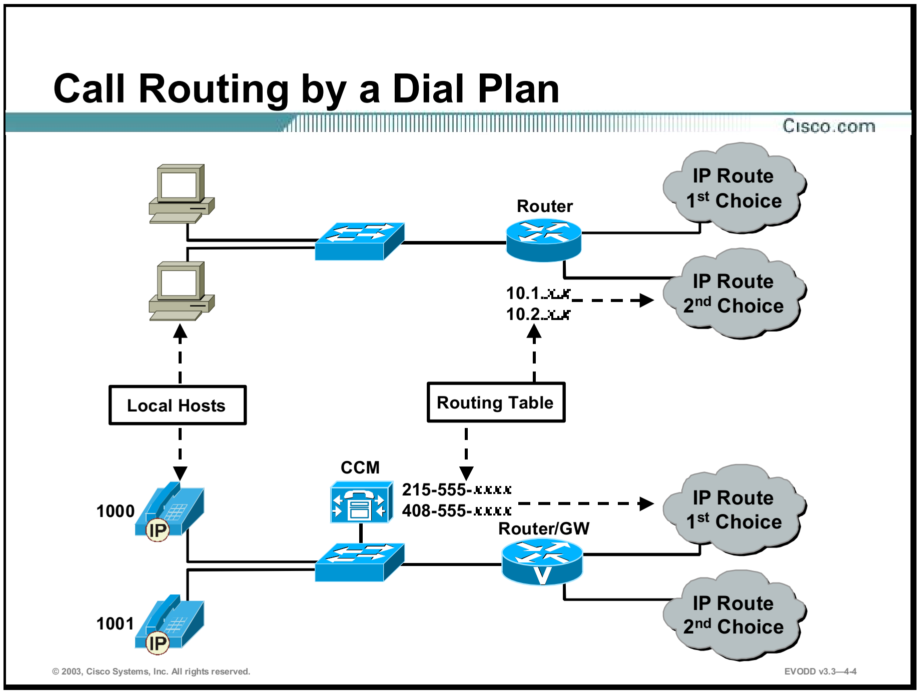
Outline

This lesson includes these topics:

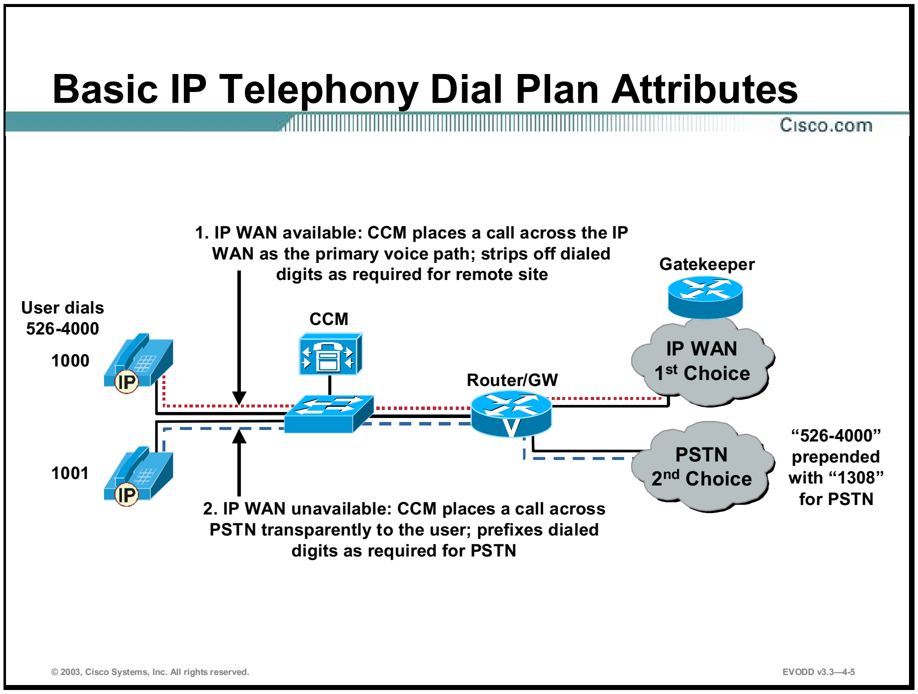
- Overview
- Dial Plan Functions
- Calling Restrictions
- Translation Patterns
- Migration
- 911 History
- Handling 911 Issues
- Voice Messaging
- Cisco Unified Messaging Solution
- Summary
- Lesson Review

Dial Plan Functions

This topic describes the functions of a dial plan and the dial plan components that you can configure in Cisco CallManager (CCM). This topic also explains how a CCM dial plan operates.



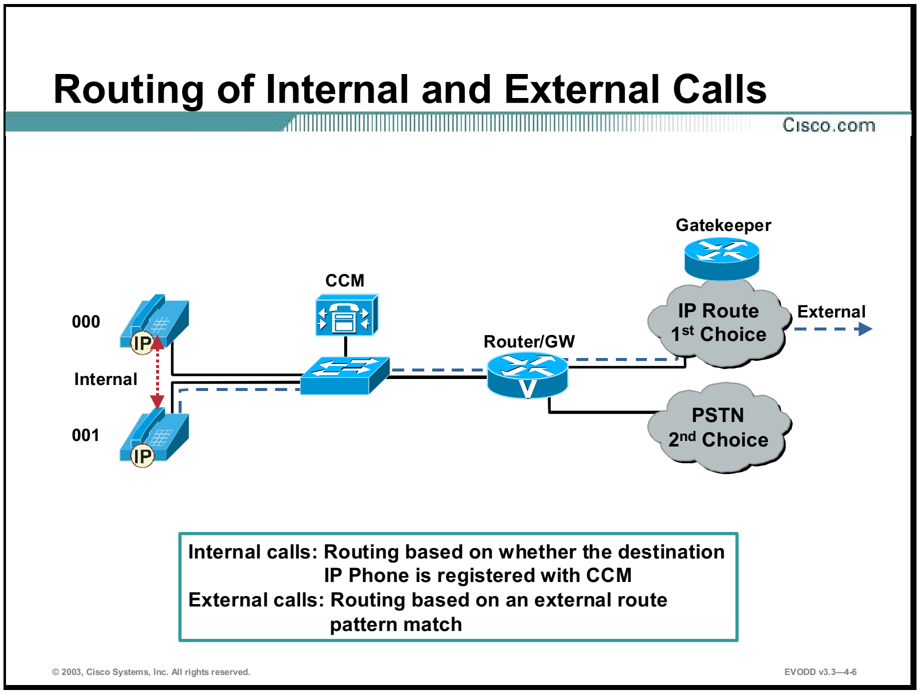
A well-designed dial plan is a vital component of any IP telephony network, and all of the other network elements rely on this plan in some fashion. Essentially a dial plan is IP routing for voice calls, as shown in the figure. IP routing and IP telephony dial plans perform similar functions in that both provide endpoint addressing, alternate path routing, and policy enforcement restrictions.



A fundamental attribute of a dial plan is its ability to route a call transparently to the dialed destination, irrespective of the available physical voice path, as shown here.

IP telephony dial plans provide these primary benefits:

- Transparent routing of calls to dialed destinations, with alternate path routing if there are multiple paths to a given destination
- Device redundancy in the event of a failure in a network element, such as a gateway, voice-mail server, or application
- Calling policies that control the destinations that selected IP Phones and users can dial
- Digit manipulation to strip or add digits to the dialed E.164 number, based on the voice path taken, either over the IP WAN or the Public Switched Telephone Network (PSTN)
- Partitioned dial plans that support overlapping dial plans



When CCM routes calls, it distinguishes between internal and external calls, as shown here. CCM routes internal calls based on the registration of the IP Phone with CCM. CCM routes external calls based on the route patterns that you configure in CCM.

When an IP Phone dials a number to place a call, CCM first analyzes the dialed number to determine if it is the number of a registered IP Phone. If the dialed number matches the number of a registered IP Phone, then CCM places the call, as long as the class of service (CoS) configuration allows for the call to be placed. (Classes of service for IP telephony are more accurately termed calling restrictions.)

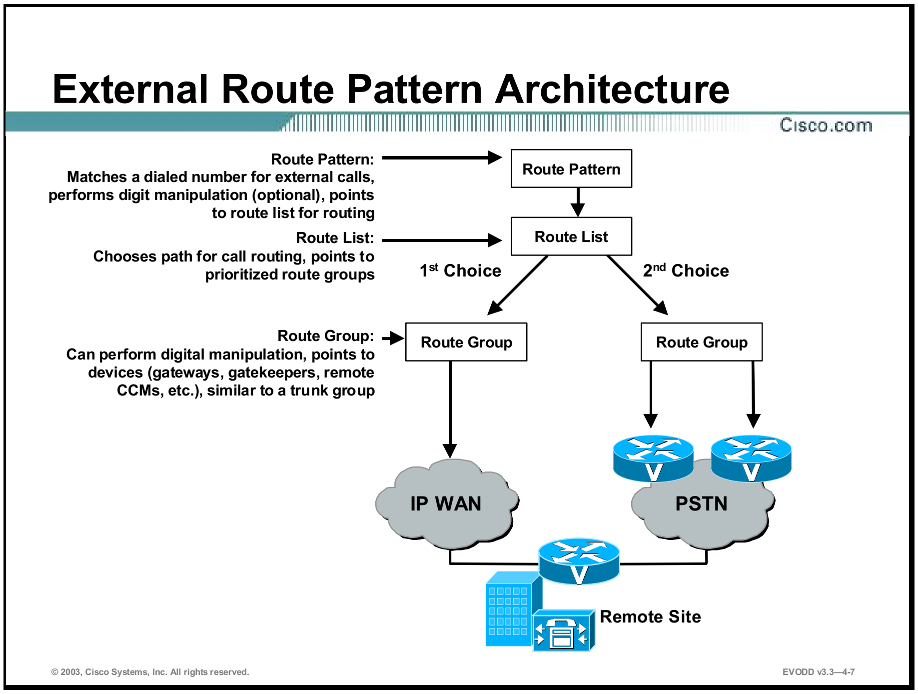
IP routing is very similar to that of two IP hosts on the same subnet sending packets to each other; in this case, the router does not have to look up the destination in its routing table.

When the IP Phone registers with CCM, it dynamically updates CCM with the new IP address while maintaining its same telephone number. Other devices that register with CCM in this and maintain a directory number include Cisco IP SoftPhones and analog telephones that are attached to MGCP gateways.

You can configure the internal dial length (the number of digits dialed) for internal calls.

When an IP Phone dials a number that does not match a registered IP Phone, CCM assumes that the call is an external call and looks in its external route table to determine where to route the call.

CCM uses route and translation pattern tables to determine where, and how, to route a call, which is very similar to the IP routing concept of static IP routes.



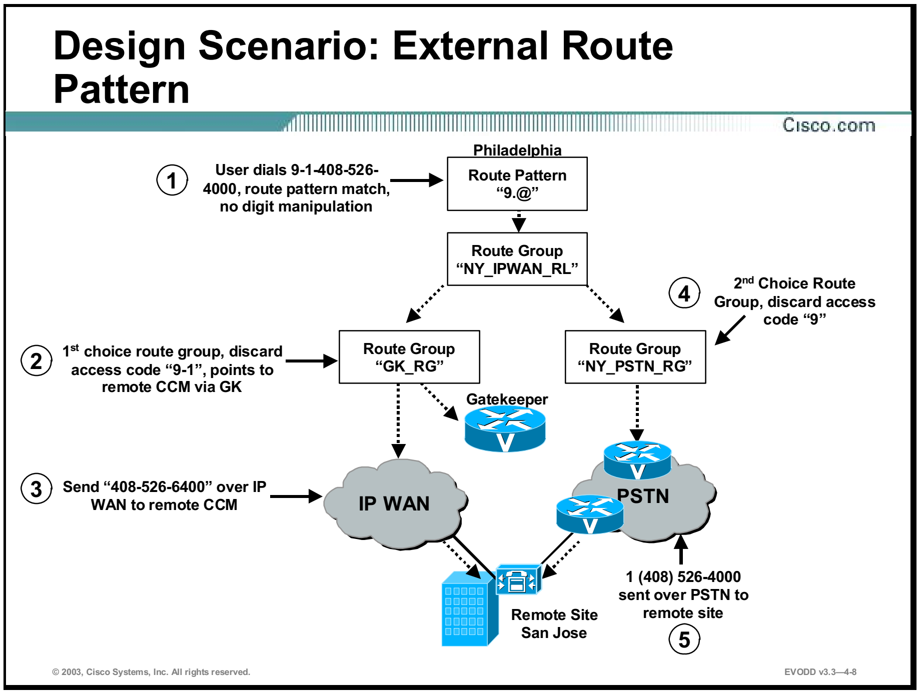
The CCM external route pattern construct is based on a three-tiered architecture that allows for multiple layers of call routing and digit manipulation. CCM searches for a configured route pattern that matches the external dialed string and uses the route pattern to select a corresponding route list. The route list is a prioritized list of the paths available for the call. The available paths make up route groups and are very similar to trunk groups in traditional PBX systems. A route pattern is best thought of as a static route with multiple paths. The figure shown depicts the three-tiered architecture of CCM dial plans.

In most cases, the route pattern directs calls to a PSTN gateway or, in the case of an IP WAN call, to an H.323 gatekeeper for delivery to a remote CCM and gateway. In addition to providing multiple prioritized paths for a call, the CCM dial plan can provide unique digit manipulations for each of these paths, based on the network requirements. Digit manipulation involves appending or removing digits from the originally dialed number to accommodate dialing habits or gateway needs. For instance, for a given dialed number, Carrier A may require 7 digits, whereas Carrier B may require 10 digits. CCM can transparently perform this manipulation.

Route groups function as trunk groups that provide access to gateways. Route lists effectively provide path redundancy by defining a prioritized list of route groups.

A typical use of a CCM route plan is to route a dialed number to the IP WAN first, if the IP WAN is down or has insufficient resources, then CCM will route the dialed number to the PSTN. You can configure Call Admission Control (CAC) to indicate that the IP WAN cannot accept the call, thus prompting the dial plan to select an alternate route for the call.

Note	Specifying alternate paths to dialed destinations applies only to route patterns and not to destinations that are IP Phones in the same CCM cluster.



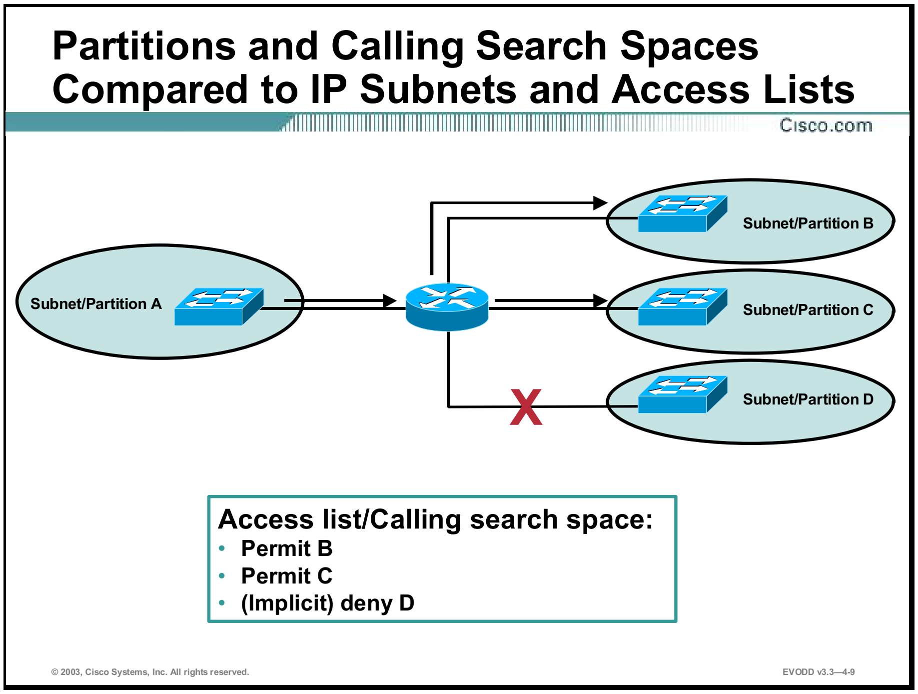
In the graphic shown, the route pattern 9.@ points to a route list that selects the prioritized paths for the call. In this case, the route list NY_IPWAN_RL attempts to send the call across the IP WAN route group GK_RG as the first choice or across the PSTN route group NY_PSTN_RG as the second choice. The route groups then point to individual devices, such as VoIP gateways.

Digit manipulation (stripping or adding digits) can be done in both the route pattern and route group. However, Cisco recommends that you perform digit manipulation in the route group (viewed from within the route list) because digit manipulation requirements may vary with the selected voice path.

Note The route pattern can point directly to a gateway for routing calls, but Cisco strongly recommends you use the complete route pattern, route list, and route group construct, because it provides the greatest flexibility for call routing, digit manipulation, and future dial plan growth.

Calling Restrictions

This topic examines how CCM provides the ability to restrict calls.



Users can be grouped into communities of interest on the same CCM, yet share the same gateways and have overlapping dial plans. This capability helps support multi-site IP WAN deployments with centralized call processing and multi-tenant deployments. CCM uses partitions and calling search spaces to implement these calling restrictions.

Partitions are groups of devices with similar accessibility. The devices in a partition are all of the entities associated with the directory numbers that users can dial, and include IP Phones, directory numbers, and route patterns. When naming a partition, choose a name that has some meaningful correlation to the group of users that the partition represents. For example, for users located in Building D in San Jose, you could create a partition called SJ-D.

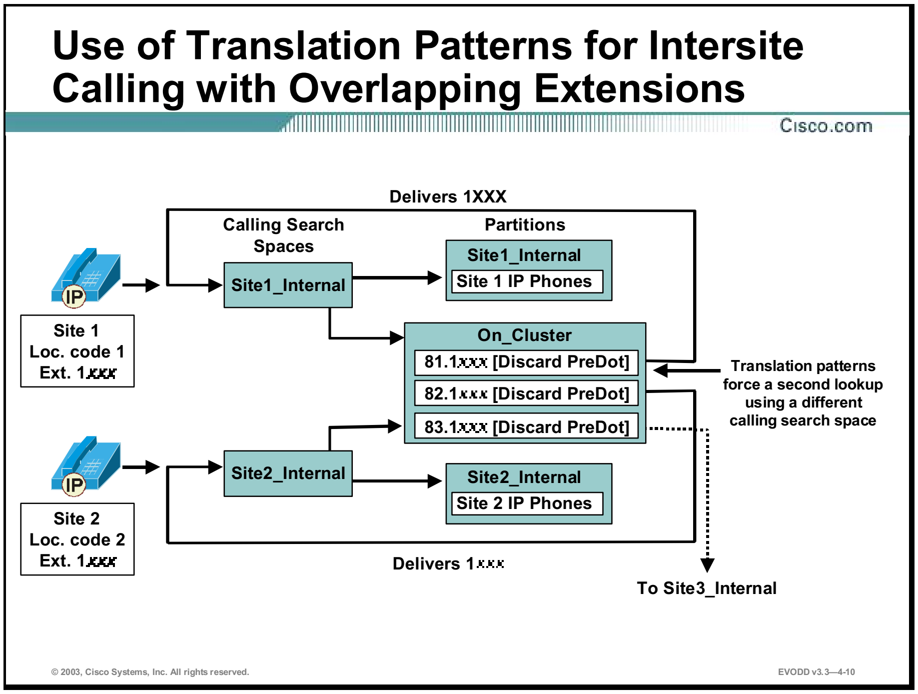
A calling search space defines the users can access each partition. A calling search space serves the same function as an access list for a subnet. You can assign calling search spaces to devices that can initiate calls, such as IP Phones, Cisco IP SoftPhones, and gateways. The calling search space determines the destinations that member devices are allowed to call.

Members of a calling search space can access only the partitions that are listed in the calling search space. Attempts to dial a directory number in a partition outside of the calling search space will fail, and the caller will hear a busy signal. If there are overlapping route patterns in different partitions within the same calling search space, CCM chooses the matching route pattern of the partition listed first in the calling search space. Otherwise, the closest match rules apply.

Partitions and calling search spaces are analogous to routers with access lists, as illustrated in the above graphic. The partition is similar to an IP subnet, where you assign users, and a calling search space is similar to an inbound access list that determines which subnets the users can reach.

Translation Patterns

This topic examines how CCM provides digit translation, which is the ability to transform a called or calling number into another number.

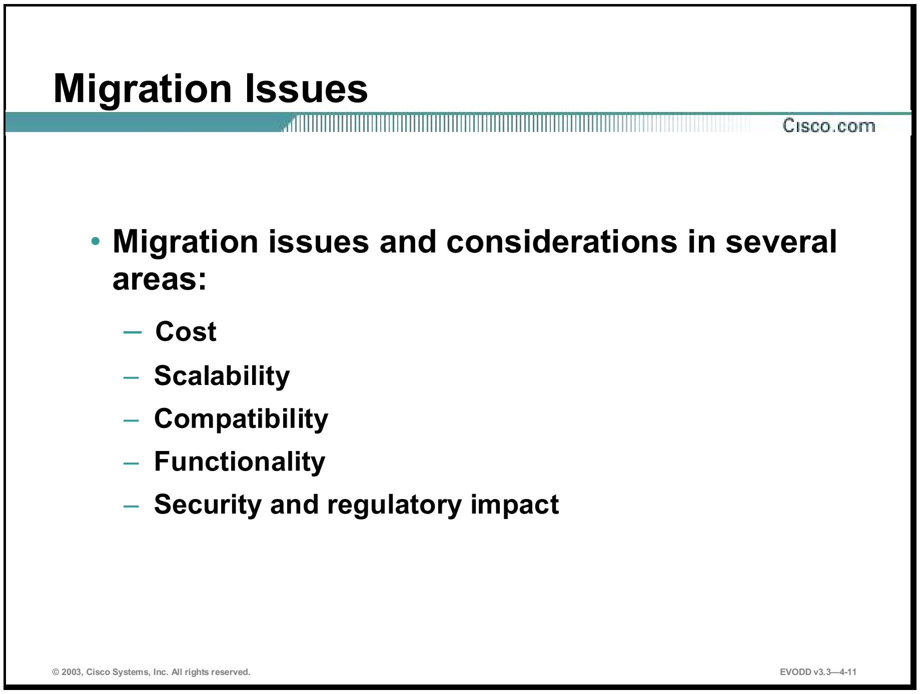


Digit translation can be used on internal and external calls, either inbound or outbound. You typically use digit translation in situations such as routing a call to an attendant at extension 1111 whenever a user dials 0, routing a call to a recorded message if the call tries to reach an unassigned Direct Inward Dialing (DID) number, or routing all external inbound calls to an interactive voice response (IVR) system.

The above graphic shows how translation patterns provide intersite dialing in the presence of overlapping extensions. For instance, if both Site 1 and Site 2 have extensions in the range 1*xxx*, partitions must be used to separate their overlapping directory numbers. To allow communication between sites, a set of translation patterns (one per site) is defined in a common partition that is visible to all users. When the user with extension 1000 at Site 1 wishes to dial the user with extension 1000 at Site 2, the user at Site 1 first dials the intersite access code (8 in this example), followed by the destination site code (2), followed by the four-digit extension of the other party (1000). This string, 821000, matches a translation pattern in the On_Cluster partition, which strips 82 and delivers 1000 to the Site2_Internal calling search space, which has access to the directory numbers at Site 2.

Migration

This topic discusses the potential issues in a time-division multiplexer (TDM) to IP telephony migration, including legacy dial-plan functions and features in the migration.



One way of identifying the many issues and considerations for a legacy to IP telephony migration is to create several categories that can be populated with issues and considerations. These categories can include:

- **Cost:** For distance-sensitive off-net calls, terminate the calls as close to the destination as possible. Use IP wherever feasible. Leverage flat-rate circuits and services by locating gateways within those areas. Consider the trade off between circuit bandwidth needs and cost. Consider the need to collect use statistics for cost allocation.

- **Scalability:** Trade off the need for devices to each maintain full directories and/or dial-plan configuration with the option for devices to ask other devices for dial-plan information. Plan now for expanded number needs, such as what kind of devices will need extension-like numbers in the IP environment. Reserve resources for voice mail, Meet-Me, Telephony Application Programming Interface (TAPI), and call park. Consider the dial-plan implications of telecommuters and teleworkers; for example, do they need virtual, on-net services?

- **Compatibility:** Consider interworking constraints that impact the dial plan, for example, interfacing with legacy PBXs and voice mail. Use a consistent dial plan among enterprise systems. Map instruments (extensions), to DIDs. Use consistent codes that are visible to the user, including operator, information and directory service, and extended features. Reserve enterprise-wide ranges for existing services including voice mail, Meet-Me conferencing, and other devices, such as paging gateways and call parking. Consider the impact of universal number portability on number validation tables.

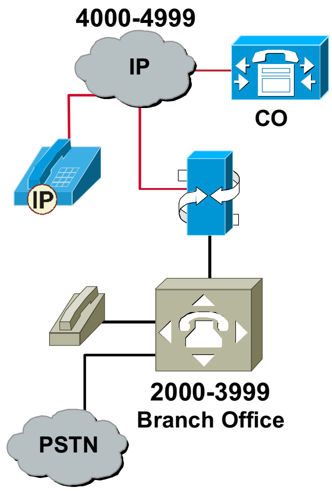
- **Functionality:** Reserve resources for new applications, such as TAPI. Provide logic for handling call failures. Implement real redundancy that recognizes the sensitivity of voice to the business. Design flexibility for potential constraints, such as flat rate, and not per-call, long distance. Design for painless administration of the legacy and IP telephony systems.

- **Security and regulatory impact:** Consider the impact of regulations and telecommunications carrier competition on the dial plan. Consider the role of security, including encryption, and privacy.

Dial Plan Migration Options: Consistent Dial Plans

Consistent dial plan between systems:

- **Map both ranges to DIDs**
- **Uniform codes for:**
  - **Outside line**
  - **Via PBX**
  - **Attendant**
  - **Other features visible to users**

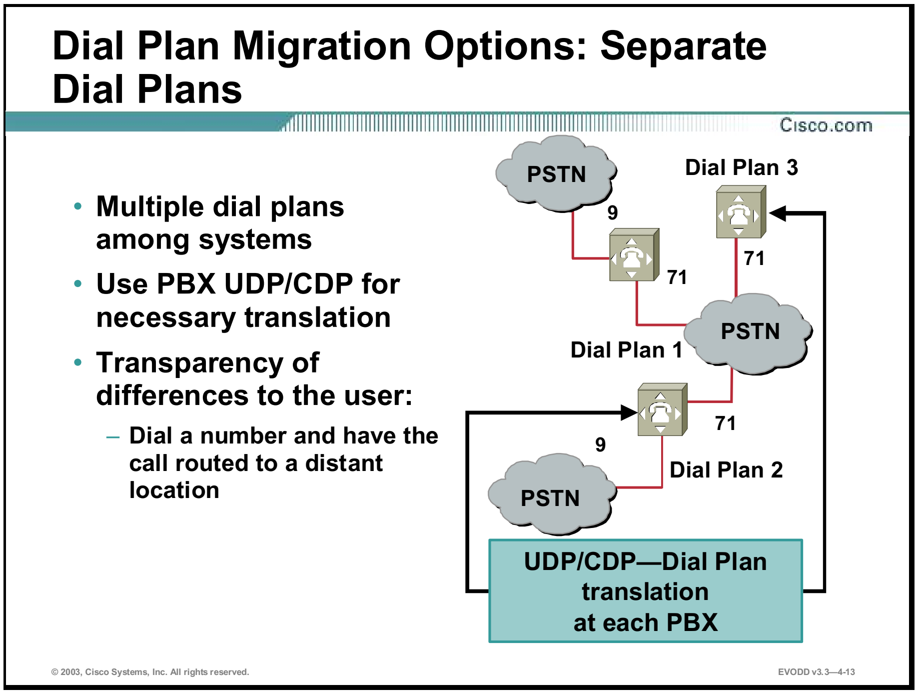


EVODD v3.3—4-12

Many migration approaches are sensitive to dial-plan issues. The most straightforward way to perform a migration is to use a consistent numbering plan. The numbering plan should map both ranges of internal extension numbers to direct inward dialing.

Great difficulty results from moving a user from their PBX system to an IP system if they lose needed functionality, or if their friends, family, and business colleagues can no longer call them directly. For example, choose the range 4000 to 4999 for IP telephones, and the range 2000 to 3999 for regular analog telephones. One recommendation is to move blocks of addresses at a time. Block moves might be moving all of the 2100 extensions, then the 2200 extensions, and then the 2300 extensions.

There must be consistency with legacy codes for an outside line, which is typically a 9 in North America, and for the operator and/or attendant, perhaps a 0.



Another high-level migration approach is to use the power of translation capabilities to permit the use of multiple dial plans. While PBXs are still available, their Uniform Dialing Plan/Coordinated Dialing Plan (UDP/CDP) capability can be used to perform translations and provide transparency to the user.

Entities, such as a branch offices and central headquarters that are networked across a WAN typically have separate DID ranges. Cisco, for example, has offices in Research Triangle Park, North Carolina, and in San Jose, California. The same DID digits cannot be obtained for both offices because the area code 408 does not exist in both places.

Access codes and tail-end drop off can be used. When calling from San Jose to someone in North Carolina external to Cisco, the call hops off the network in RTP using a local gateway. This gateway might be an ISDN circuit to the PSTN.

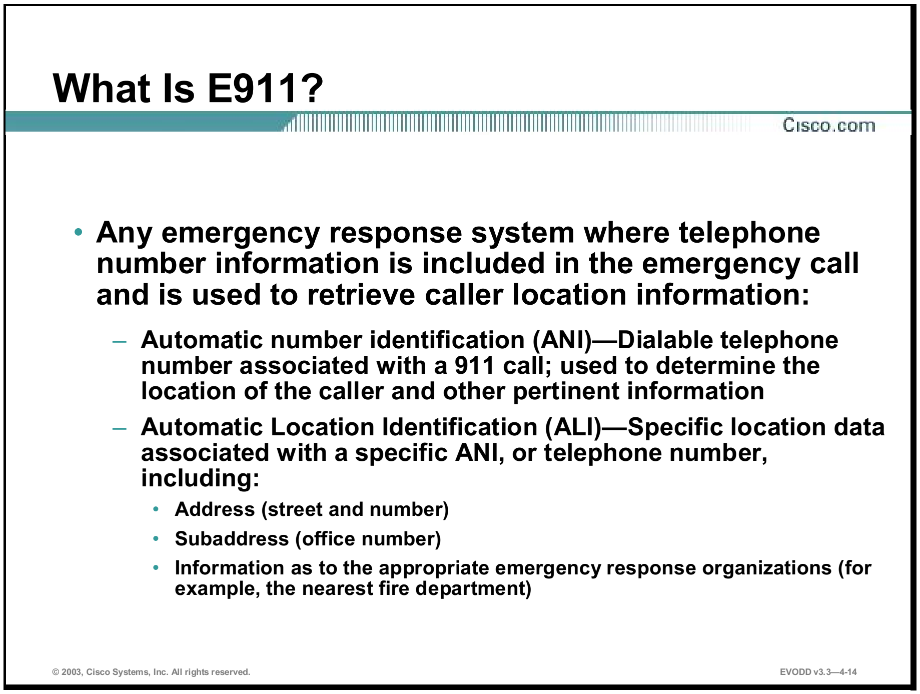
Note A tail-end drop off is sometimes called onnet to offnet.

The dial plan can make use of a local attendant, or if the local office is too small or is geographically remote, the attendant can be in another location.

A common practice in companies, regardless of the size of the office, is to use four-or five digit extensions that map to the DID digits.

911 History

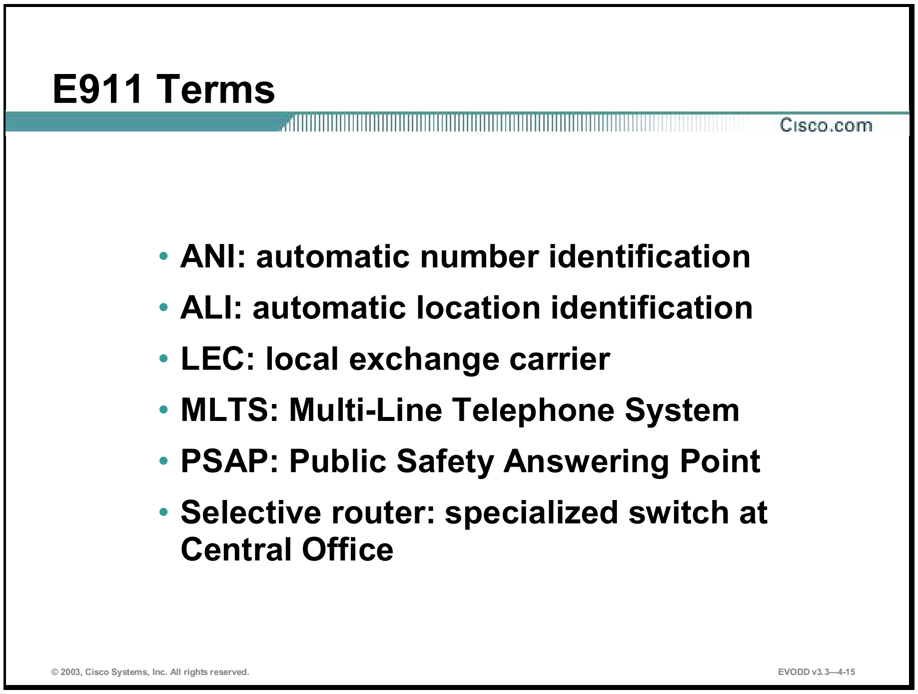
This topic discusses the history of the basic 911 service. Basic 911 emergency service began in the United States in 1968, and was designed as an easy-to-remember access code for emergency services. Enhanced 911 (E911) was developed to correct an inherent weakness within the basic service.



With basic 911 service, each call is typically processed by the local exchange carrier (LEC) and routed from the central office to the nearest Public Safety Answering Point (PSAP). The attendant answering the emergency 911 call collects information about the person calling, the location and nature of the emergency, and then tries to give instructions to the caller, while emergency personnel are dispatched to the location. Relying on the caller to provide this information is the inherent weakness of basic 911 service.

The E911 design eliminates the need for the caller to provide location information. With E911, automatic number identification (ANI) uses the telephone number to find the address where the call is being made. This address can be used to route the call to the closest PSAP. This process eliminates the need for the PSAP to locate the closest emergency help.

The LEC uses the telephone number to search its database and locate the name and exact location of the person to whom the telephone is registered. This feature, called Automatic Location Identification (ALI), minimizes errors in reporting the location of the emergency. In many cases, the person calling for help does not know the address, is unfamiliar with that location, or may even make a mistake in providing the address. In such cases, automatically providing the address is crucial.

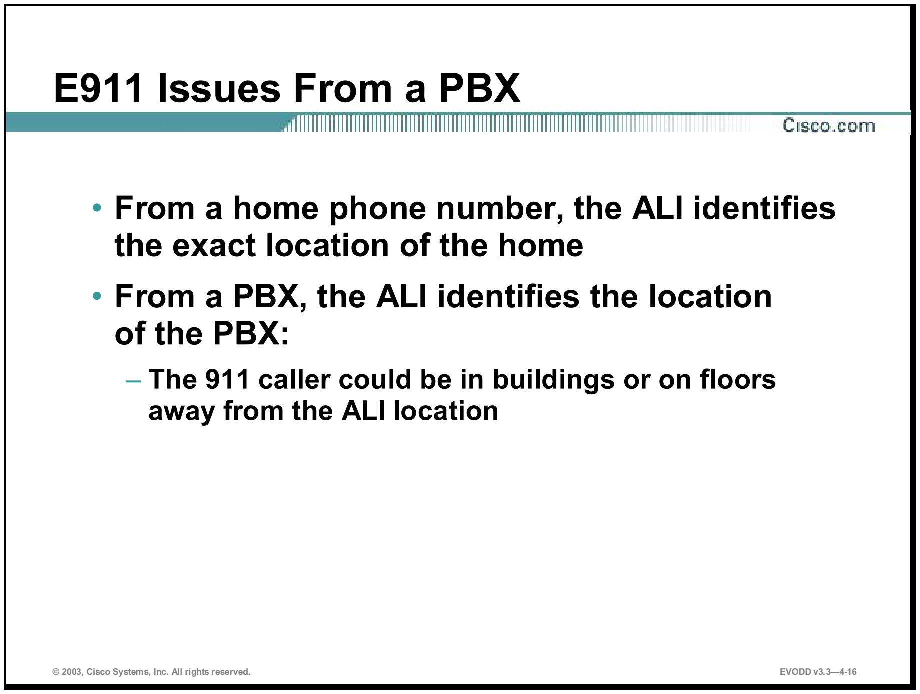


The following is a list of common terms associated with E911:

- **ANI:** automatic number identification
- **ALI:** automatic location identification
- **LEC:** local exchange carrier
- **MLTS:** Multi-Line Telephone System
- **PSAP:** Public Safety Answering Point
- **Selective router:** specialized switch at the central office

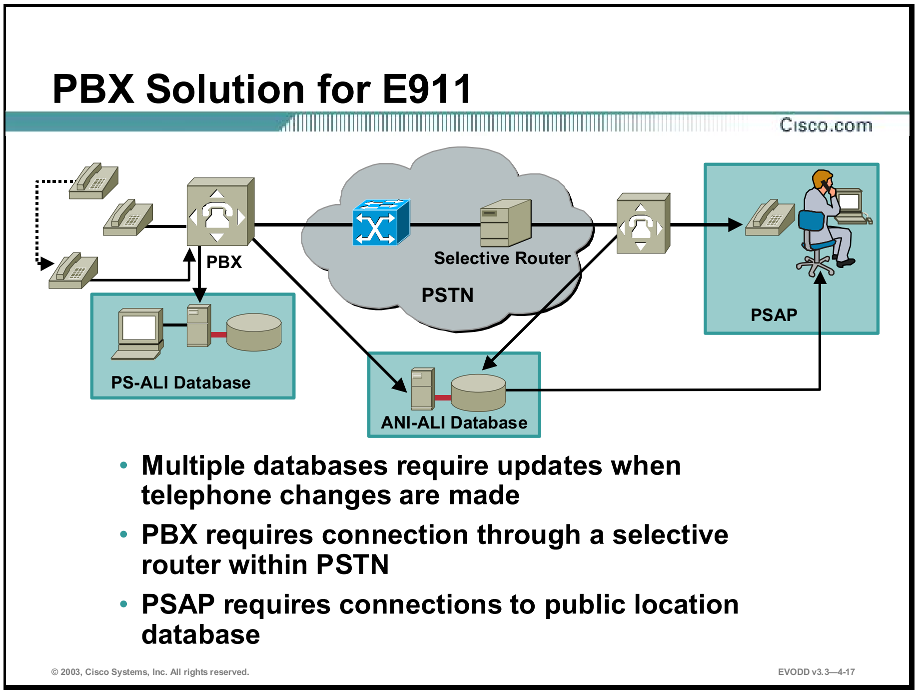
Handling 911 Issues

This topic describes issues that must be addressed regarding E911. The solutions to these issues are similar, regardless of whether you have a PBX or Cisco IP telephony environment.



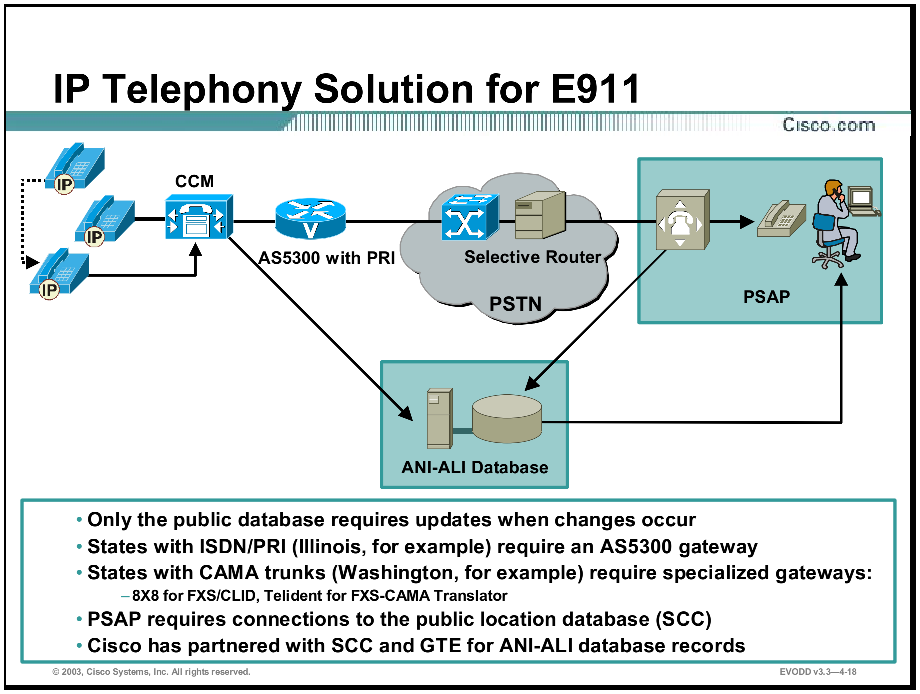
When you place a 911 call from home, the telephone number associated with the house identifies the location to the operator. The ALI links only to the specific telephone number, and the ANI.

With a PBX system, many telephone extensions or stations share telephone lines. When a 911 call comes through a PBX to the operator, the only address the operator sees is the address of the PBX, not the location of the caller. The caller could be on floors or in buildings far away from the PBX, which is where emergency personnel will arrive. The location information cannot link to a specific number because PBX contains shared telephone lines and the displayed operator number always changes. E911 cannot help in this situation.



Here is an example of how a PBX supports E911. When you move a telephone to another office, you need to update the PBX database. You also need to update the ALI database on the premises. Notify the service provider, maintaining the ALI database for the dispatch center, of the record update as well.

A PBX can now route the call to the central office, and then route the call to the closest dispatch center or PSAP. Because the PSAP pulls information from the ANI-ALI database, the PSAP knows the location of the caller.



The Cisco Emergency Responder approach is similar to the PBX approach. When a telephone is moved, the CCM database must be updated, which may happen automatically. The local database with the automatic locator and identifier information also needs to be updated. The updated information is then sent to the service provider, such as SCB or Verizon.

When a 911 call is placed, the call goes out via CCM, over ISDN in the scenario shown, to a selective router and on to the PSAP. As with the PBX solution, the PSAP pulls the location information from the ANI-ALI database and determines the exact location of the call.

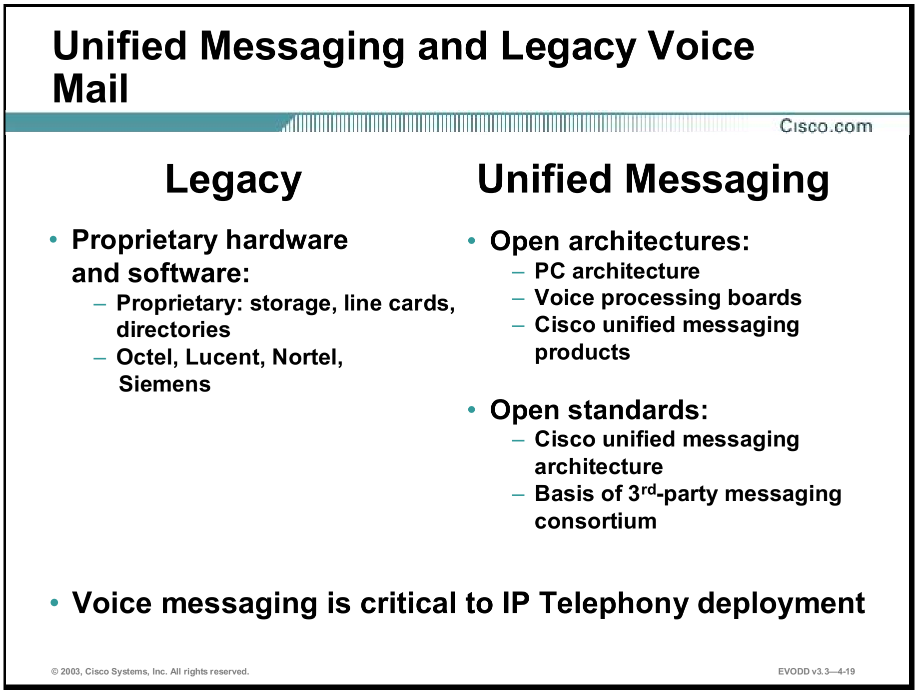
The solution is different for customers requiring Centralized Automated Message Accounting (CAMA) trunks. Instead of an AS5300 gateway, a specialized gateway is required. Two Cisco partners offer the complete solution: 8 X 8 offers the Foreign Exchange Station (FXS)/calling line identification (CLID) box and Telident offers the CAMA trunk interface.

Note The Cisco Emergency Responder is the Cisco IP telephony component that specifically addresses E911. More information can be found at: http://www.cisco.com/warp/public/cc/pd/unco/cer/. Contact your local 911 service provider for specific state regulations regarding 911 service.

Caution E911 issues are fraught with peril and possible legal ramifications. Therefore, unless you are very familiar with E911 issues and legalities, and are very knowledgeable about the Cisco Emergency Responder, you should consult experts in these areas.

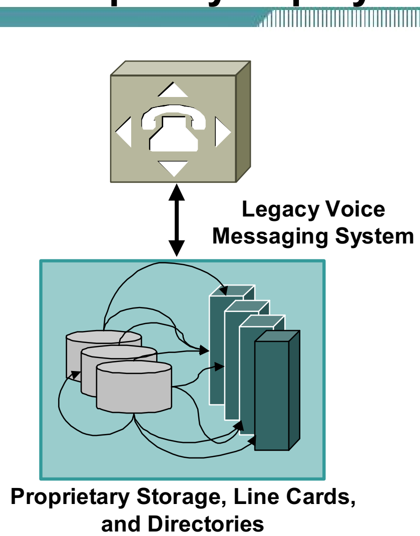
Voice Messaging

This topic describes unified messaging. Unified messaging is the technology that provides a single repository for multiple message types. Voice messaging is a key subset of unified messaging.



For any IP telephony deployment, voice messaging, commonly referred to as voice mail, is critical and is a mandatory component of the deployment.

Voice Messaging Critical to IP Telephony Deployment



- **Proprietary integration:**
  - **Proprietary software and hardware (for example, Octel, Lucent, Nortel)**

 EVODD v3.3—4-20

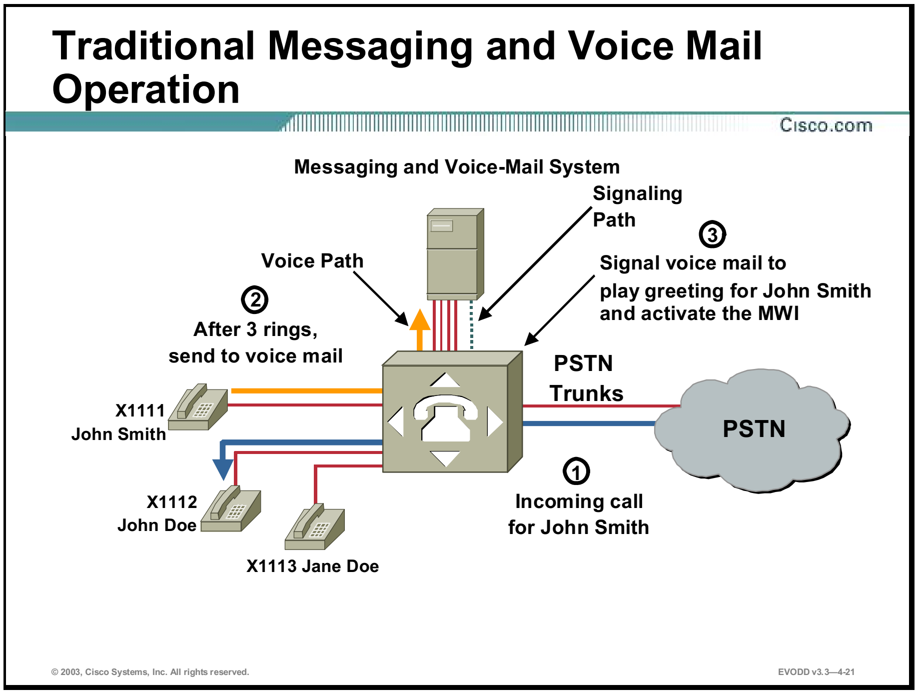
Current enterprises usually rely on voice-mail capabilities. Legacy voice-messaging systems are proprietary.

The Cisco open standards voice mail approach provides a way to simplify the architecture and operations for systems. The legacy approach makes integration of new features and operation difficult.

Note	Certain Cisco systems interoperate with some legacy systems, but it is important that you understand the details. For example, if a legacy voice-mail system interconnects to a legacy PBX via station-side line cards, perhaps the ports on the voice-mail system could be interconnected to FXS ports on a Cisco router or switch that supports Simplified Message Desk Interface (SMDI) protocol.

If a customer wants to transfer voice mail over to an IP telephony deployment at a later date, it is important to understand the coexistence issues between legacy voice mail and IP telephony.

Caution	You must always consider voice mail when designing and IP telephony deployment.



A typical voice-mail environment is shown here. Telephones are connected to a PBX or key system, and the voice-mail system is, most likely, attached to the PBX.

There are two types of communication paths between the PBX and the voice-mail system: an audio or voice path, and a signaling path. This scenario can be used to explain the two paths:

1. A call comes into the PBX system for John Smith.

2. After three rings, the call is forwarded to voice mail, and the PBX signals the voice-mail system to play the greeting that John has recorded.

3. Then the PBX activates the Message Waiting Indicator (MWI) on the telephone.

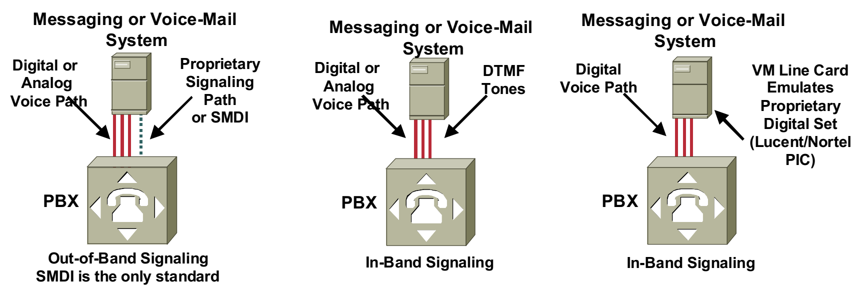
Information is transmitted on the signaling path. The signaling path can be a logical or physical connection. The information can include data; such as the number being called belongs to John Smith, the name of the person who is calling John Smith, and why the call was sent to voice mail.

The voice-mail system then determines which greeting to play for the call to John Smith. The caller can then leave a voice message in the voice mailbox.

Finally, the voice mailbox, using the signaling channel, instructs the PBX to light the MWI on the telephone.

Typical Connection Types for Legacy Voice Mail

Cisco.com



- **SMDI/signaling channel delivers:**
  - **Called party, calling party, message waiting information**
  - **Reason for sending to voice mail, forward all, forward busy, forward no answer, direct call**

 EVODD v3.3—4-22

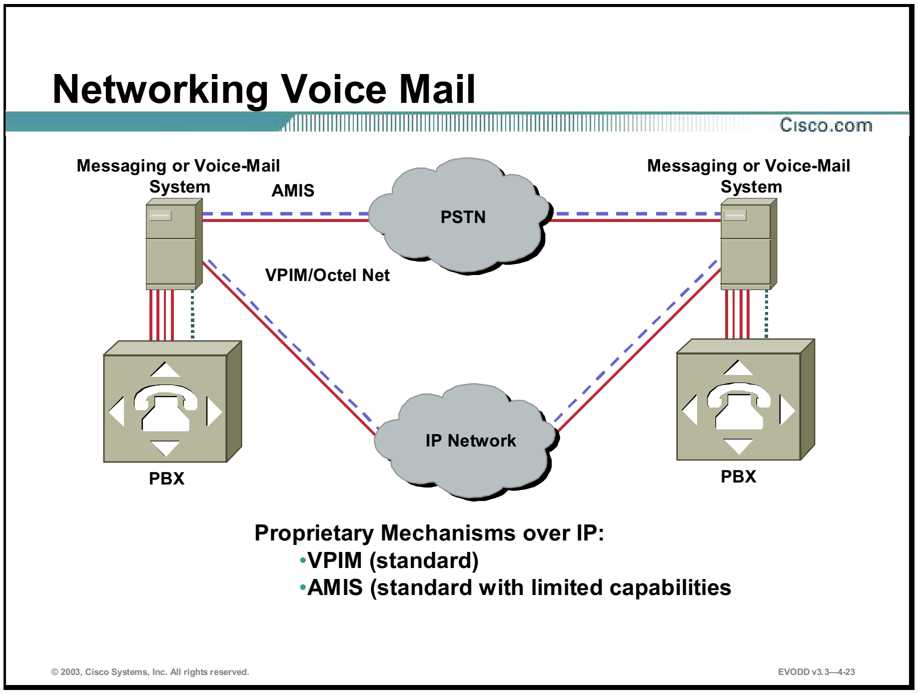
There are different approaches for the signaling connection between the voice-mail system and the PBX.

The left portion of the diagram shows the external signaling path described in the previous example. This signaling path can be proprietary, or SMDI, a Bellcore standard signaling method. SMDI is the only standards-based signaling method between PBX and voice mail. All other mechanisms are proprietary.

The center portion of the diagram shows a slight variation of SMDI, where the information about the called party, the reasons why, and other data, are sent via dual tone multifrequency (DTMF) tones in-band along with the audio flow. Often the information precedes the actual message.

The right portion of the diagram illustrates another very common approach. The voice-mail system contains line cards, such as those from Lucent, Nortel, or Siemens, that perform emulation of the target PBX.

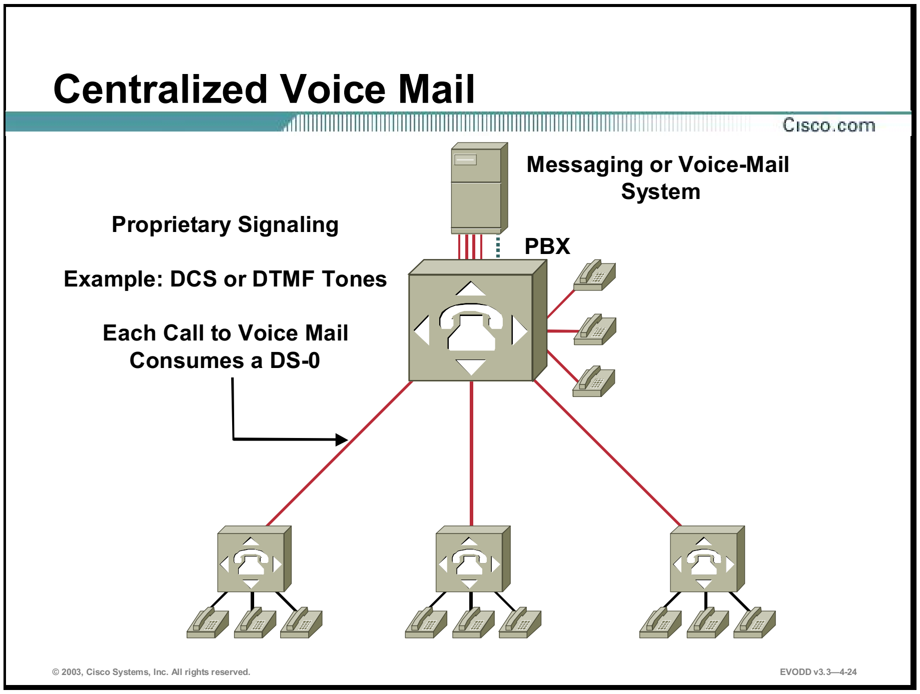
Whether the signaling is SMDI or proprietary, the key information that is sent to the voice-mail system includes, the called party, the calling party, message-waiting information, and the reason why the call was not answered.



The Audio Messaging Interchange Specification (AMIS) defines a pervasive industry method for enabling disparate voice messaging systems to internetwork, or exchange voice mail, via analog connections, typically through the PSTN. The receiving voice-mail system answers a telephone call and receives the analog voice data, including DTMF tones, for the voice mailbox.

Note Most major vendors currently support AMIS Analog (AMIS-A).

Voice Profile for Internet Mail (VPIM) is a protocol that is used for internetworking various voice-mail systems using the TCP/IP protocol suite. Simple Mail Transfer Protocol (SMTP) and the Multipurpose Internet Mail Extensions (MIME) VPIMv2 define that all message interchanges must utilize G.726. VPIMv3 supports G.711, G.726, G.723.1, and Microsoft-Global System Mobile (MS-GSM) for message encoding.

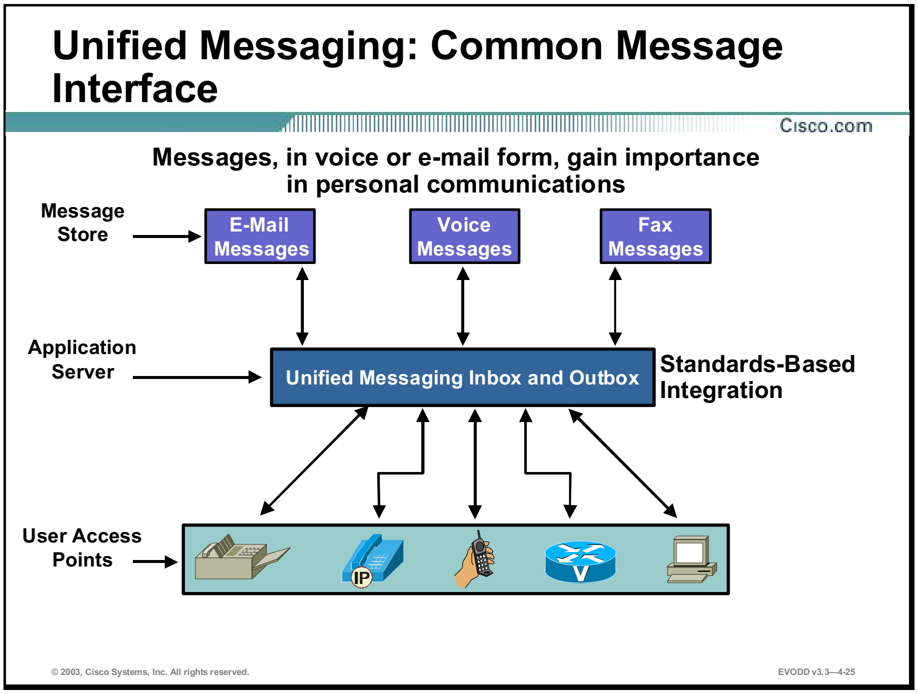


Another common topology is the classic hub and spoke topology, where a centralized voice-mail configuration is used. This configuration is typically seen where, for control or cost reasons, a company has decided not to place voice-mail systems at their remote locations

Centralized systems typically use a proprietary signaling approach. When a call comes in to a branch or remote site, and the telephone that is called is busy or does not answer, the call is forwarded to the central site and a message is left for the remote user. It is important to note that for each call that is not answered at a remote site, the call will consume a telephone circuit (analog trunk or DS-0) between the remote site and the central location.

Cisco Unified Messaging Solution

This topic describes how e-mails, faxes, voice mails, and pages all reside within a single repository.

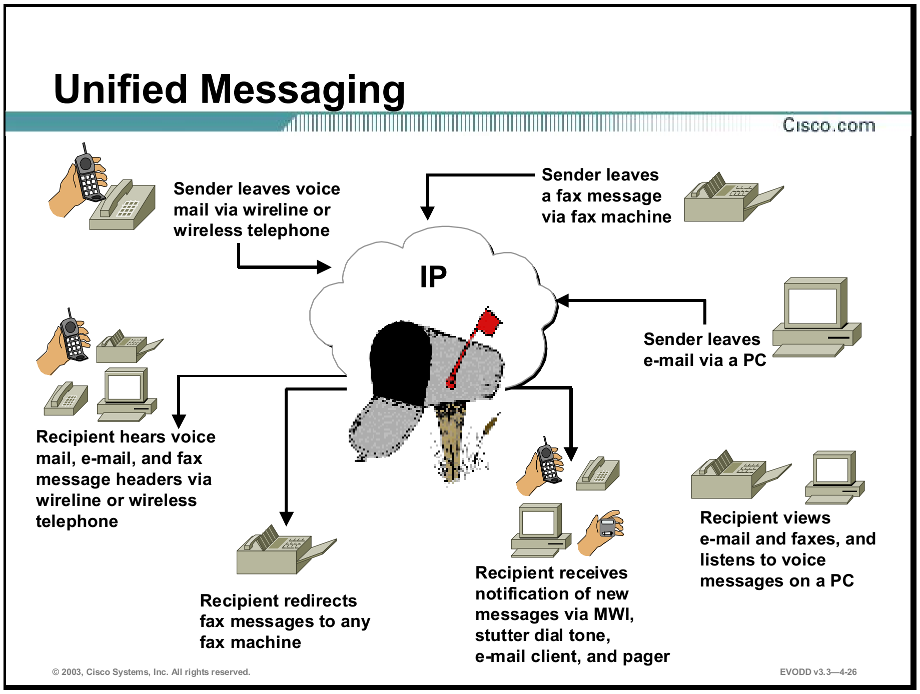


The Cisco unified messaging architecture can be understood as a three-tiered architecture, where e-mail, voice mail, and faxes are combined into a single system.

In unified messaging, the user has an inbox, where all types of messages are delivered. The user may access a message via an IP Phone, a handheld terminal, a full-featured PC, or a purpose-built interface.

The unified messaging inbox and outbox hold normalized messages, and perform normalization. This function, shown here as the middle layer, is performed by the Cisco unified messaging application server, called Cisco Unity, which runs on Windows 2000.

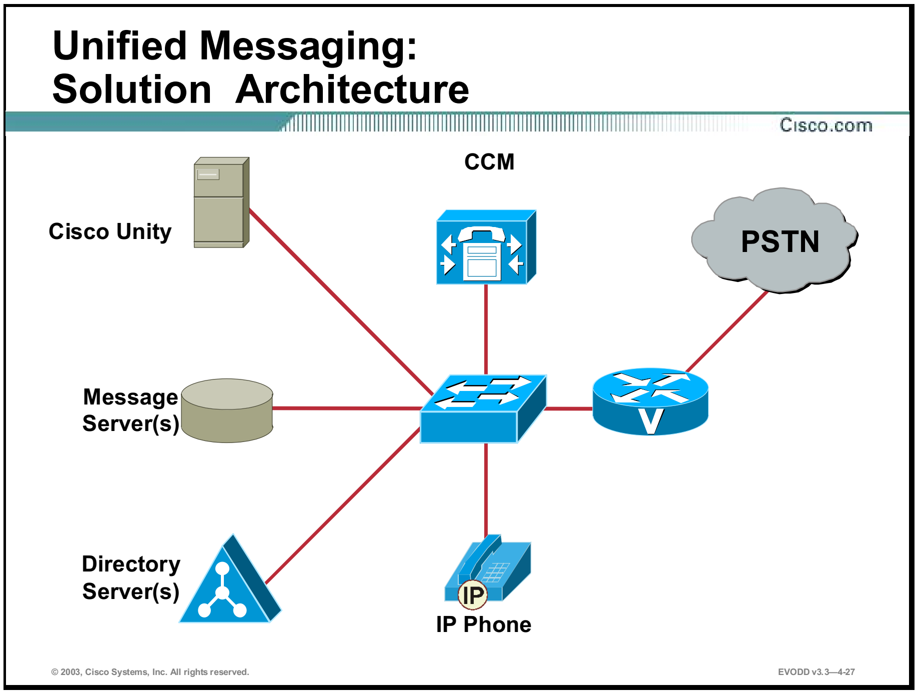
Unified Messaging



Unified messaging enhances productivity by changing the nature of the access device. For example, the user can listen to a voice message in the traditional manner, over the telephone, or the user can go to a web browser or e-mail client and listen to the message over a multimedia device, such as a PC, laptop, or PDA.

E-mail messages, once available only through a text-based or graphical interface, can now be heard via a telephone user interface with text-to-speech translation.

Faxes can be deposited into the message repository and redirected to any fax machine.



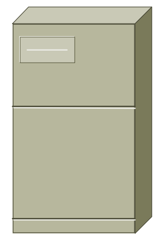
Each Cisco unified messaging solution is comprised of three main components: the Cisco Unity Server, the directory server, and the message store. These three components can run on the same server hardware, or they can be distributed on two or more network accessible hosts. The Cisco open architecture provides maximum flexibility in the choice of components. Particularly, it allows for the deployment of the Cisco unified messaging solution into other types of networks.

Cisco Unity Server

Cisco.com

Unity Windows NT/2000-based Communications Server:

- **Voice mail:**
  - **Unified messaging**
  - **Integrated faxing**
  - **Networking**
  - **VoIP**



 EVODD v3.3—4-28

The Cisco Unity server runs on Windows 2000 and uses Microsoft Exchange 2000 to store messages. Cisco Unity offers unified messaging, integrated faxing, using a supported third party faxing product, and networking, which allows the connection of Cisco Unity servers in branch offices across the country. Cisco Unity seamlessly integrates with CCM and supports many legacy PBX systems.

Cisco Unity includes these features:

- E-mail, voice, and fax messages are delivered to a subscriber e-mail inbox, giving users centralized communications control.
- E-mail, voice and fax messages can be accessed from a desktop PC or laptop computer using the Internet, or from any touchtone telephone.
- A text-to-speech module reads e-mail messages over the telephone in clear, spoken U.S., U.K. and Australian English, French, German, Dutch, or Spanish.
- E-mail, voice and fax messages can be sent to anyone who can receive e-mail over the Internet.
- A VCR-style, Graphical User Interface (GUI) interface lets you play, rewind, pause, or fast forward messages with a few clicks.
- Faxes can be stored for onscreen viewing or printing from any networked PC, and faxes can be forwarded to any fax machine from a touchtone telephone, via a fax server.

- A browser-based, personal administrator supports customized message notification options, allowing users to respond quickly to messages. This feature also allows IT staff to enable end users to manage their own accounts, saving time and decentralizing routine administration.

- Compound messaging capability provides the option to combine different media, such as attaching a Word file to a voice message.

- Global addressing speeds up the communications process.

- All message types can be downloaded for off-line response.

- Voice and fax messages can be saved, along with e-mail, in public or personal Microsoft Exchange and Outlook folders for a complete record of communications.

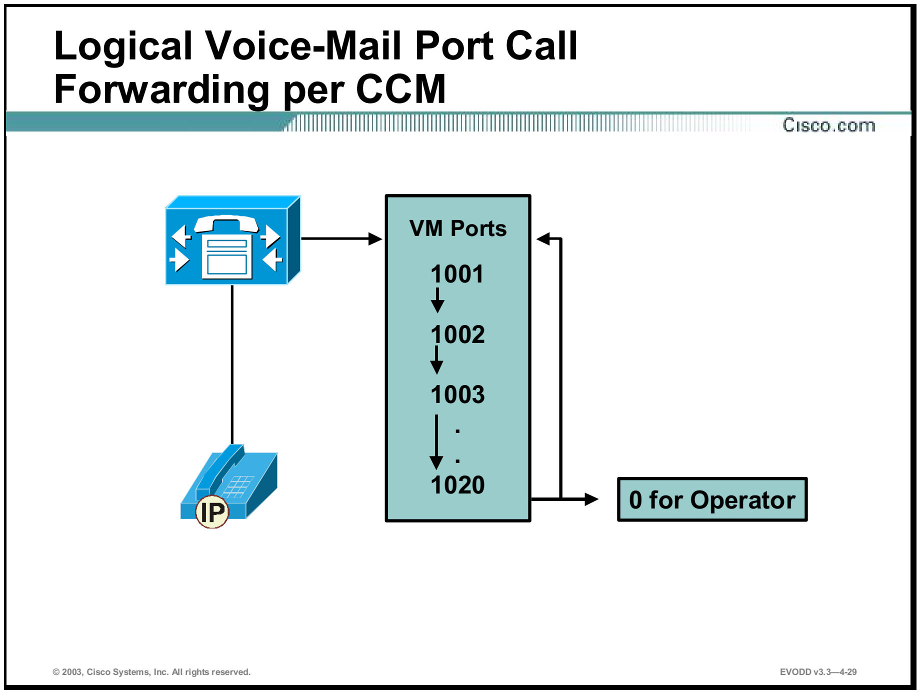
- Microsoft Exchange Inbox Assistant rules can be applied to voice mail and faxes.

- A browser-based system administration interface enables maintenance from any PC on the network, saving time, expense, and effort.

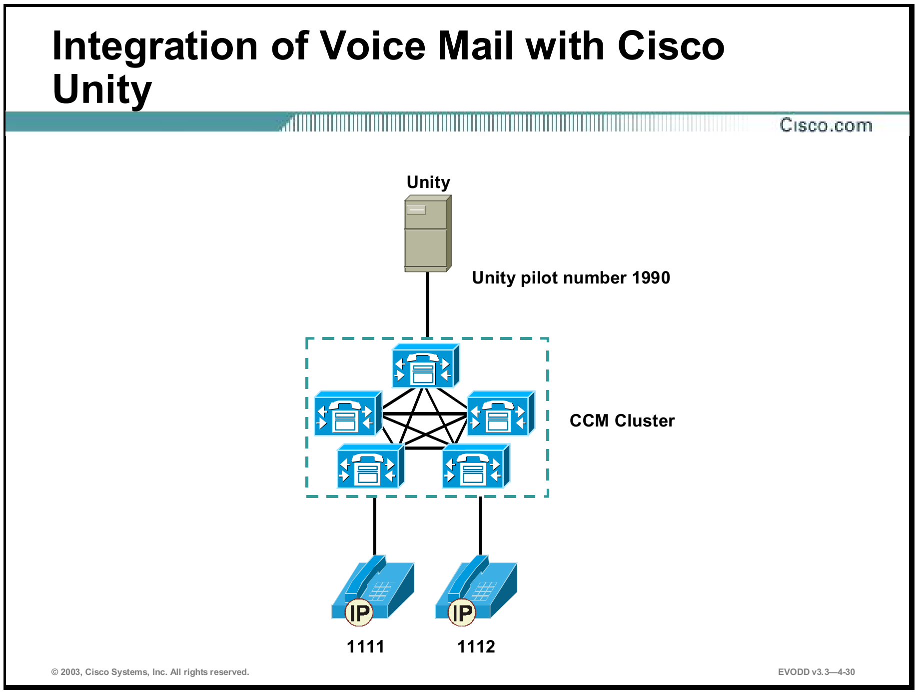
- Truly unified architecture allows IT staff to set one backup procedure, one message storage policy, and one security policy.

- Fault-tolerant system tools include robust security, file replication, event logging, and optional software Redundant Array of Independent Disks (RAID) levels 0-5.

- International product offering localized versions in multiple languages, including Dutch, four dialects of English (Australian, New Zealand, U.K., and U.S.), French, German, Norwegian, and Spanish. Depending on the language, the products feature everything from system prompts and subscriber conversations to the browser-based administration consoles and product documentation in the customer language of choice.



The example shows how CCM processes call forwarding busy (CFB) and forward no answer (FNA). In this example, extensions 1001 to 1020 handle this processing. Within the CCM, all users configured with voice mail have 1001 as their FNA and forward busy (FB) numbers. Extension 1001 has its FNA and FB numbers configured as 1002. Port 1002 has its FNA and FB configuration entered as 1003, until port 1020 is queried. Port 1020 has its FNA and FB configured as either 0 for the operator, or 1001 to cycle back through the directory numbers to see if any ports are now free By default, CCM cycles through up to 20 FNA and/or FB ports before sending a busy signal to the caller. You can configure up to a maximum of 200 ports.



Cisco Unity integrates with CCM using Skinny Client Control Protocol. This method has a major advantage over SMDI in that Skinny Client Control Protocol is IP-based; therefore no analog ports are used for the voice path. This method simplifies both server and network design.

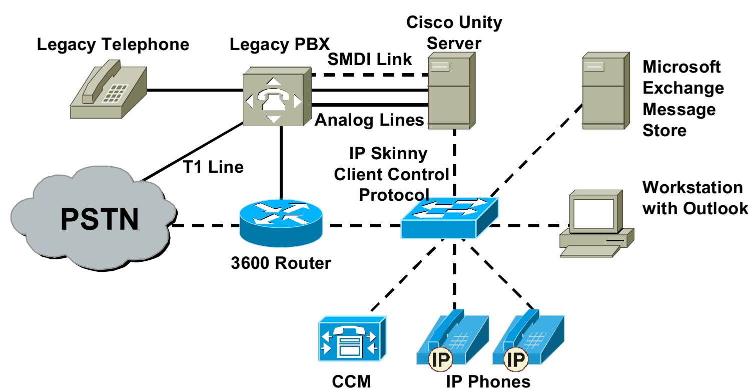
Voice mail integration through Cisco Unity provides these features and capabilities:

- Integration with CCM through Skinny Client Control Protocol (similar to an IP Phone)
- Use of the Voice Port Wizard in CCM to assign directory numbers and configure voice-mail ports
- Messages button on IP Phones to automatically dial the voice-mail pilot number (multiple pilot numbers are also possible)
- Multiple MWI on/off directory numbers within the same CCM cluster

Cisco Unity can also support multiple CCM clusters, thereby offering a centralized voice-mail service.

Cisco Unity Dual Switch Integration

Cisco.com



EVODD v3.3—4-31

Cisco Unity also provides the capability to integrate with both a traditional time-division multiplexing (TDM) PBX system and CCM simultaneously, as shown in the diagram. This dual integration capability can assist you in migrating from a traditional PBX system to CCM and, if desired, in moving to Cisco Unity prior to deploying CCM.

Cisco Unity supports dual integration with the following PBX systems:

- Lucent or Avaya Definity G3 (analog)
- Lucent or Avaya Definity G*x* (digital PBX Link)
- Nortel Meridian 1 (digital PBX Link)
- NEC NEAX2000 (MCI)
- NEC NEAX2400 (MCI)
- Centrex (SMDI)
  - AT&T 1AESS
  - AT&T 5ESS
  - Nortel Networks DMS100
- Ericsson MD-110 (serial)

- Mitel (analog ONS)
  - SX200
  - SX2000

Cisco Unity Platforms

Cisco.com

Platform	Supported Unified Messaging Users
Cisco ICS7750 - SPE310 Module (1 GB RAM)	500
Cisco MCS 7847	2200
Compaq ProLiant - DL380 G1/G2 - Single Processor	1599
Compaq ProLiant - DL380 G2 - Dual Processor	2200
Compaq ProLiant - ML370 G2 - Dual Processors	2200
Compaq ProLiant - DL580 G1 - Quad Processors	7500
Compaq ProLiant - DL570 G1 - Dual Processors	7500
Compaq ProLiant - ML570 G1 - Quad Processors	7500
Dell OptiPlex GX-150	499
Dell PowerEdge 1400 SC	1174
Dell PowerEdge 2500 - Dual Processor	2200
Dell PowerEdge 4400	6000
Dell PowerEdge 4600 - Dual Processor	6000
Dell PowerEdge 6400 - Quad Processor	7500
Dell PowerEdge 6450 - Quad Processor	7500
Dell PowerEdge 6600 - Quad Processor	7500
Dell PowerEdge 6650 - Quad Processor	7500

EVODD v3.3—4-32

Cisco Unity is supported on a specified set of platform, from vendors including Cisco, Compaq, Dell, and IBM. In this slide, a subset of the supported models from Cisco Compaq and Dell is shown, along with the number of unified messaging users supported by each platform. The above numbers indicate the maximum number of supported users (the message store may be external to the Unity server).

Note A more comprehensive listing of supported Cisco Unity platforms is located at: http://www.cisco.com/warp/public/cc/pd/unco/un/prodlit/ucutp_st.htm

Cisco Unity Platforms (Cont.)

Cisco.com

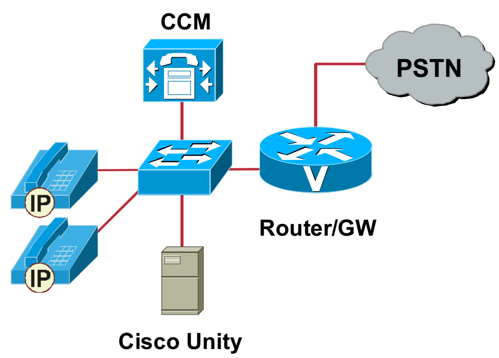
Platform	Supported Unified Messaging Users
IBM x232 - Single Processor	1599
IBM x232 - Dual Processor	2200
IBM x342 - Single Processor	1599
IBM x342 - Dual Processor	2200
IBM x250 - Quad Processor	7500

EVODD v3.3—4-33

In the figure above, a subset of supported models by IBM is shown, along with the number of unified messaging users supported by each platform.

Design Scenario: Single Site

Cisco.com

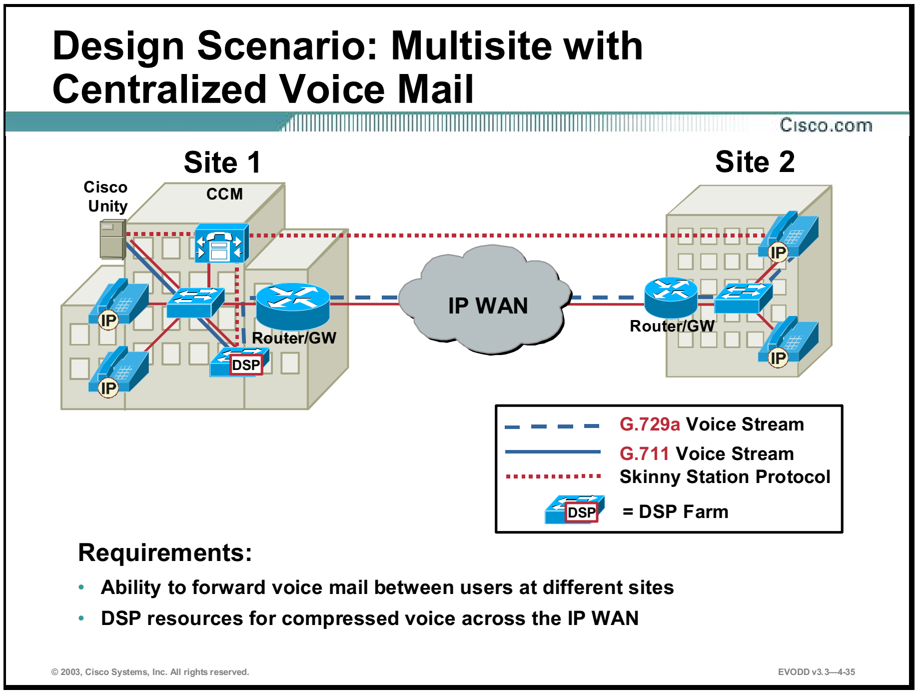


- **Requirements:**
  - Ability to exchange voice mail between users
  - G.711 voice only

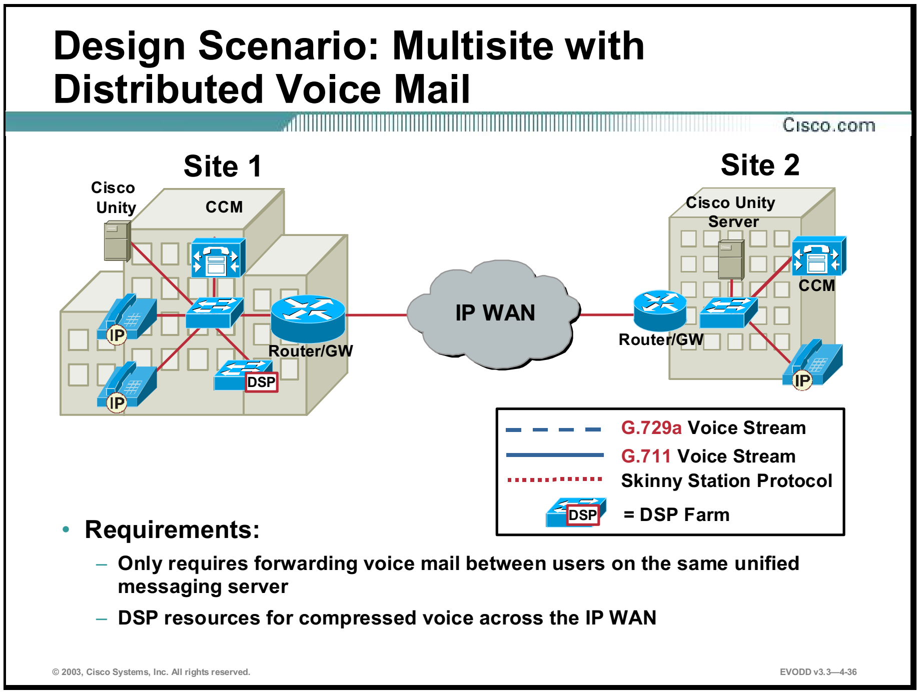
EVODD v3.3—4-34

The single-site deployment of a Cisco IP telephony voice-mail solution has all of the messaging components on a single server. This server runs the messaging components: Cisco Unity, directory server, and message store.

The solution shown here is appropriate for smaller enterprises (for example, a single site with less than 500 users). A single unified messaging server provides voice mail for all CCM users.



The configuration shown is appropriate for a small company, multisite environment. The total user base for voice mail is less than 500, and this scenario supports the ability to forward voice mail between users. This solution provides centralized call-processing and voice mail. All telephones can access centralized voice mail using G.729 via the digital signal processor (DSP) resources. An individual at Site 2 calling Site 1 dials the CCM, which enables the connection to both the DSP resources transcoding farm and the unified messaging server.



This example shows multiple sites with distributed voice mail. With the Cisco Unity networking feature, the servers can communicate with each other, allowing the servers to exchange messages with each other. The DSP farms are used for transcoding Cisco Unity server-to-server messages to G.729 if there is a need for low bandwidth across the IP WAN.

Reference For more UNITY case studies, refer to the following link: http://www.cisco.com/univercd/cc/td/doc/product/voice/c_unity/whitpapr/index.htm

Additionally, Cisco Learning Partners offer the following UNITY courses:

- Cisco UNITY System Administration: http://www.cisco.com/pcgi-bin/front.x/wwtraining/CELC/index.cgi?action=CourseDesc&COURSE_ID=1855

- Cisco UNITY System Engineer: http://www.cisco.com/pcgi-bin/front.x/wwtraining/CELC/index.cgi?action=CourseDesc&COURSE_ID=1856

- Cisco Unity Voice Mail for the ICS 7750 Installation: http://www.cisco.com/pcgi-bin/front.x/wwtraining/CELC/index.cgi?action=CourseDesc&COURSE_ID=1715

Summary

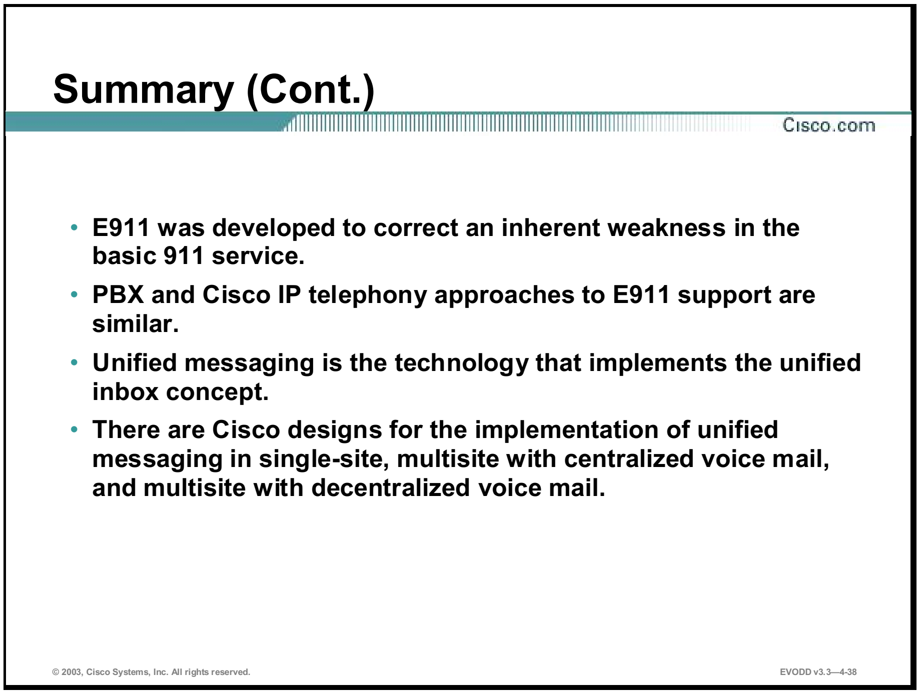
This topic summarizes the key points discussed in this lesson.

Summary

Cisco.com

- Dial plans govern the handling of both incoming and outgoing calls for a voice system.
- Dial plans are implemented using several tables that contain the logic for handling and routing both incoming and outgoing calls.
- Dial plan requirements can include support for abbreviated dialing, as well as redundant paths that are transparent to the calling party.
- There are several issues involved in migrating to a dial plan for VoIP.

EVODD v3.3—4-37



Next Steps

After completing this lesson, go to:

- Voice Over Data Characteristics module

References

For additional information, refer to these resources:

- Configuring Dial Plans, Dial Peers, and Digit Manipulation:

  http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122cgcr/fvvfax_c/vvfpeers.htm

- Planning for Emergency Responder:

  http://www.cisco.com/univercd/cc/td/doc/product/voice/respond/res_1_1/admin/e911plan.htm

- Cisco Unity:

  http://www.cisco.com/warp/public/cc/pd/unco/un/

Lesson Review

This practice exercise reviews what you have learned in this lesson.

Q1) A dial plan is analogous to which of these items?

A) router routing table

B) access-list

C) subnet

D) CoS

Q2) Which of these options best describes a partition?

A) partitions define which users can access which calling search spaces

B) partitions are CoS groups

C) partitions are groups of devices with similar accessibility

D) partitions are subnets

Q3) Which of the following do CCMs use to address problems with overlapping dial plans?

A) partitions

B) calling search spaces

C) CoS restrictions

D) translation patterns

Q4) Which three of the following are considered PBX to CCM migration issues? (Choose three.)

A) cost

B) ROI

C) scalability

D) security and regulatory impact

Q5) Which of these E911 features provide location information to a 911 operator?

A) ANI

B) ALI

C) PSAP

D) MLTS

Q6) Which Cisco product addresses E911 concerns?

A) NetRanger

B) Emergency Responder

C) PSAP-IP

D) SRST

Q7) What is the only standards-based signaling protocol used between a PBX and a voice-mail system?

A) H.323

B) QSIG

C) Q.931

D) SMDI

Q8) What are the three components of the Cisco unified messaging solution? (Choose three.)

A) Cisco Unity server

B) directory server

C) CCM server

D) message store

5

Voice Over Data Characteristics

Overview

There are many options for shuttling data and voice communications throughout an enterprise. At one extreme, you can run separate, parallel networks—one for data, one for voice—with each network optimized for the traffic type it was designed to support. This option is the typical enterprise network of today. At the other extreme are new ways of running packetized voice alongside data on a single network that maximizes network bandwidth use and yields large savings on WAN service costs. This module will teach you how Voice over Data technology minimizes the effects of loss, delay, jitter, and echo. You will learn how codecs work and how they differ in design and efficiency. You will also learn how fax over data operates to handle real-time fax.

Upon completing this module, you will be able to:

- Describe how voice is compressed
- Describe the impact of loss, delay, and jitter
- Describe how echo is kept from degrading voice quality
- List special requirements for Fax over Data

Outline

The module contains these lessons:

- Delay Characteristics
- Compression Technologies and Packet Compensation

Delay Characteristics

Overview

Achieving high quality voice requires addressing a large number of variable conditions when planning a network design. This lesson will teach you about several delay characteristics of Voice over Data. You will learn how to calculate a delay budget.

Importance

Since delay results in poor voice quality, it is important to understand the characteristics of data and the causes of delay. Also, determining the amount of acceptable delay is critical to user acceptance.

Objectives

Upon completing this lesson, you will be able to:

- Identify the characteristics of Voice over Data
- Describe the causes of fixed delay
- Describe the causes of variable delay
- Calculate a delay budget
- Describe the method used to assign numerical values to voice quality

Learner Skills and Knowledge

To fully benefit from this lesson, you must have these prerequisite skills and knowledge:

- A fundamental understanding of Cisco AVVID components
- A fundamental understanding of voice services offered by a PBX

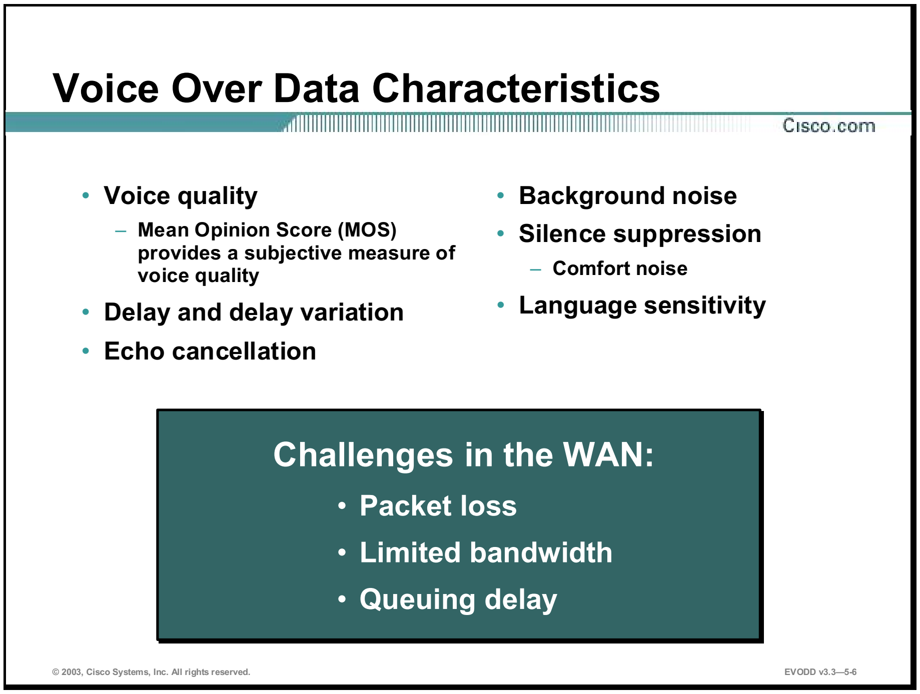
Outline

This lesson includes these topics:

- Overview
- Delay Components
- Fixed Delay
- Variable Delay
- Delay Budget
- Impact of Delay on Quality
- Summary
- Lesson Review

Delay Components

This topic discusses components that contribute to the overall delay budget of a network (such as the total end-to-end delay).



There is several measurable Voice over Data characteristics. One of the primary measurable characteristics is voice quality. Some of the design considerations surrounding voice quality include delay and delay variations, packet loss, background noise, and language sensitivity.

The mean opinion score (MOS) is a subjective measure in which voice quality is determined by the trained ear. This measure is similar to the judging seen in figure skating; if someone falls down the quality is known to be less whether or not the technical reasons are known. A more current measure of speech quality is Perceptual Speech Quality Measurement (PSQM), as defined in ITU standard P.861. PSQM allows test equipment to perform quality measurements.

Delay and delay variation is a critical design area. Delay variation comes into play when it is time to reassemble the packets at the playout end. If the packets arrive at different rates, the playout software waits until all the pieces arrive so that all that the packets can play out in their proper order. This waiting, called dejitter buffering, causes additional delay.

Packet loss is a similar problem, but a delayed packet is just as bad as a lost packet.

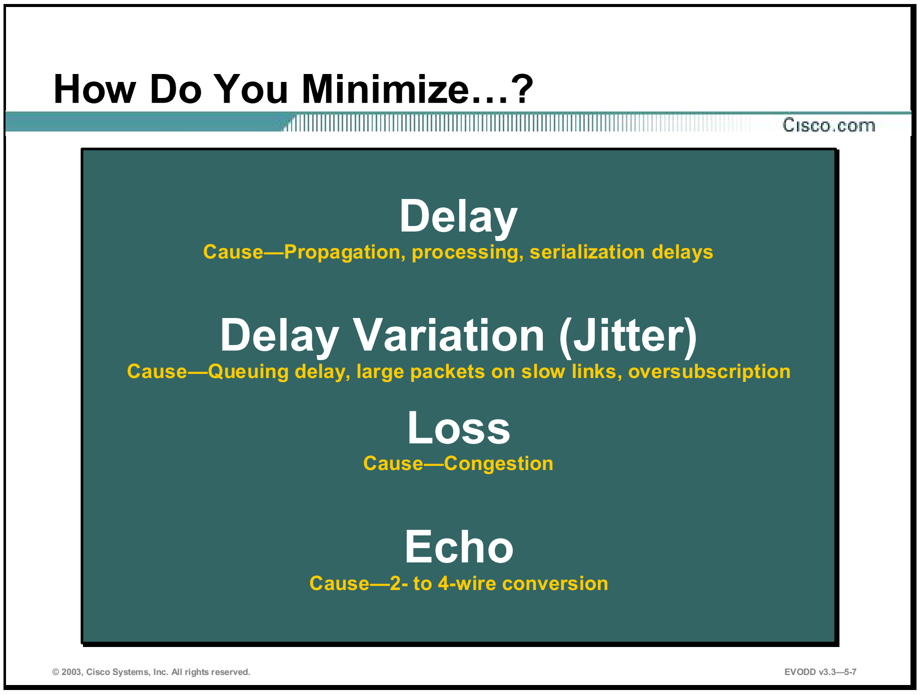
Echo is a problem in all types of voice networks. Echo is caused by signal reflections of the speaker's voice from the far-end telephone equipment back into the speaker's ear. The solution is primarily a matter of tuning, called "echo cancellation."

Background noise is another problem. When neither party is speaking, the silence can be disconcerting. People are accustomed to hearing some background noise, which is different in an airport than in a doctor's office.

One of the techniques used to save bandwidth is to first recognize silence. An indication of the silence type and duration can be sent and replayed at the receiving end. The problem that must be solved is approximating background noise so that the listener hears something that sounds realistic. Suppressing silence in this manner can save considerable bandwidth since most conversations are about 65 percent silence (AT&T statistics cite 62 percent).

Finally, language sensitivity can cause difficulties in quality. Codecs that rely on codebooks encounter sounds from languages that lead to inexact pattern matches.

These issues, coupled with issues from the WAN such as packet loss, queuing delays, and bandwidth limitations, mean that achieving high quality voice requires addressing a large number of variable conditions.

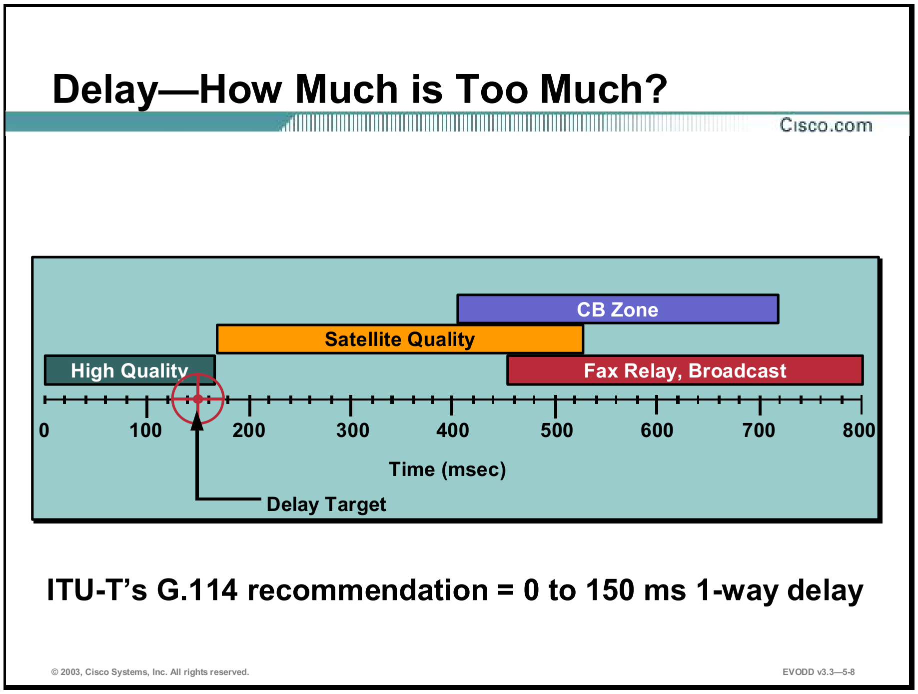


Delay can increase the user's awareness of echo and can cause talker overlap. Echo becomes a significant problem when the round-trip delay becomes greater than 50 ms. Voice over Data systems must address the need for echo control and implement some means of echo cancellation because echo is perceived as a significant quality problem.

Talker overlap (or the problem of one talker overlapping the other talker's speech) becomes significant if the one-way delay becomes greater than 250 ms. Calls go to walkie-talkie type mode if delays are bad. The end-to-end delay budget is the major constraint and driving requirement for reducing delay through a packet network. Propagation, processing, and/or serialization introduce delay.

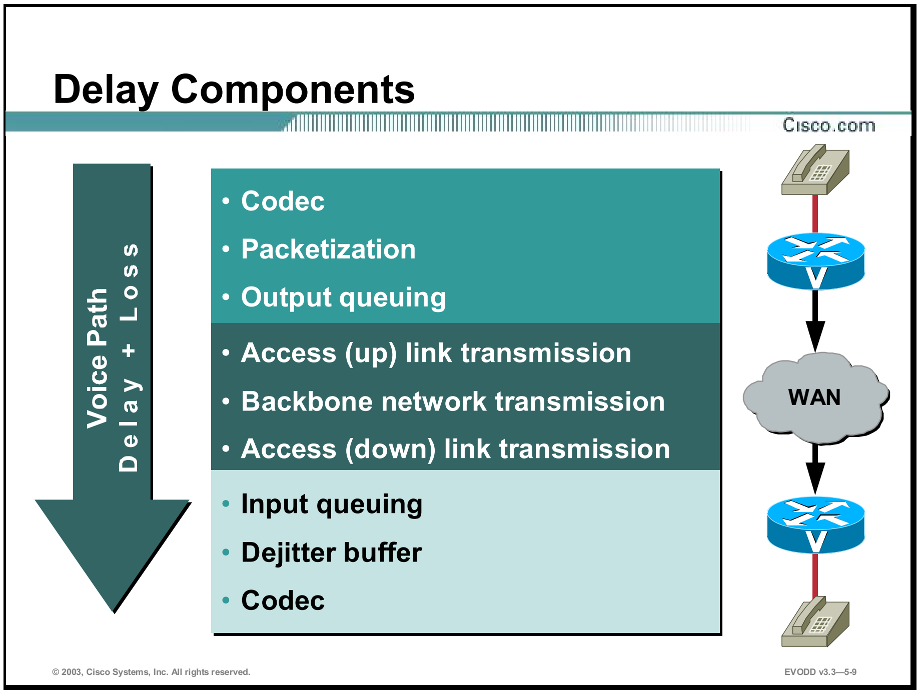
Lost packets, usually caused by congestion in the WAN, result in speech dropouts, or a stutter effect. The stutter effect is caused when the playout side tries to mitigate the problem by repeating previous packets.

Delay variation, or jitter, is caused by queuing delay, line over-subscription, or a condition known as large packet freeze out. The playout side attempts to accommodate the late packets by buffering them, thus smoothing out the differences in interarrival times and playing the voice samples out in a steady fashion. This strategy induces delay because of its inherent buffering. After a certain amount of time, a late packet is deemed lost and the process must continue.



This illustration attempts to quantify levels of quality against delay measured in milliseconds. The International Telecommunication Union (ITU) specifies a 150-ms end-to-end delay as the threshold for high-quality voice. Voice crossing satellite links encounter additional delay, and experience a drop in quality that users experience because of the distances covered. Cellular quality is in a similar range, with users who are willing to compromise some quality for convenience. Citizens band, fax relay and broadcast are at the higher end of the spectrum.

Voice over Data for the enterprise targets the high-quality end of the spectrum. The 150-ms or less delay is the threshold that governs design considerations.



The illustration shown here characterizes the delays encountered across the voice path. These delays start with the codec itself, which collects a sample to digitize and compress. This process takes time, and then the samples are packetized. If a voice information frame has two 10-ms samples, the first sample has already accumulated a 10-ms delay when the second one is being processed. These packets arrive in an output queue that adds additional delay.

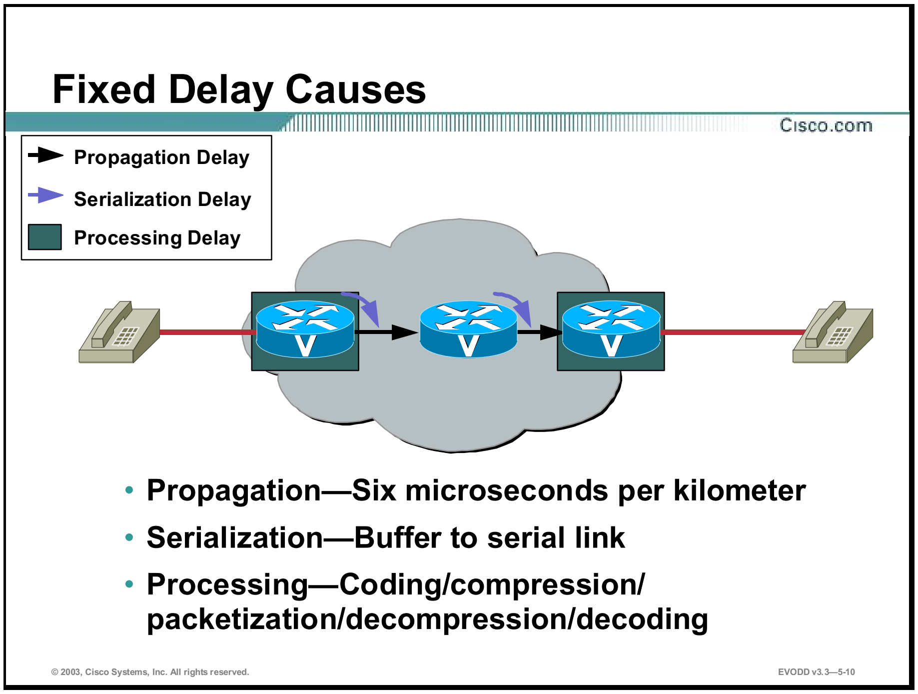
The transmission delays include the time it takes to serialize the bits in the packets onto the transmission medium. This time can vary widely. The propagation delay of the actual transmission must be accounted for as well as the output queues at the other end.

Finally, on the receiving end, packets are queued, run through a dejitter buffer, and then played out. A dejitter buffer can add a significant delay to the process.

The calculation of these delay components is part of creating an end-to-end delay budget to discover where problems may be encountered.

Fixed Delay

There are two types of delay: fixed and variable. This topic discusses fixed delay.



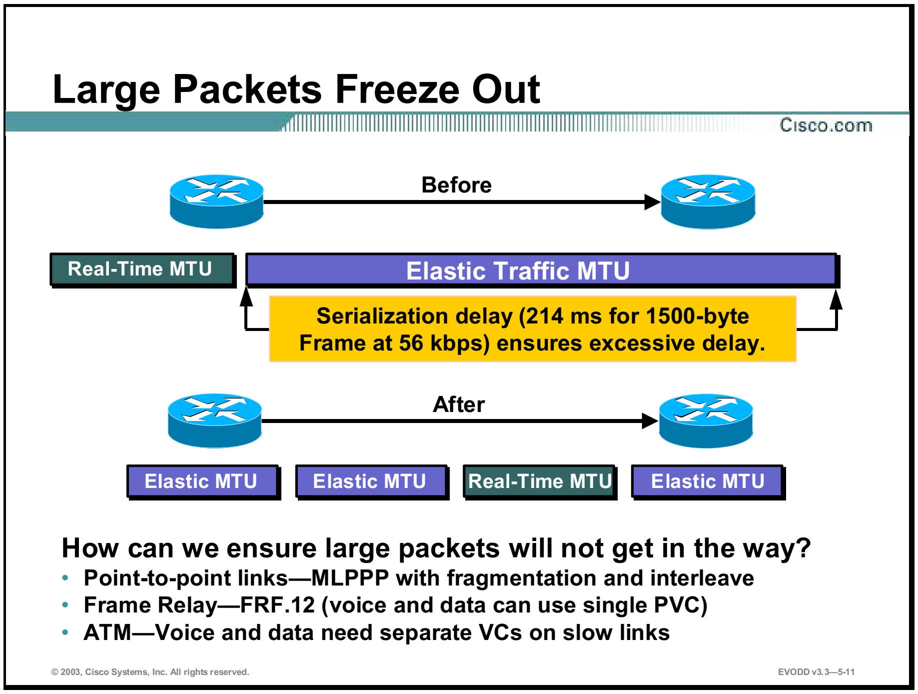
Fixed delay includes propagation, serialization, and processing delays. Fixed causes are predictable and can be determined when the connection is established to calculate a delay budget.

Processing delay is caused by the need to collect a frame of voice samples to be processed by the voice coder. The processing delay is related to the type of voice coder used and varies from a single sample time (0.125 microseconds) to many milliseconds. A representative list of standard voice coders and their frame times follows:

- G.726—ADPCM (16, 24, 32, 40 kbps) 0.125 microseconds
- G.728LD—CELP (16 kbps) 2.5 milliseconds
- G.729CS—ACELP (8 kbps) 10 milliseconds
- G.723.1—Multi Rate Coder (5.3, 6.3 kbps) 30 milliseconds

The actual process of encoding and collecting the encoded samples into a packet for transmission over the packet network causes additional processing delay. The encoding delay is a function of both the processor execution time and the type of algorithm used. Often, multiple voice coder frames will be collected in a single packet to reduce the packet network overhead. For example, three frames of G.729 code words, equaling 30 ms of speech, may be collected and packed into a single packet.

Serialization delay is a function of packet size and line speed. Propagation delay is a function of distance and can be calculated as 6 microseconds per kilometer.



Serialization delay becomes a major problem on low-speed links when small voice frames get caught behind large data frames. This delay leads to a condition called large packet freeze out and ensures that delay problems occur for any delay sensitive traffic, but especially voice traffic.

In the example shown here, a 1500-byte frame on a 56-kbps circuit takes 214 ms to play out. A frame carrying a voice packet queued behind this large frame is delayed so long that it will probably be considered a lost packet. A mechanism is required to break large frames into smaller fragments to allow the smaller voice frames to get out in a timely fashion. The actual mechanisms that fragment frames are dependent upon the Layer 2 technology. These quality of service (QoS) mechanisms are mandatory for assuring that the transport infrastructure is capable of delivering voice quality.

Fixed Frame Serialization Delay Matrix

Cisco.com

Frame Size

Link Speed	1 Byte	64 Bytes	128 Bytes	256 Bytes	512 Bytes	1024 Bytes	1500 Bytes
56 kbps	143 µs	9 ms	18 ms	36 ms	72 ms	144 ms	214 ms
64 kbps	125 µs	8 ms	16 ms	32 ms	64 ms	128 ms	187 ms
128 kbps	62.5 µs	4 ms	8 ms	16 ms	32 ms	64 ms	93 ms
256 kbps	31 µs	2 ms	4 ms	8 ms	16 ms	32 ms	46 ms
512 kbps	15.5 µs	1 ms	2 ms	4 ms	8 ms	16 ms	23 ms
768 kbps	10 µs	640 µs	1.28 ms	2.56 ms	5.12 ms	10.24 ms	15 ms
1536 kbps	5 µs	320 µs	640 µs	1.28 ms	2.56 ms	5.12 ms	7.5 ms

µs = microsecond ms = millisecond

EVODD v3.3—5-12

This matrix shows the relationship between link speed and frame size. When calculating a delay budget, refer to the link speed and frame size numbers.

Serialization delay is a particular problem on low-speed links. A 1024-byte frame requires 144 ms to be sent on a 56-kbps circuit. Serialization delay alone consumes most of the 150 ms target delay budget.

Understanding the impact of these numbers to network design is critical, particularly in capacity planning. Many design mistakes appearing in network implementation trace back to this part of the delay calculation.

Serialization delay is defined as the time it takes to put the bits on the wire. As illustrated by the following formula, decreasing access speed or increasing packet size negatively affects serialization delay:

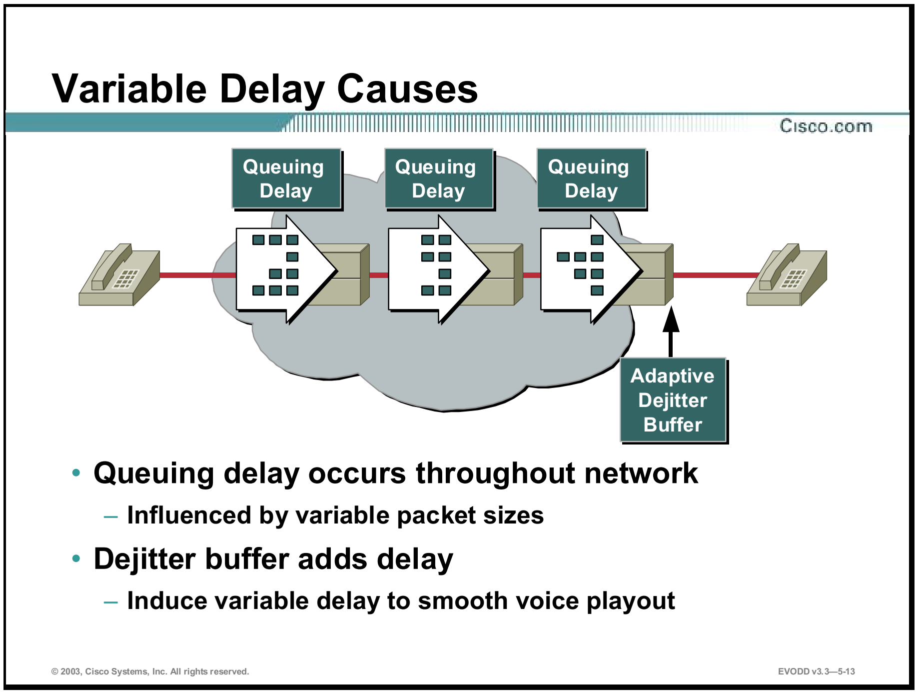
- Serialization delay = (# bits in packet/access rate)

As an example, an 8-byte packet (that is, 640 bits) that exits a serial interface running at 64,000 bps would have a serialization delay of 10ms (that is, 640/64000).

Traffic shaping components such as Committed Information Rate (CIR) or Peak Information Rate (PIR) never affect serialization. Serialization is always computed using the physical port speed. For example, it takes 125 microseconds to place one byte on a 64-Kb circuit (8 bits / 64000 bits/sec = 0.000125). The same byte placed on an OC-3 circuit will take 0.05 microseconds (8 bits/155000000 bits/sec = 0.00000005).

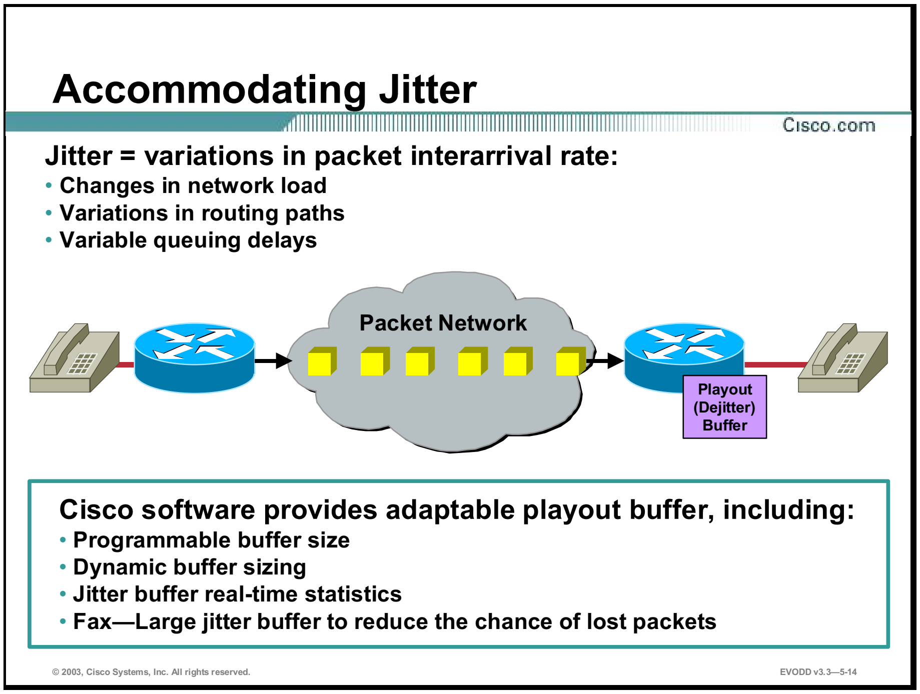
Variable Delay

This topic describes approaches for minimizing variable delay. Variable delay components include network queuing delay and the delay induced by a dejitter buffer at the playout end of a call.



Queuing delay is a function of the link capacity in the network and the processing that occurs as the packets transit the network. The jitter buffer adds delay to smooth out the packet delay variation that each packet is subjected to as it transits the network. This delay can be a significant part of the overall delay because packet delay variations can be as high as 70 ms to 100 ms in some Frame-Relay networks and IP networks.

Queuing delay is influenced by packet sizes and congestion in the WAN. Solutions involve QoS mechanisms, such as class-based weighted fair queuing (CBWFQ), IP RTP Priority, and low latency queuing (LLQ).



The delay problem is compounded by the need to remove jitter. Jitter is a variable interpacket timing caused by the network that a packet crosses. Delay problems are compounded by the need to remove jitter. To remove jitter you must collect packets and hold them long enough to allow the slowest packets to arrive in time to be played in the correct sequence and timing. This collection and holding process causes additional delay.

The two conflicting goals—minimizing delay and removing jitter—have propagated various schemes to adapt the jitter buffer size to match the time varying requirements of network jitter removal. This adaptation has the explicit goal of minimizing the size and delay of the jitter buffer, while at the same time maintaining the even spacing of packet delivery.

There are two approaches to adapting the jitter buffer size. The approach selected will depend on the network type.

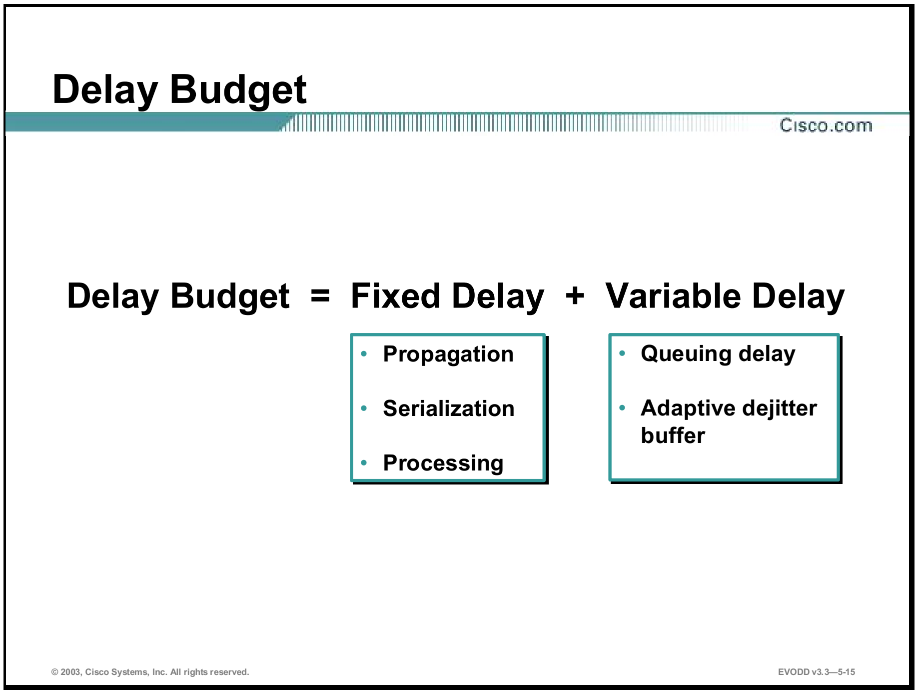
The first approach is to measure the variation of packet level in the jitter buffer over a period of time and to incrementally adapt the buffer size to match the calculated jitter. This approach works best with networks that provide a consistent jitter performance over time (such as ATM networks).

The second approach is to count the number of packets that arrive late and create a ratio of these packets to the number of packets successfully processed. This ratio is then used to adjust the jitter buffer to target a predetermined allowable late packet ratio. This approach works best for networks with highly variable packet interarrival intervals (such as IP networks). Cisco hardware uses this strategy.

In addition to the techniques described, the network must be configured and managed to provide minimal delay and jitter, enabling a consistent QoS.

Delay Budget

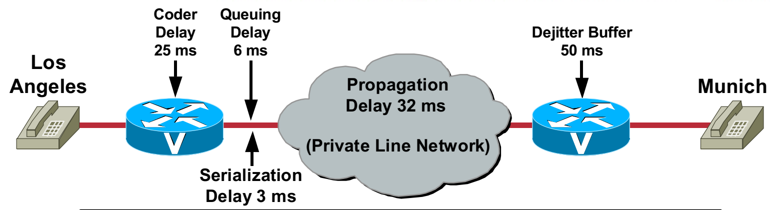
This topic demonstrates delay budget calculation.



Fixed delay includes propagation, serialization, and processing delays. Currently, configuring voice activity detection (VAD) adds an additional fixed delay of 5 ms for processing on the outbound voice channel. This delay needs to be considered in the delay budget if VAD is used. Variable delay includes queuing delay and the dejitter buffer delay.

Calculate Delay Budget

Cisco.com



	Fixed Delay	Variable Delay
Coder Delay G.729 (5 msec look ahead)	5 msec	
Coder Delay G.729 (10 msec per frame)	20 msec	
Packetization Delay—Included in Coder Delay		
Queuing Delay 64-kbps Trunk		6 msec
Serialization Delay 64-kbps Trunk	3 msec	
Propagation Delay (Private Lines)	32 msec	
Network Delay (Public Frame Relay Svc)		
Dejitter Buffer		50 msec
Total	60 msec	56 msec

 EVODD v3.3—5-16

The illustration shows a sample delay budget calculation (assuming no VAD, which would add an additional 5 ms).

A delay budget can be calculated with a general understanding of the fixed and variable delay components. The delay budget is the amount of delay permissible in the planned network while adhering to voice quality objectives.

Review the example shown and assume the delay budget is 200 ms. Note that this example is from a private Frame Relay network running over leased lines. For a public service such as public Frame Relay, you should contact the service provider for their guaranteed delay figures and use those figures in the delay budget.

First, fill a Frame Relay frame with two 10-byte voice frames. The coder delay for G.729 voice compression is an initial 5 ms for a look-ahead, plus 20 ms for the two 10-byte frames.

Packetization delay is the rate at which a packet is filled. This delay is typically governed by the speed at which voice samples are played out. Standard voice is transmitted at 64 kbps or one byte every 125 microseconds. Packetization delay is included in the coder delay.

Queuing delay is variable; with a 64-kbps trunk, the delay will equal 3 ms per 20-byte packet already in queue. Two voice packets in queue were used, for a total queuing delay of 6 ms. This assumption is variable and cumulative based on the number of devices in the network.

Serialization delay is the time it takes to play out a 20-byte packet onto the 64-kbps trunk. The propagation delay on a 64-kbps private line from Los Angles to Munich is 32 ms. Assuming the adaptable dejitter buffer was set to a nominal playout value of 50 ms, that much delay is induced on the playout.

Consequently, there are 60 ms of fixed delay and 56 ms of variable delay, equaling 116 ms plus the network queuing delay from the Frame Relay network. This example is likely to deliver quality voice calls over this Voice over Frame Relay setup, because the total delay budget of 116ms is below the International Telecommunication Union-Telecommunication Standardization Sector (ITU-T) G.114 recommendation for a maximum one-way delay of 150 ms.

Impact of Delay on Quality

This topic describes how numerical values are assigned to measure voice quality.

Voice Quality Guidelines

Cisco.com

MOS Values

Score	Quality	Description of Impairment
5	Excellent	Imperceptible
4	Good	Just perceptible, not annoying
3	Fair	Perceptible and slightly annoying
2	Poor	Annoying but not objectionable
1	Bad	Very annoying and objectionable

Voice Quality Measurements

Codec Type	Mean Opinion Score	Delay (ms)
G.711	4.1	0.75
G.726	3.85	1
G.728	3.61	3.5
G.729	3.92	10
G.729a	3.9	10

EVODD v3.3—5-17

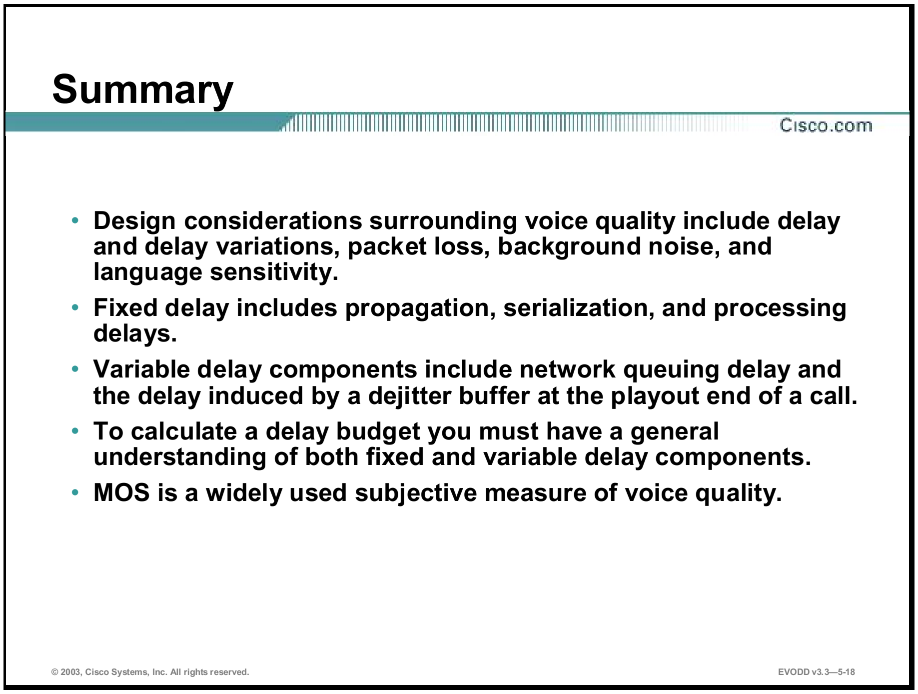
The MOS is a widely used subjective measure of voice quality, as defined by ITU-T's recommendation P.800. Voice samples are recorded according to ITU-T P.50. These voice samples are then played back to a group of men and women, who score the quality of the playback on a scale of 1 through 5.

Scores of 4 to 5 are deemed toll quality; 3 to 4 are communication quality; and scores less than 3 are considered synthetic quality.

In the illustration shown here, the top table describes the differences in MOS values. The bottom table shows a variety of codec types and their corresponding MOS values. Notice that the MOS for the G.729 codec is 3.92. However, the MOS score for G.729 changes to 2.68 if the G.729 call is placed through three tandem encodings, which introduce delay.

Summary

This topic summarizes the key points discussed in this lesson.



Next Steps

After completing this lesson, go to:

- Compression Technologies and Packet Compensation lesson

References

For additional information, refer to these resources:

- Understanding Delay in Packet Voice Networks:

  http://www.cisco.com/warp/public/788/voip/delay-details.html

- Synopsis of Basic VoIP Concepts:

  http://www.cisco.com/univercd/cc/td/doc/product/lan/cat4224/sw_confg/voipintr.htm

Lesson Review

This practice exercise reviews what you have learned in this lesson.

Q1) Echo becomes a significant problem when the round-trip delay becomes greater than what amount of time?

A) 10 ms

B) 50 ms

C) 125 ms

D) 150 ms

Q2) Which three of the following components are fixed delay components? (Choose three.)

A) propagation delay

B) queuing delay

C) serialization delay

D) processing delay

E) packetization delay

Q3) Which two of the following QoS mechanisms can be used to influence queuing delay? (Choose two.)

A) CBWFQ

B) WRED

C) IP RTP Priority

D) cRTP

Q4) What is the ITU-T's G.114 recommendation for maximum one-way delay?

A) 10 ms

B) 20 ms

C) 125 ms

D) 150 ms

Q5) What effect does tandem encodings have on a MOS value?

A) It increases the MOS value.

B) It decreases the MOS value.

C) It has no effect on the MOS value.

D) Tandem switching uses a different MOS scale (such as, 1-10).

Compression Technologies and Packet Compensation

Overview

Voice coding, decoding, compression, and decompression are handled by the coder/decoder (codec). Depending on the type of packet network used, lost packets can also be a problem in today's Voice over Data network. This lesson teaches you about waveform encoding and how to implement voice codecs and hybrid coders. This lesson also examines quality degradation from lost packets.

Importance

Compression can affect the quality of voice and the required bandwidth. In order to determine the proper network resources, you must understand the compression option available. In addition, when sending voice over a data network, it is essential for packets to be received in a timely manner and legacy features such as fax capability must be preserved.

Objectives

Upon completing this lesson, you will be able to:

- List voice compression technologies
- Describe the characteristics of waveform encoding
- Describe hybrid coders functionality
- Describe speech interpolation
- Compare the relative quality of codecs
- List potential reasons for lost packets
- Explain the cause of echo

- Describe how Cisco provides support for fax transmission across a Voice over Data network

Learner Skills and Knowledge

To fully benefit from this lesson, you must have these prerequisite skills and knowledge:

- A fundamental understanding of codecs and their relative bandwidth requirements
- A fundamental understanding of telephony electronics (such as, 2-wire vs. 4-wire circuits)

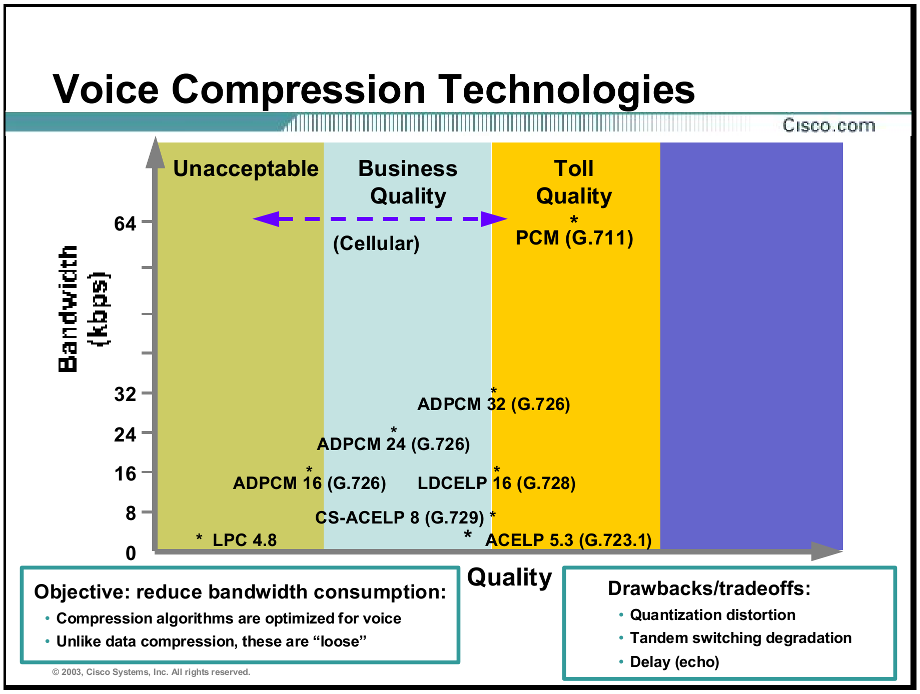
Outline

This lesson includes these topics:

- Overview
- Speech Coding
- Waveform Encoding
- Hybrid Coders
- Voice Compression
- Voice Activity
- Packet Loss
- Echo
- Fax Support
- Summary
- Lesson Review

Speech Coding

This topic discusses coder/decoder (codecs) in Voice over Data products.



Codecs use compression algorithms to reduce bandwidth consumption. Bandwidth consumption is illustrated on the Y-axis. Voice quality, which is measured subjectively by mean opinion score (MOS), is reflected on the X-axis.

Uncompressed digital voice, pulse code modulation (PCM), requires 64 kbps of bandwidth. If this voice stream is compressed as indicated on the Y-axis, that same voice can be transmitted in 8 kbps or less. Transmitting voice in 8 kbps is accomplished with coding algorithms, or loose algorithms that are optimized for voice. However, the tradeoff is in quality, as indicated on the X-axis. Generally, the higher the bandwidth, the higher the quality of voice becomes.

Voice compression schemes have varying bit rates. To understand voice compression methods, you must understand the three main forms of digital speech coding, which are waveform, vocoders, and hybrids.

Coding Algorithm Reference

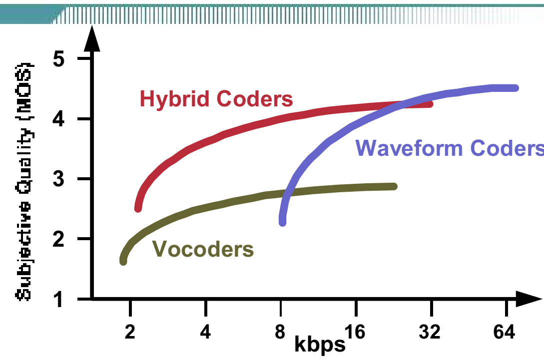
- Pulse code modulation (PCM)
- Adaptive differential pulse code modulation (ADPCM)
- Linear predictive coding (LPC)
- Code excited linear prediction (CELP)

- Low delay CELP (LDCELP)
- Algebraic CELP (ACELP)
- Conjugate-Structure ACELP (CS-ACELP)

Mean Opinion Scores

Cisco.com

Measuring Mean Opinion Scores Covered in ITU-T P.800 Series



Rating	Level of Speech Quality	Distortion
5	Excellent	Imperceptible
4	Good	Just perceptible but not annoying
3	Fair	Perceptible and slightly annoying
2	Poor	Annoying but not objectionable
1	Unsatisfactory	Very annoying and objectionable

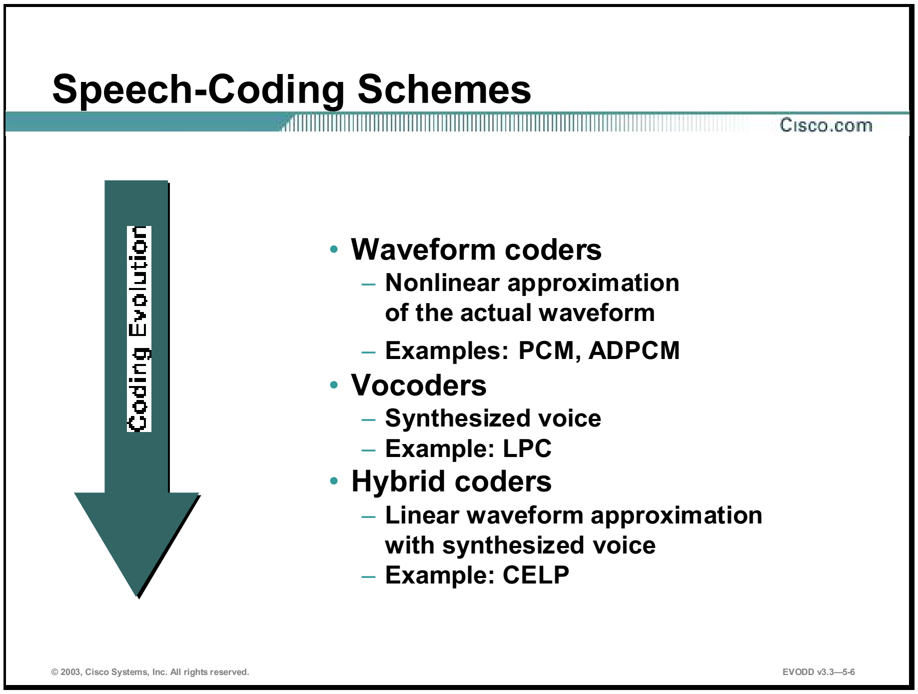
EVODD v3.3—5-5

The graph illustrates the Mean Opinion Score (MOS) approach to measuring voice quality.

Perceptual Speech Quality Measurement (PSQM) is a newer, more objective measurement, which is quickly overtaking MOS scores in the industry as the quality measurement of choice for coding algorithms. PSQM, as per ITU standard P.861, also provides a rating on a scale of zero to five. However, zero is rated as better quality, and five is the worst quality. Currently, various vendors have test equipment that is capable of providing a PSQM score for a test voice call over a particular packet network.

The key advantage to these compression schemes is that they require only a fraction of the bandwidth needed by traditional coding implementations. Perceived voice quality at the far end is good, because of the speed and power of the digital signal processors (DSPs). For example, 8-kbps CS-ACELP compares favorably with 32-kbps ADPCM in standardized tests of voice quality.

The graph shows the difference between the three types of coders: vocoders, hybrid coders, and waveForm coders.



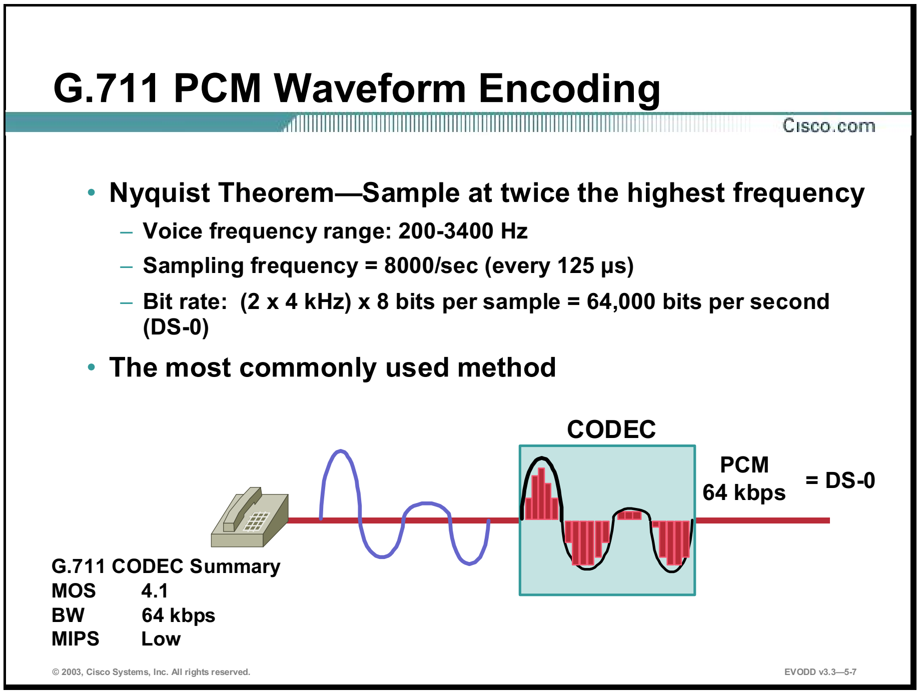
Waveform coders produce a nonlinear approximation of the actual waveform. The coders take samples of the form often enough that the original sound is approximated without the listener being able to tell the difference.

Vocoding schemes are low bit rates (2.4 kbps is possible) that sound synthetic and training is necessary to become accustomed to the voice. These are typically military applications where battlefield conditions are unpredictable and the synthesized output is always the same. Most of these approaches are half duplex.

The hybrid coding schemes are part of what is called Analysis-by-Synthesis (AbS) coding, which defines all generally used speech coding techniques in the 4.8 kbps to 16 kbps range. Since AbS coding continuously analyzes and learns to anticipate what a speech waveform should look like in the near-term (5 ms) future, it is a much higher quality than simple analysis or synthesis. A feedback loop allows the codec to continuously learn. These hybrid coders play a major role in Voice over Data implementations.

Waveform Encoding

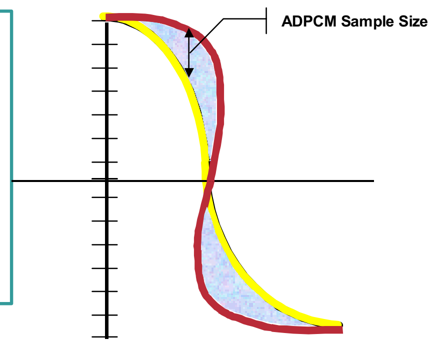
This topic explains the Nyquist Theorem. The Nyquist Theorem uses a sampling rate of twice the highest frequency and is the most commonly used method for waveform encoding.



Using the Nyquist Theorem, analog voice signals can be sampled at rates of 8 kHz, because the frequency range for voice is less than 4 kHz. If it is necessary to use 8-bit samples, the resulting bandwidth requirement is 64 kbps (for example, 8 kHz * 8 bits = 64 kbps). This rate is termed digital signal level 0 (DS-0) using PCM.

G.726 ADPCM

- Adaptive differential PCM
- Waveform coding scheme
- Adaptive: Automatic companding
- Differential: Encode the changes between samples only
- Rates and bits per sample:
  - 32 kbps = 8 kbps x 4 bits/sample
  - 24 kbps = 8 kbps x 3 bits/sample
  - 16 kbps = 8 kbps x 2 bits/sample



Codec Summary

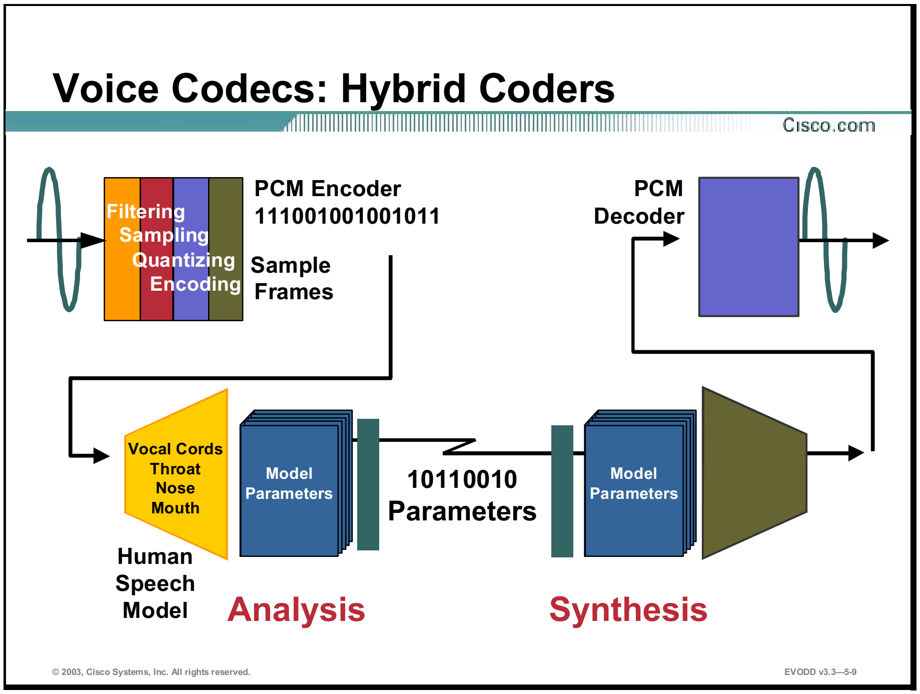
MOS	**3.85**
BW	**16, 24, or 32 kbps**
MIPS	**Medium**

EVODD v3.3—5-8

ADPCM is a form of pulse code modulation that produces a digital signal with a lower bit rate than standard PCM. ADPCM produces a lower bit rate by recording only the difference between samples and adjusting the coding scale dynamically to accommodate large and small differences. Some applications use ADPCM to digitize a voice signal so voice and data can be transmitted simultaneously over a digital facility normally used for voice or data only. G.726 ADPCM is usually found in traditional telephony use.

Hybrid Coders

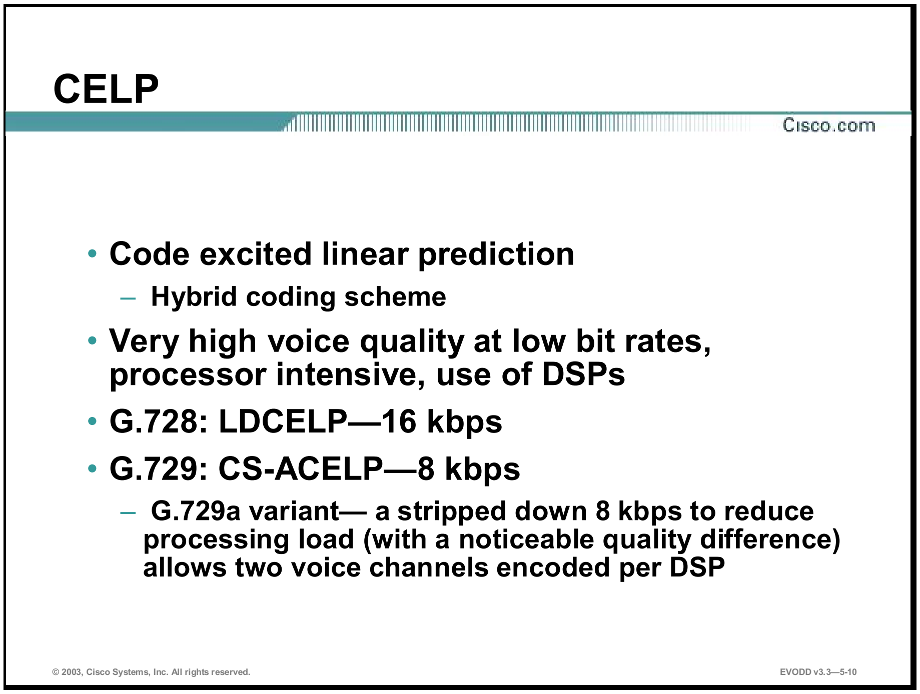
This topic examines hybrid coders. Early coders were based on sampling (PCM and ADPCM). The first evolution in coders was an attempt to use a model and synthesize voice linear predictive coding (LPC). The current generation of coders uses the best of both techniques, to provide a hybrid approach.



ADPCM and other coding schemes encode each sample of voice as a predefined value related to a specific frequency. Since every sound must be coded, whether it is the same or different than the last one, a large amount of bandwidth is required. In some cases, coding with fewer bits (the adaptive differential in ADPCM) reduces the required amount of bandwidth, but this coding severely impacts perceived voice quality due to loss of fidelity.

The CELP coders (LDCELP and CS-ACELP) are different from earlier generations of voice coding algorithms such as ADPCM. The CELP algorithms are voice-modeling schemes. The human voice box is modeled in the software loaded in the DSPs. The algorithms work by comparing incoming voice samples with idealized human speech models. Because human voices can only vary in certain ways from moment to moment, the next sample can be predicted to some degree. The difference from the model is coded as a small number of bytes of data and sent to the far-end DSP. At this point, the far-end DSP uses the computed differences to the standard model to synthesize voice for output to the listener. In other words, a much smaller amount of data can be sent because only the changes to the model are being transmitted, not the complete sample. The human voice output is synthesized from the model, not decoded from the data stream.

An analogy would be to send a text message to Mom's cell telephone using abbreviations instead of complete sentences. For example, by merely sending Mom the letters A, D, and B, she might be able to use her own “model” to determine that: (A) the kids are fine, (D) working at Cisco is terrific, and (B) the caller is sorry he or she does not call more often.

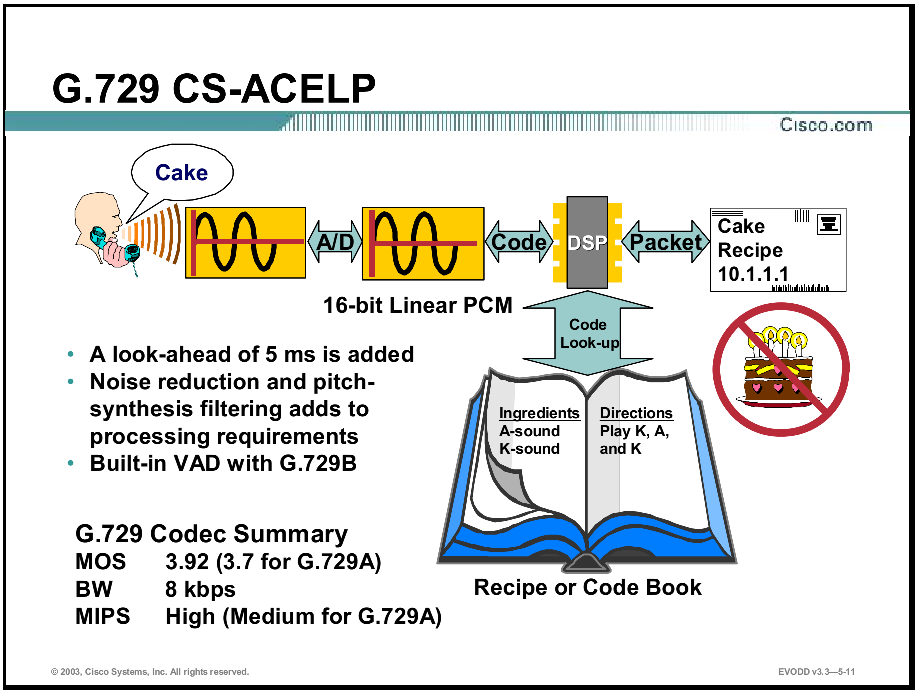


CELP algorithms use a hybrid-coding scheme. CELP coders are optimized for speech and can lower the bit rate without losing quality. CELP algorithms include ACELP, LDCELP, and CS-ACELP.

The coders cause up to a 25 ms delay. A linear filter is applied to the sound produced by the vocal cords. The coder then plays a sound closest to the one that sounds like the speaker. The filter breaks the sound down into voice and unvoiced sounds, and takes 20 ms to reproduce the voice.

CELP coders are used in most wireless telephones, Voice over Frame Relay (VoFR), and Voice over IP (VoIP).

With CELP, a codebook is built of different excitation sequences. A codebook allows a codec to substitute frequently used patterns with a smaller “code,” thereby using less bandwidth. The algorithm matches the codebook that most closely matches the signal.



G.729 uses the CS-ACELP algorithm. CELP voice compression techniques apply additional knowledge of speech and a codebook. The speech is analyzed and compressed with a vector quantizer. The excitation is transmitted as an index into a large vector quantizer codebook, along with a gain term to control its power. Typically, the codebook index is represented with about 10 bits (to give a codebook a size of 1024 entries), and the gain is coded with about 5 bits.

This technique has produced relatively good speech at bandwidths as low as 2.4 kbps. The CELP standard is at 4.8 kbps, but some drawbacks included computational power and introduced delay.

Examples of voice coder types include:

- **G.728:** Describes LDCELP. Voice compression requires only 16 kbps of bandwidth. CELP voice coding must be transcoded to a public telephony format for delivery to or through telephone networks.

- **G.729:** Describes further improvements known as CS-ACELP compression. This compression enables voice to be coded into 8-kbps streams. There are three forms of this standard, which include G.729 and two variations, G.729A and G.729B. All three forms provide speech quality as good as that of 32-kbps ADPCM.

- **G.729A:** Uses a less processor-intensive algorithm than G.729, at the expense of a slight degradation in voice quality.

- **G.729B:** Has embedded voice activity detection (VAD) and comfort noise support.

Note In addition to the basic technology, voice card selection requires compatibility verification of the Cisco IOS router, the voice feature card VCWare, and the DSPWare.

G.728 LDCELP: Low-Delay CELP

Cisco.com

- Similar to CS-ACELP except LDCELP uses smaller codebook and operates at 16 kbps to minimize delay to between 2-5 ms (no look-ahead)
- 10-bit codeword is produced from every 5 samples of speech from the 8-kHz input
- 4 of these 10-bit codewords are called a subframe which takes approximately 2.5 ms to encode
- 2 of these subframes are combined into a 5-ms block for transmission

G.728 Codec Summary

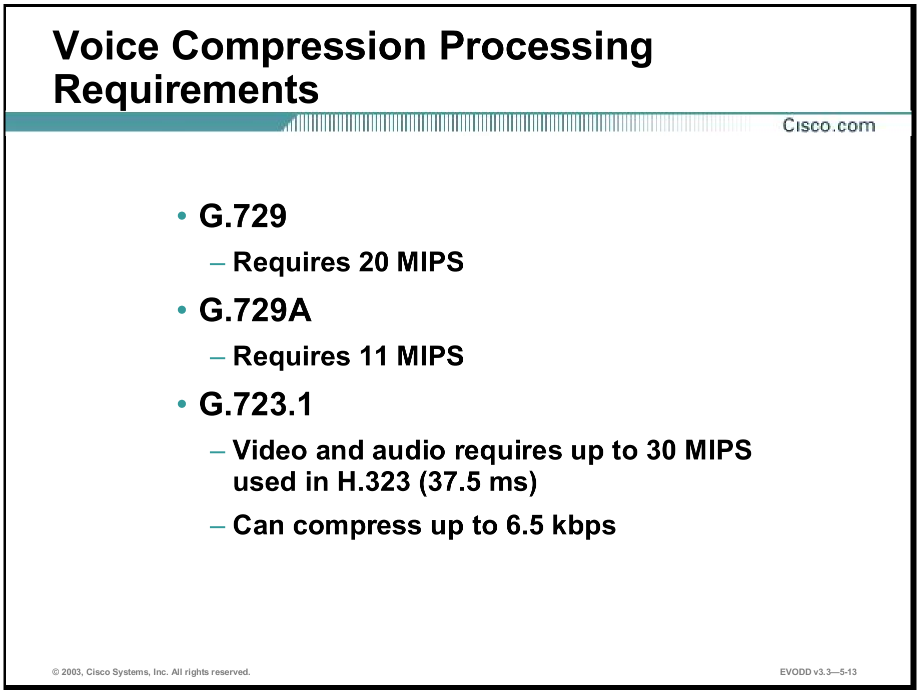
MOS	3.61
BW	16 kbps
MIPS	High

EVODD v3.3—5-12

G.728 LDCELP provides low delay but at a high-processing cost. The delay is 2 to 5 ms, with no look-ahead. Currently, IP telephones do not support G.728 and the standard is not used.

Voice Compression

This topic details the processing power required to support various codecs.



Another tradeoff when compressing voice is that CPU processing power is required. CPU processing power is measured in millions of instructions per second (MIPS) and becomes important when trying to squeeze multiple channels (calls) on a single DSP. For example, G.729 requires 20 MIPS processing per channel, but G.729a is a lower-complexity codec and is almost twice as effective. From a network design perspective, this affects the choice of codecs only, which affects the number of calls allowed over the same hardware configuration.

Digital Speech Interpolation

Cisco.com

- Voice activity detection (VAD)
- Removal of voice silence
- Examines voice for power, change of power, frequency, and change of frequency
- All factors must indicate the voice fits into the window before cells are constructed
- Automatically disabled for fax/modem

EVODD v3.3—5-14

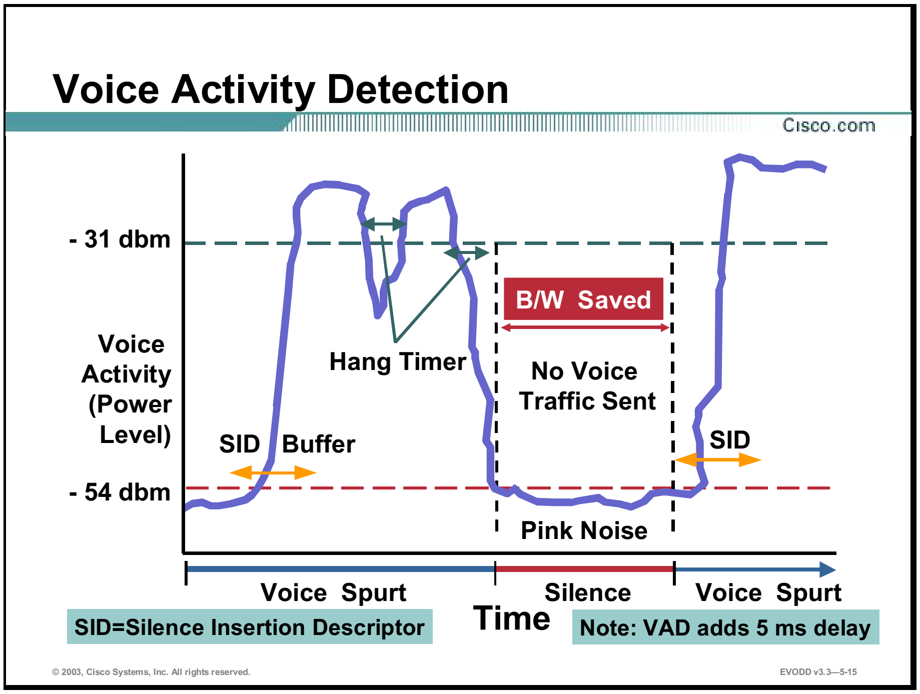
Another characteristic of codecs is called digital speech interpolation (DSI). DSI scans the speech pattern for silence, then inserts and sends an indication of the silence to the other end. These indicators are called Silence Insertion Descriptors (SIDs).

DSI is very effective in real-time bandwidth efficiencies. In some instances, VAD can remove 65 percent of voice packets and renders a full-duplex conversation half-duplex as far as the network is concerned. All unused bandwidth is made available to other bursty applications.

VAD is built into some codecs, like G.729B. Some users turn off VAD because the trained ear notices sounds clipped at the beginning of speech after a silence period.

Voice Activity

This topic examines voice activity.



VAD and the dejitter buffer are part of the voice activity. If VAD is enabled, whether part of the codec or through Cisco IOS software, SID packets are played out in the dejitter buffer. The value, of course, is bandwidth savings because up to 65 percent of a voice conversation is silence. If you send an indicator of the silence, instead of the actual sounds of silence, you can experience significant bandwidth savings.

True silence is uncomfortable to the listener, who might think the call is dropped ("Are you still there?"). SIDs are used to approximate normal background noise. White noise is a canned background noise. Pink noise approximates the ongoing background noise in a conversation. For example, speaking from an airport has different background noise than speaking from the office. Therefore, pink noise may be more appropriate than white noise for certain situations.

If VAD is enabled, a 5-ms look-ahead buffer is required. Switching VAD on or off is a configuration option. VAD should be calculated into the delay budget, if used, because it adds 5 ms of delay on the outbound voice channel.

Voice Codec Cheat Sheet

Cisco.com

Encoding	Quality MOS	Native Bit Rate kbps	DTMF	Dual Comp	CPU	Music On Hold
G711 PCM	A 4.1	D 64	A	A	A	A
G726 ADPCM	B 3.85	C 32	B	B	B	B
G728 LDCELP	B 3.61	B 16	B	C	C	C
G729 CS-ACELP	A 3.92	A 8	B	B	C	C
G729a CS-ACELP	B 3.7	A 8	C	C	B	D
G723.1 ACELP	C 3.65	A 5.3	C	D	C	D

A = Excellent B = Good C = Fair D = Poor

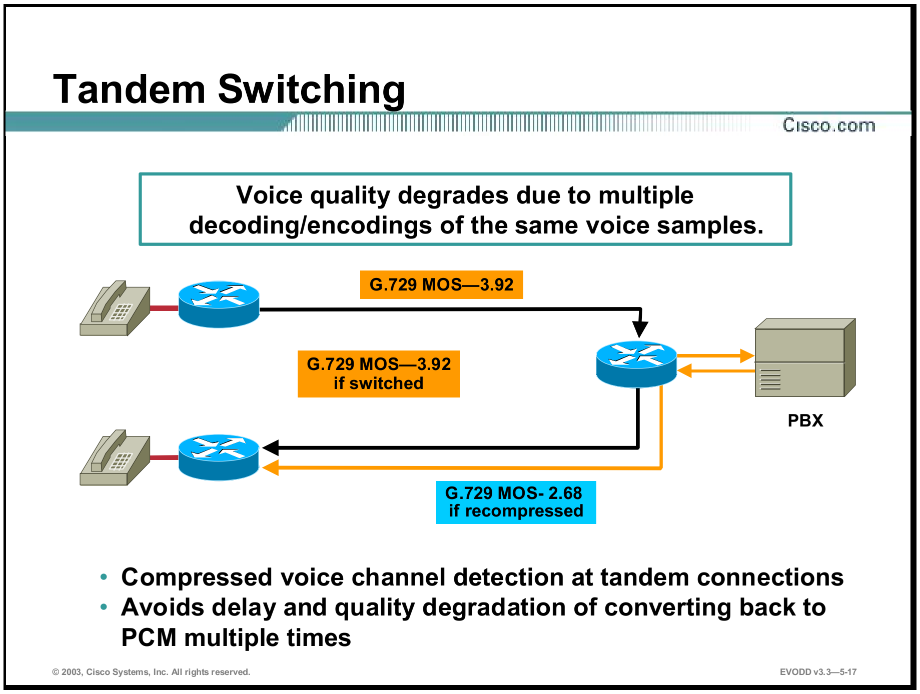
Dual Comp = Dual Compression

EVODD v3.3—5-16

In the voice arena, the price of voice processing hardware has fallen dramatically as the MOS generated by different voice compression algorithms has increased considerably.

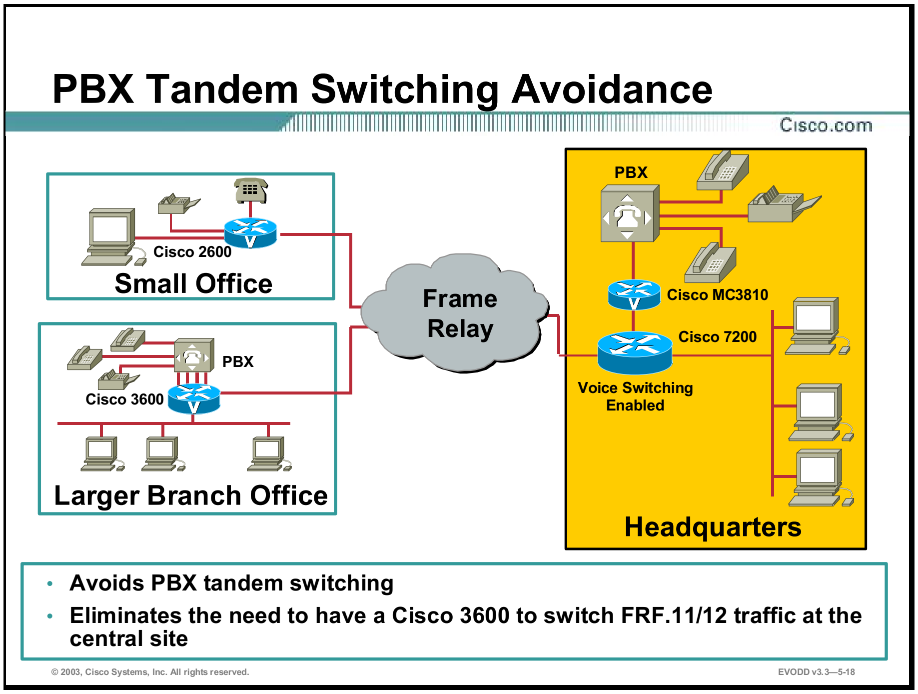
For example, if the MOS of compressed voice rises, then the amount of bandwidth required to carry voice drops. This enables new business applications like Internet commerce and customer service, augmented by Web-based voice services.

This matrix shows the characteristics of different codecs in comparison to others. Dual tone multifrequency (DTMF) refers to how well the encoding supports tones used for touch-tone dialing. Compressing DTMF tones can cause tones to become unintelligible to the applications that require tones for control, such as voice mail or interactive voice response (IVR). In addition, since music on hold (MOH) behaves differently than speech, its perceived quality may be less than speech when using high-compression codecs.



Dual compression refers to the performance of the different codecs in tandem switching scenarios. A problem with voice quality is encountered in tandem switching. This problem occurs when a compressed voice stream is decompressed at the router, transferred to the PBX, forwarded to another location, and recompressed at the router en route to the final destination.

Cisco equipment is designed to recognize tandem hops and switch the compressed voice to the end destination without decompressing and recompressing.



An integral part of providing multi-service connectivity to remote sites is the ability to switch voice between sites.

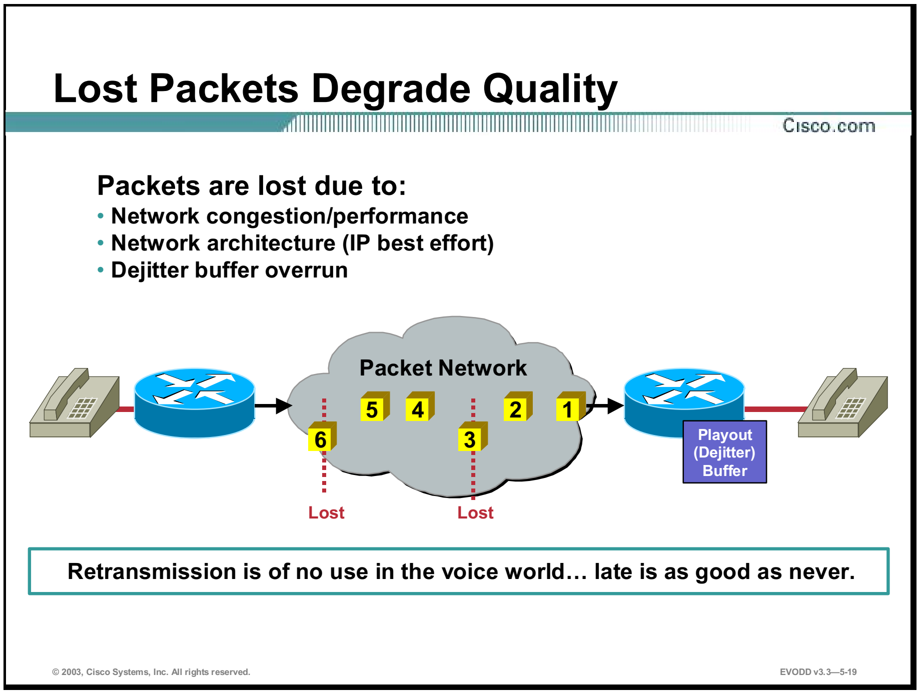
The VoFR/VoIP products of some vendors rely on the PBX to perform this tandem switching of voice calls from one location to another. However, most telecom managers generally perceive tandem switching in a negative light for two reasons. First, tandem switching ties up two trunks on the PBX by taking a call in on one line and switching it to another. Second, tandem switching of VoIP/VoFR though a PBX can cause degraded voice quality. Voice coming into the PBX must be decompressed and then recompressed going out of the PBX.

Cisco integrated multi-service architecture allows VoFR/VoIP calls to bypass the traditional tandem PBX switch. Cisco performs this tandem PBX bypass in the router, avoiding the use of the valuable PBX trunk lines and multiple voice encoding.

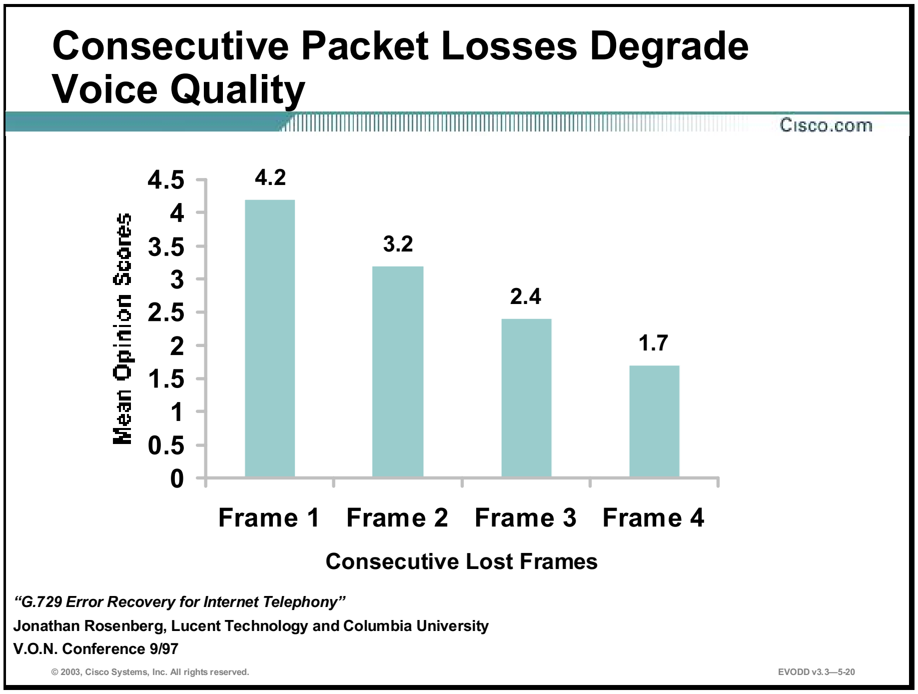
Note	Cisco tandem PBX bypass feature is especially beneficial in the deployment of large VoFR/VoIP networks.

Packet Loss

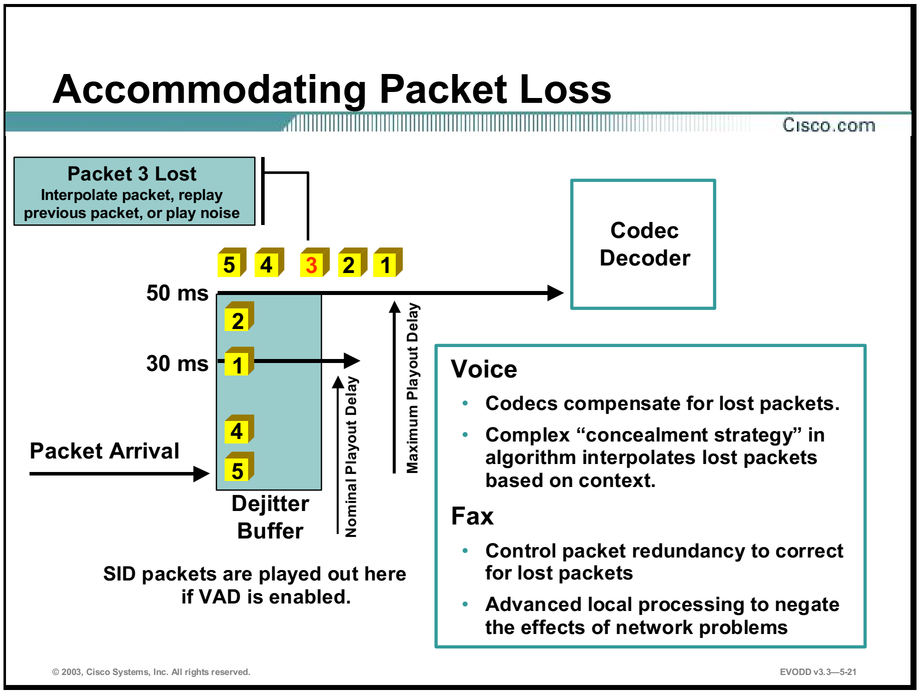
This topic identifies potential causes for packet loss. Lost packets can be a severe problem, depending on the type of packet network that is being used.



Because IP networks do not guarantee service, they will usually exhibit a much higher incidence of lost packets than ATM networks. Data frames are not time sensitive and dropped packets can be appropriately corrected through the process of retransmission. Loss of multiple voice packets, however, cannot be recovered in a way that is transparent to the calling parties.



This chart shows the drop in quality, as measured by MOS, when lost frames occur consecutively.



The dejitter buffer is a queue that induces delay, and smoothes out the inevitable variance in packet interarrival times. The size of this queue is both variable and adaptable in Cisco IOS software. The nominal playout delay is the average time allowed for packets to be played out to the codec decoder module, which converts packets into a voice stream. The maximum playout delay is the maximum time allowed for playout to occur. The values of these two settings balance how much delay is induced with voice quality. Using an optimizing algorithm, Cisco IOS automatically adapts these values to network conditions.

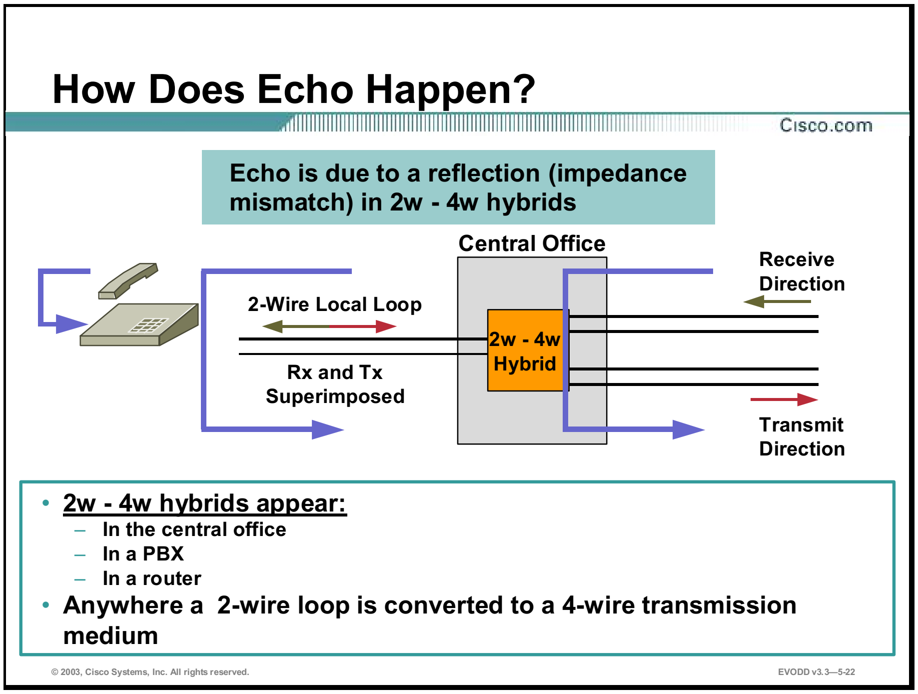
The three methods that can be used to play out a lost packet include: repeat the previous packet, insert some predetermined noise, or interpolate the sound based on the previous and following sound packet characteristics. The latter method provides the least drop-off in quality.

The illustration shows packets arriving in the dejitter buffer out of order: 2, 1, 4, 5, (3 does not show). When the nominal threshold is reached for packet 2, the playout software knows there is a problem since 1 has not arrived. The playout software waits until the maximum playout delay is reached. By then, 1 has arrived, and 1 and 2 are played out in proper sequence. However, 3 never arrives before that threshold is reached. Therefore, the lost packet algorithm interpolates packet 3 and plays this generated packet out in its place, followed by 4 and 5, which arrived in a timely fashion.

If VAD is enabled, SID packets are sent. SID packets cause the far-end router to generate comfort noise, so that the user does not think the call has disconnected.

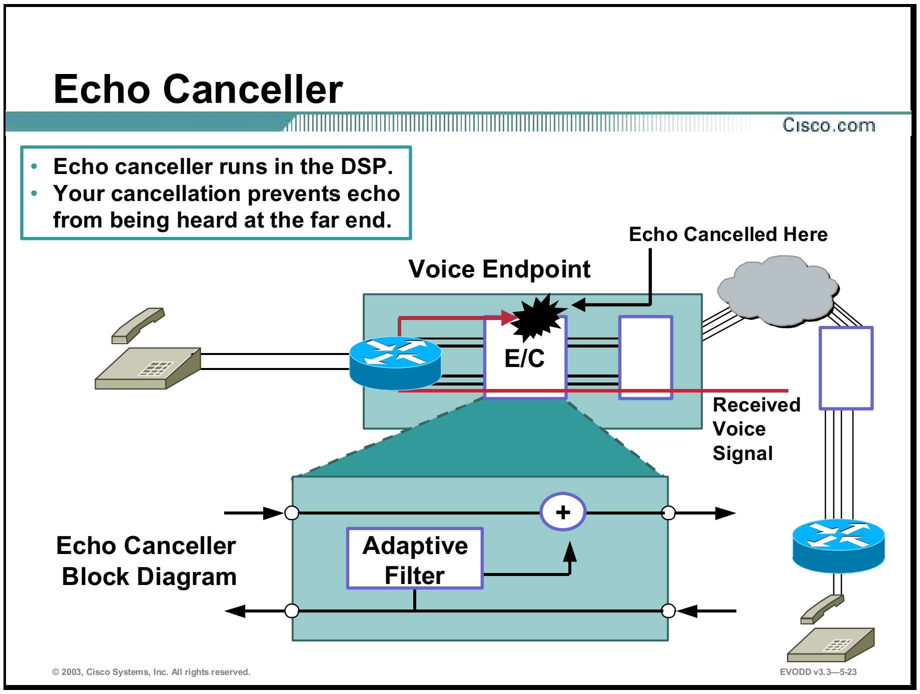
Echo

This topic addresses how to reduce echo in a Voice over Data network.



Echo in a telephone network is caused by signal reflections generated by the hybrid circuit that converts between a 4-wire circuit (a separate transmit and receive pair) and a 2-wire circuit (a single transmit and receive pair).

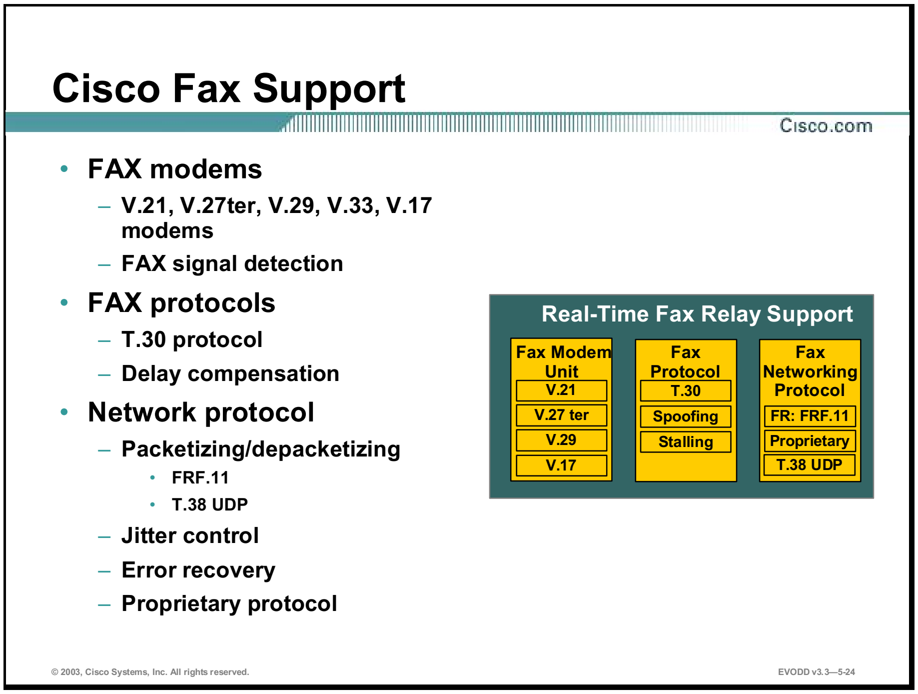
The reflections of the speaker's voice are heard in the speaker's ear. Echo is present even in a conventional circuit-switched telephone network. However, it is acceptable because the round-trip delays through the network are less than 50 ms and the echo is masked by the normal side tone that every telephone generates.



Echo becomes a problem in Voice over Data networks, because the round-trip delay through the network is typically greater than 50 ms. Cisco routers can cancel echo by generating a signal that is approximately 180 degrees out of phase with the detected echo. When these two signals are combined, they nearly cancel each other out.

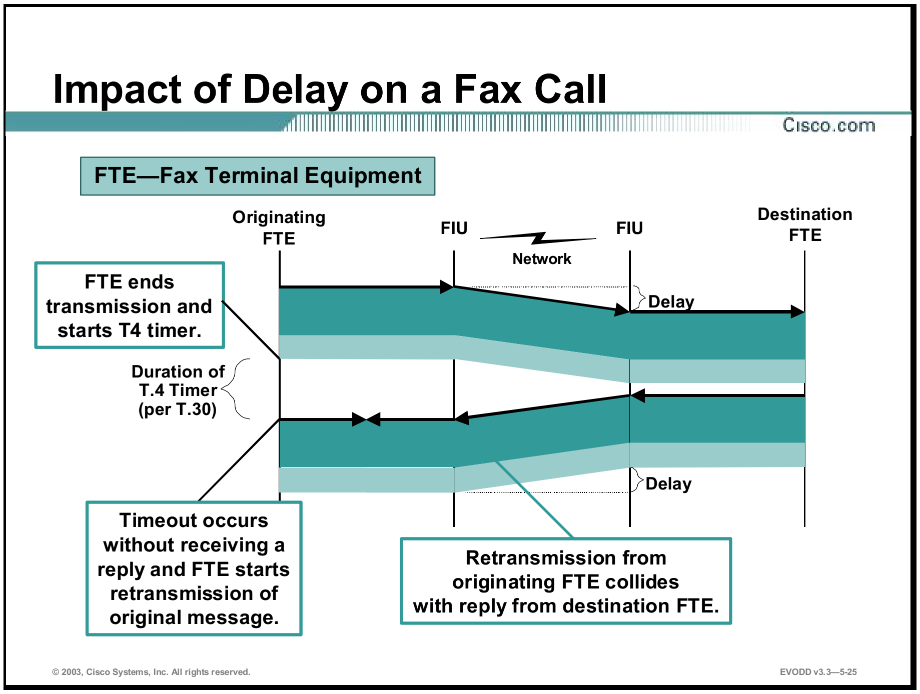
Fax Support

This topic addresses the Cisco approach to supporting fax transmission across a Voice over Data network.



Fax is highly sensitive to analog signal quality. On the analog-to-network side, DSPs must demodulate analog fax signals and packetize data per the network protocol. On the network-to-analog side, the DSP must depacketize data per the network protocol and re-modulate the analog signals. If voice and fax are enabled on a given channel, the sending side recognizes the fax, switches to fax mode, and sends the fax. The receiving side detects the fax preambles and switches to fax mode. There is no fax/voice negotiation between the endpoints; instead, the initiating end recognizes and switches the mode.

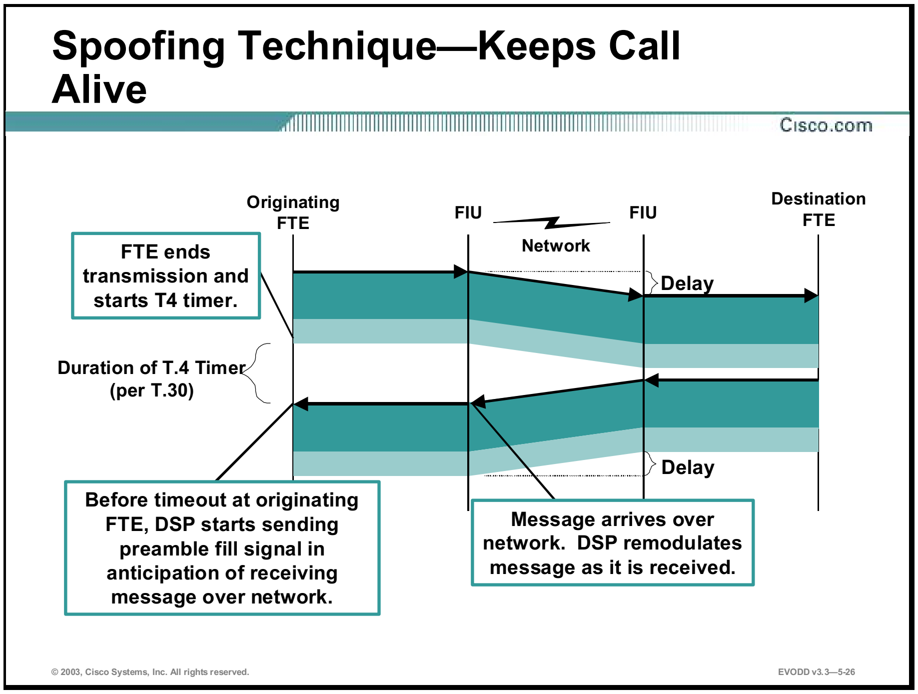
T.30 is the standard analog fax protocol. Network protocols supporting fax transmission over a data network include T.38 UDP, FRF.11, or T.38 encapsulated in Real-Time Transport Protocol (RTP).



Fax calls are more sensitive to latency than voice calls. The fax protocols in use were established with low-latency PSTN connections in mind. The calling fax machine expects a response from the receiving fax machine, and if it does not receive that response in a set time (usually 3 seconds in automatic mode), it retransmits the original message. This constant retransmission could render a fax relay solution over packet networks inoperable without proper compensation techniques built into the fax over data solution.

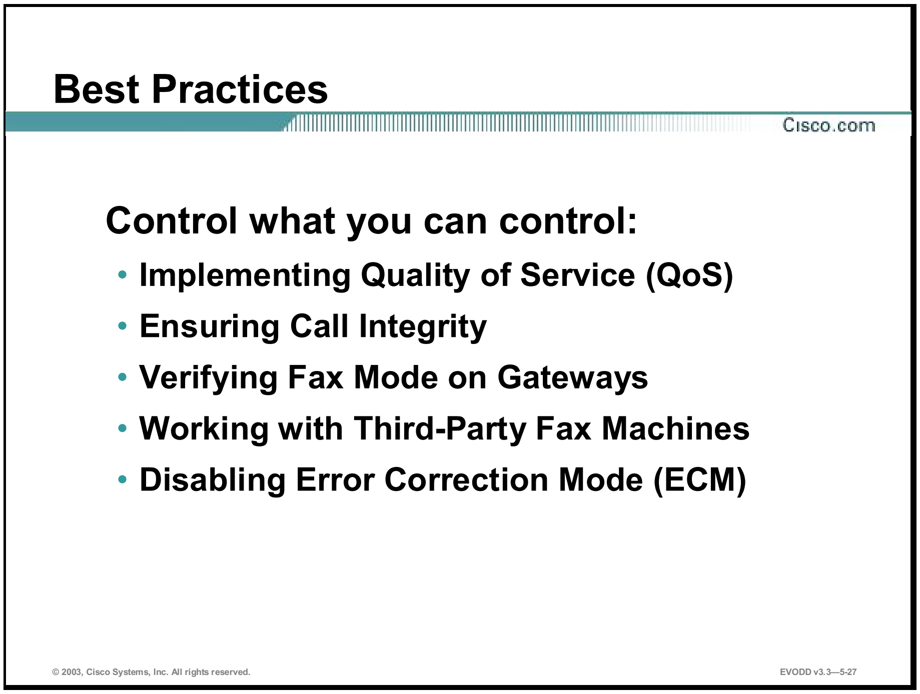
Delay in the DSP can occur with analog signal detection, packet framing, delay and jitter compensation, and interoperability compensation. Network delays occur with transport protocols, packet retransmissions, and congestion/routing problems.

Fax Terminal Equipment (FTE) devices use synchronous modems. If network packets are late, something must still be transmitted to the modem. There is no built-in flow-control once the transmitting of page data starts, and no way to tell the FTE to wait or slow down. An adaptable dejitter buffer can help alleviate this problem.



The Cisco fax solution provides a robust, patented solution to this latency issue via spoofing techniques. The transmitting fax interface unit spoofs the transmitting fax machine by providing the initial call response prior to its arrival from the destination FTE, and then completes the response as it is received from the destination FTE.

Fax uses low-speed redundancy packets and high-speed packets for the data payload. For signaling, the DSP sends three copies of each data byte to ensure reliability.



Network designers can focus on five areas to improve voice quality:

Implementing Quality of Service (QoS)

QoS tools can assist you in lessening packet drops and delays to improve fax transmission. On Cisco voice gateways the QoS parameters do not distinguish between fax and voice calls. However, by using QoS tools, you can still improve the quality of service for IP telephony transitions, including voice and fax traffic.

Ensuring Call Integrity

These tips can help ensure the integrity of the fax calls:

- Use Call Admission Control (CAC) to ensure calls for which the required quality of service cannot be guaranteed are not admitted.
- Use G.711 across the WAN to avoid any potential CAC issues.
- Disable call waiting on all dedicated modem and fax ports

Verifying Fax Mode on Gateways

The default fax mode on all Cisco IOS gateways is Cisco fax relay. For best performance, verify that you have Cisco fax relay on both the originating and terminating gateways. If two IOS gateways have differing transports, they will negotiate to use Cisco Fax relay.

Working with Third-Party Fax Machines

Cisco fax relay has been tested in multiple network scenarios with several brands of Group 3-compatible fax machines, including:

- Xerox WC-450
- Brother FAX 560
- Cannon Multi Pass C-775
- Hewlett Packard Office Jet 710
- Ricoh Fax 1900L
- Sharp UX 340LM
- Panasonic KX-FP-81
- Okidata OKI fax 5400
- Pitney Bowes 9920 and 9930

These fax machines, configured for T.30 signaling, were found to be interoperable with Cisco Fax Relay.

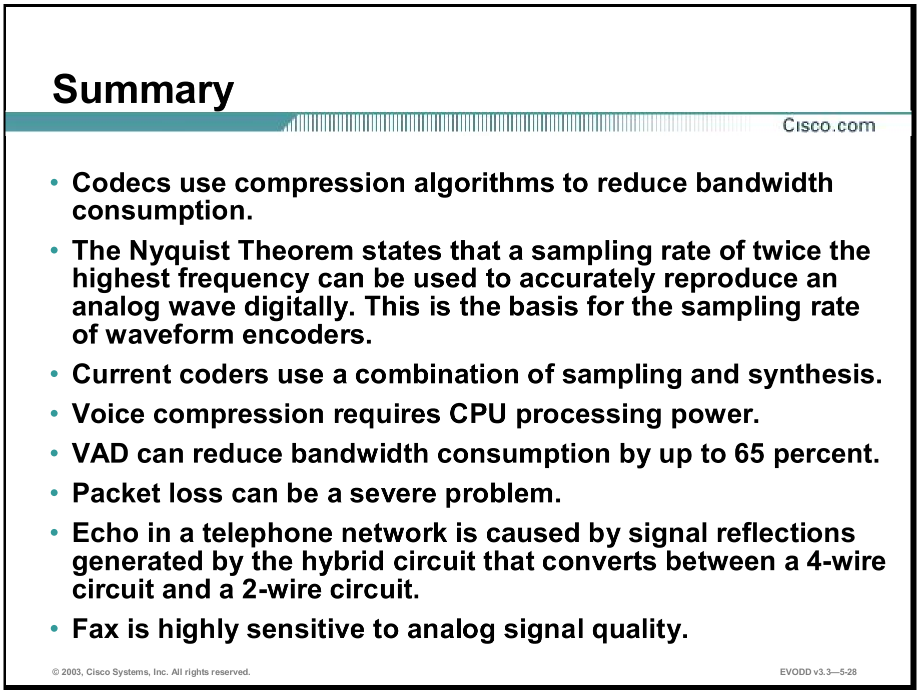
For fax rates greater than 14.4kbps, Cisco voice gateways switch to fax pass-through mode even if they are configured for Cisco fax relay. However, fax machines with ECM enabled are highly sensitive to network condition such as packet delay and packet loss.

Disabling Error Correction Mode (ECM)

Error Correction Mode (ECM) is intended to eliminate errors in the fax transmission. By default, ECM is enabled on Cisco IOS voice gateways. With ECM enabled, if packets drop, the receiving fax machine generates retransmission requests. If the packet drop is excessive, the call duration increases, and the call might be dropped.

Summary

This topic summarizes the key points discussed in this lesson.



Next Steps

After completing this lesson, go to:

- Voice Over Data Migration module

References

For additional information, refer to these resources:

- Waveform Coding Techniques:

  http://www.cisco.com/warp/public/788/signalling/waveform_coding.html

- Solving Voice Problems:

  http://www.cisco.com/univercd/cc/td/doc/product/voice/ics7750/tblshoot/tsvoice.htm

Lesson Review

This practice exercise reviews what you have learned in this lesson.

Q1) Which three of the following types are voice coders? (Choose three.)

A) waveform coders

B) vocoders

C) hybrid coders

D) linear coders

Q2) G.711 is an example of which voice coder type?

A) waveform

B) vocoder

C) hybrid

D) linear

Q3) G.729 is an example of which voice coder type?

A) waveform

B) vocoder

C) hybrid

D) linear

Q4) Which of the following codecs require 20 MIPS of CPU processing power?

A) G.729

B) G.729A

C) G.729B

D) G.723.1

Q5) Background noise varies based on location. Which of the following is synthetically generating noise approximating the appropriate background noise for a conversation?

A) white noise

B) purple noise

C) red noise

D) pink noise

Q6) Which three of the following methods accommodate a lost packet? (Choose three.)

A) expand the playout time of the previous packet

B) interpolate the packet

C) replay the previous packet

D) play noise

Q7) What is the primary cause of echo?

A) a dB level that is too high on the transmit side

B) a dB level that is too high on the receive side

C) an impedance mismatch in 2-wire to 4-wire hybrid circuits

D) Electromagnetic Interference (EMI)

Q8) What is the standard analog fax protocol?

A) T.120

B) T.30

C) FRF.11

D) T.38

6

Voice Over Data Migration

Overview

A quality of service (QoS)-enabled design is the prime prerequisite of a migration from a traditional PBX-based telephony system to a Voice over Data system. This module discusses key migration considerations emphasizing the Cisco QoS tools. This module also presents a migration case study, detailing a large-scale migration, to bring together the various design considerations presented throughout the course.

Upon completing this module, you will be able to:

- Identify the recommended QoS tools available to ensure voice quality in Voice over Data networks
- Define prioritization, traffic-shaping, and link efficiency tools
- Define admission control tools
- Specify where to use QoS tools
- Apply knowledge of QoS tools in Voice over Data design
- Identify existing PBX features to migrate
- Select appropriate LAN and WAN components to support voice requirements
- Specify Cisco CallManager (CCM) and Cisco Unity requirements to support a specified number of users

Outline

The module contains these lessons:

- Voice Quality Overview

- Comprehensive Design Strategies

Voice Quality Overview

Overview

You can prevent potential problems while running Voice over Data networks by establishing quality of service (QoS) for the enterprise campus and the WAN. This lesson teaches you about the Cisco QoS features. You will learn to design and migrate Voice over Data networks using QoS tools to ensure voice quality.

Importance

Users accustomed to PBX-based telephony systems are not tolerant to voice delay, voice packet drops, or jitter. Therefore, QoS is one of the most critical design considerations to preserve voice quality in a migrated Voice over Data network.

Objectives

Upon completing this lesson, you will be able to:

- Describe requirements for Voice over Data networks
- Identify the recommended QoS tools available to ensure voice quality in Voice over Data networks
- Describe prioritization tools
- Describe slow link efficiency tools
- Describe traffic-shaping tools
- Describe how to enable the campus infrastructure with QoS
- Describe how the Cisco IP Phone works with Cisco Catalyst switches
- Define admission control tools
- Specify where to use QoS tools

Learner Skills and Knowledge

To fully benefit from this lesson, you must have these prerequisite skills and knowledge:

- *Building Cisco Remote Access Networks* (BCRAN) course
- An understanding of legacy queuing techniques, including: custom queuing, priority queuing, and weighted fair queuing

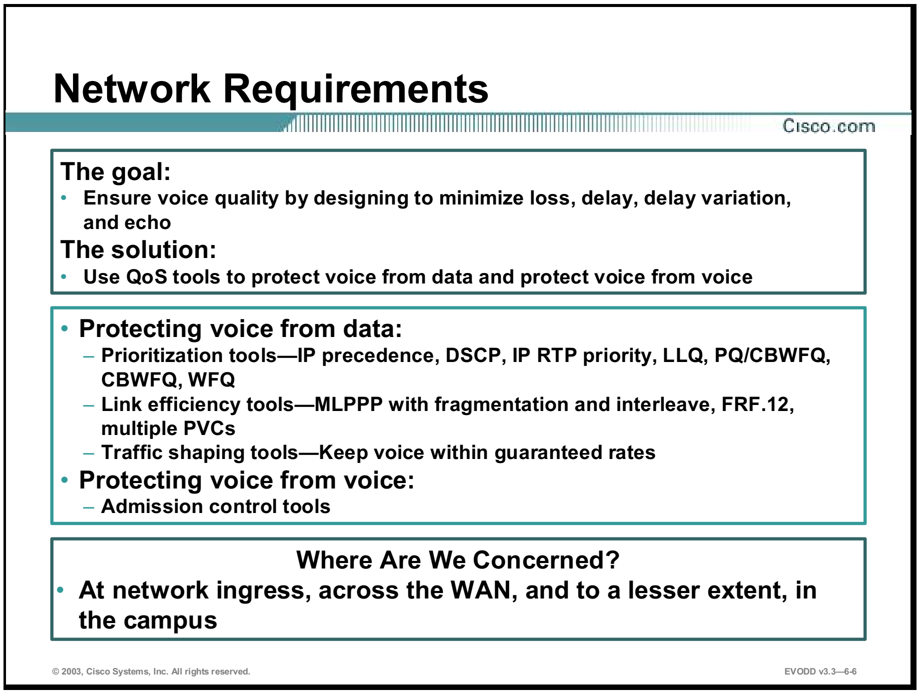
Outline

This lesson includes these topics:

- Overview
- Network Requirements
- Identifying QoS Tools
- Prioritization Tools
- Slow Link Efficiency Tools
- Traffic-Shaping Tools
- Prioritization
- Inside the IP Phone
- Admission Control
- Transport Options
- Summary
- Laboratory Exercise: QoS
- Lesson Review

Network Requirements

This topic details the requirements of Voice over Data networks, which are highly sensitive to delay. To ensure voice quality, the network must minimize loss, delay, delay variation, and echo.



Excellent voice quality can traverse an IP WAN, regardless of media or very low data rates, when prerequisite tools are applied. For this level of quality to occur, you must protect packets carrying voice payload from packets carrying other types of data payloads such as FTP, Telnet, or HTTP. In addition, you must constrain the amount of voice traffic introduced to the network to prevent oversubscribing a defined voice capacity.

Protecting voice from data

Because voice traffic competes with data traffic for timely delivery, you must apply certain tools to ensure voice quality. These tools fall into four primary areas:

- Classification
- Prioritization
- Slow-speed link efficiency
- Traffic shaping

Protecting voice from voice

Voice traffic can impede other voice traffic if available bandwidth is oversubscribed, even with voice-friendly quality of service (QoS) tools enabled. You can ensure voice quality by using recommended design practices.

QoS Tools

Cisco.com

- **Current recommendation:**
  - **Low latency queuing (LLQ) or priority queuing/class-based weighted fair queuing (PQ/CBWFQ)**
  - **IOS 12.0(7)T for IP; 12.1(2)T for Frame Relay**
- **Evolution of classification and queuing tools:**
  - **Weighted-fair queuing (WFQ)**
  - **Class-based WFQ**
  - **Priority queuing WFQ (IP RTP Priority)**
  - **Priority queuing/CBWFQ (PQ/CBWFQ)**

The mechanisms of QoS allow you to control the mix of bandwidth, delay, jitter (variances in delay), and packet loss in the network. With this control, you can deliver a network service (such as Voice over IP [VoIP]), define different service level agreements (SLAs) for divisions, applications, or organizations, or simply prioritize traffic across a WAN.

Having a voice-enabled infrastructure in place is a key prerequisite of designing a Voice over Data solution. The QoS mechanisms must allow for the prioritization of voice traffic and ensure its timely delivery ahead of other data.

You must understand the mechanisms used in planning and implementation to have a successful QoS deployment. These mechanisms vary by Cisco IOS release, and they add resource demands on the routers involved. As a designer, you must understand the following requirements:

- Which products you need to support voice
- Which Cisco IOS release you must use
- Whether you need additional RAM to support the processing overhead

Each subsequent Cisco IOS release builds upon previous tools to continually improve QoS. Currently at the top of the evolutionary QoS chain is low latency queuing (LLQ), also called priority queuing or class-based weighted fair queuing (PQ-CBWFQ).

Traditional Data Networks

Cisco.com

Traditional data traffic *tolerates*:

- **Bursty data flow**
- **First-come, first-serve access**
- **Data rate adaptive to network conditions**
- **Brief outages**

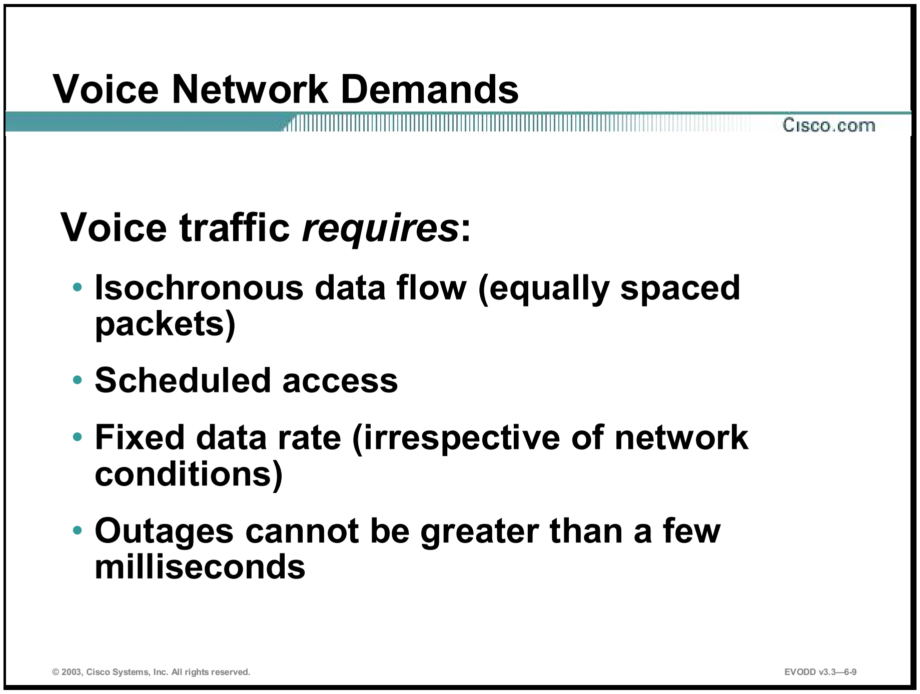
EVODD v3.3—6-8

When assessing the problems of running a Voice over Data network, you must first consider the characteristics of the traditional data network.

A traditional data network is based on bursty, or uneven, data flow. The data comes in packets that take as much bandwidth as possible. Access is on a first-come, first-served basis, so the packets that arrive first get the bandwidth they need. As a result, the data rate is adaptive to network conditions.

Note	Several data protocols have evolved over time to adapt to the bursty nature of data networks.

Many routing protocols converge when there is a link or hardware failure, ranging from a few seconds to a few minutes. These brief outages are survivable; for example, a delay of a few seconds is generally not noticeable if you are retrieving your e-mail. An outage can cause delays that are several minutes in length and still be serviceable.



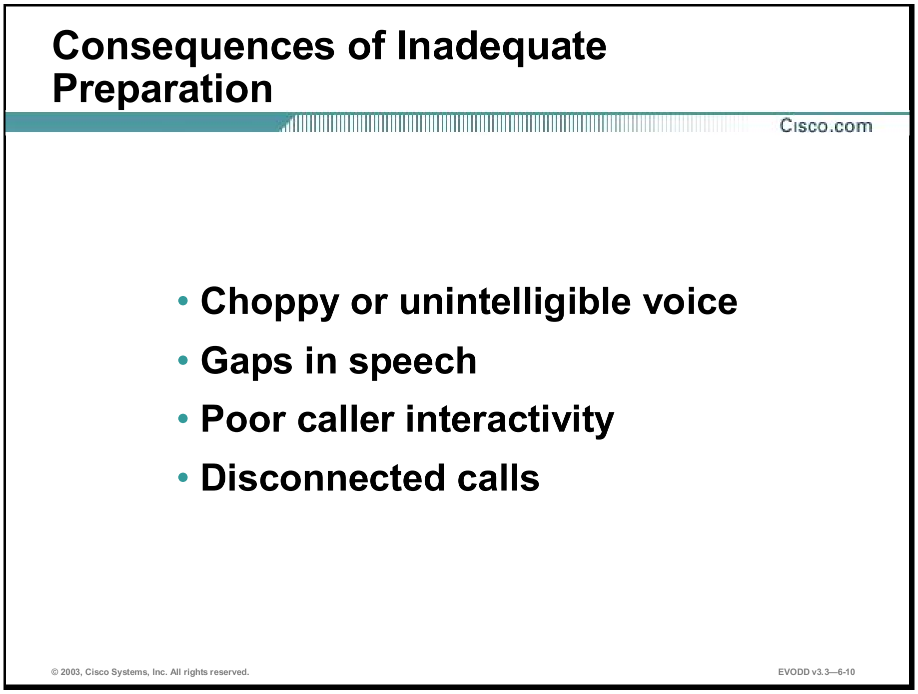
In voice and/or data networks, the data flow is isochronous, which means that the packets are spaced at equal intervals. Traditional networks are arranged in time slots known as time-division multiplexing (TDM). The access is scheduled. Specific bandwidth is also reserved for voice data.

Note The data rate is not adaptive to network conditions.

Outages of greater than a few milliseconds are not permissible. Users do not want to wait for the data network to come back in order for voice to be available, so silences are not acceptable on voice calls.

Identifying QoS Tools

This topic identifies QoS tools to combat the potential problems of running Voice over Data networks.



A choppy or unintelligible voice could result from a data network that is unprepared for voice running over it. For example, digital cell phone users may experience these gaps in speech when going through a tunnel or dropping out of transmitter range. Gaps in speech are particularly troublesome when pieces of speech are interspersed with silence; the speech literally disappears.

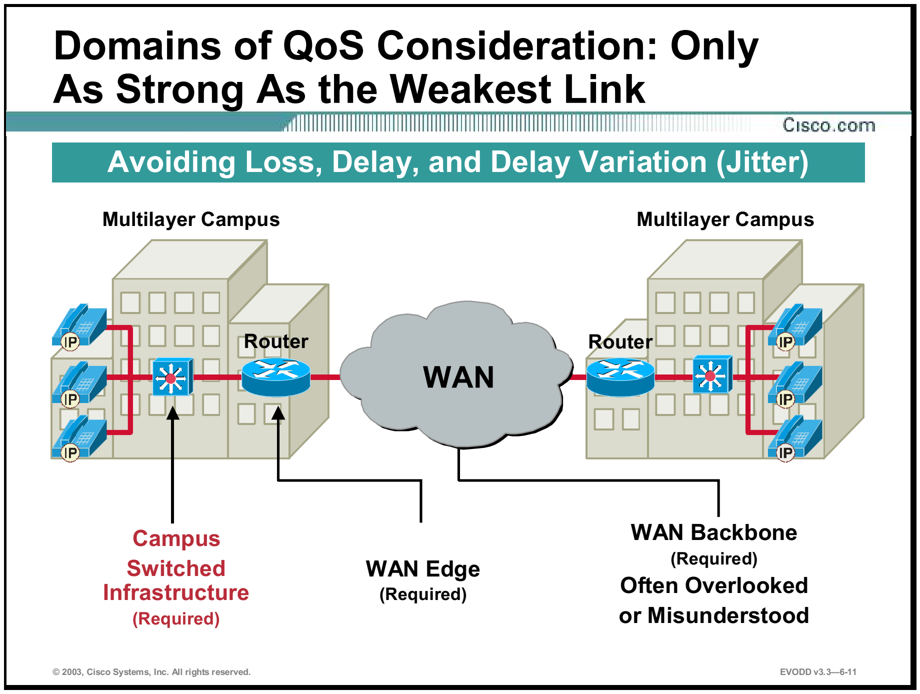
Example

A user dials 68614 in a voice mail system. In this case, the gaps in speech are gaps in the tone. The 68614 becomes 66881114 because the network perceives the gaps in the speech as pauses in the touch-tones.

Delay causes poor caller interactivity. Sometimes called the walkie-talkie effect, delay can result from a voice collision. For example, speech may take so long to reach the far end of the line that by the time the response comes back, the caller has spoken again and the voices collide. This problem is typically experienced on satellite links using traditional telephone services with long delays between callers talking.

Disconnected calls are the worst cases. Long gaps in speech may cause callers to hang up, or signaling problems may disconnect calls.

Note These events are unacceptable in the voice world but are quite common for inadequately prepared data networks that are attempting to carry voice.

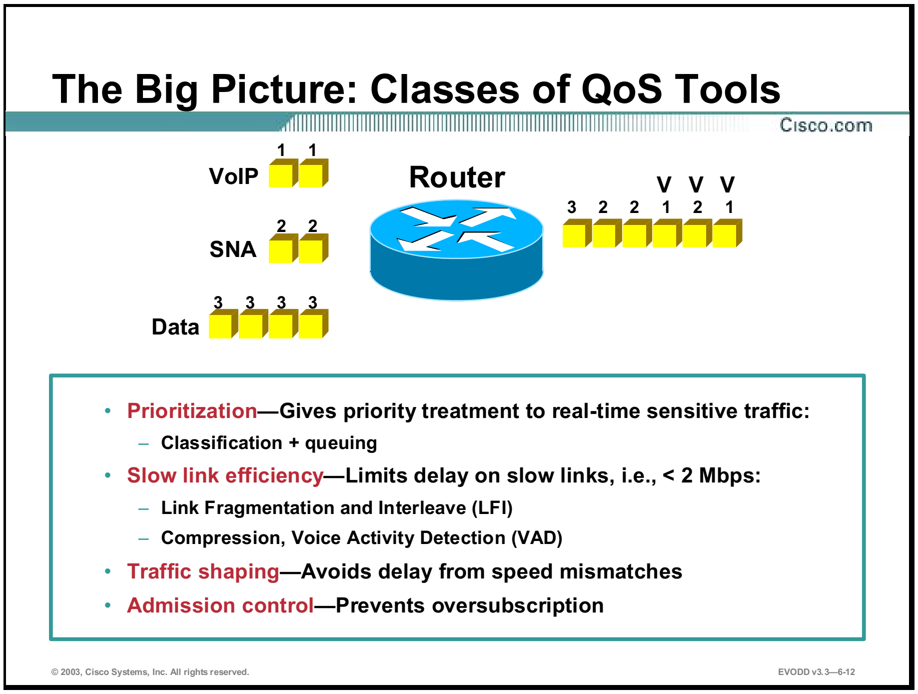


The WAN edge, or router egress, and the WAN itself are the most important QoS areas in the toll bypass arena and in end-to-end IP telephony.

Note	These areas must be planned, and the proper QoS tools must be enabled.

Higher bandwidths—up to 100 Mbps, Gb, or ATM speeds—prevail for the campus. QoS is important as bandwidth requirements increase.

The carrier is probably the most forgotten and most easily misunderstood component of QoS. You must negotiate a contract for carrier capacity, and then you may design a Voice over Data network within the specifications of the contract.



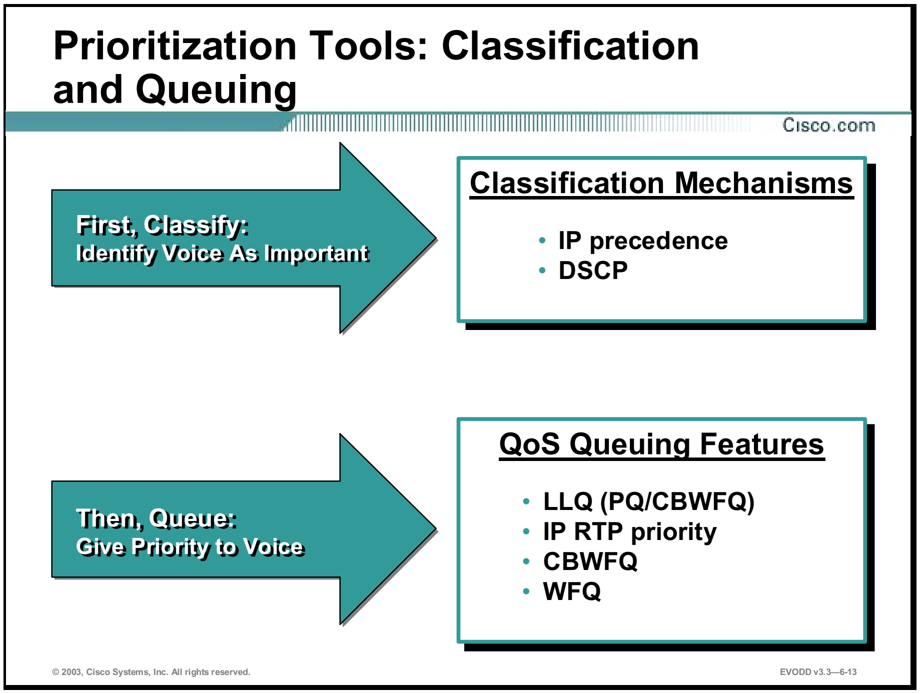
As voice and data leave the router, you have three areas of concern: prioritization, slow-speed links, and traffic shaping.

Prioritization protects voice from data by ensuring that the voice packets are out in front of the data packets. Several tools are available for prioritization.

On slow-speed links, large data frames (such as 1500 bytes) may take an excessive amount of time to be sent out of a serial interface, which delays the transmission of voice frames.

You can address admission control, or the need to prevent oversubscription of voice, through planning and design.

Note	Traffic shaping coexists with the WAN.



The prioritization tools are classification and queuing.

Classification

Before a network can handle traffic according to its unique requirements, the network must identify or label it. Numerous classification techniques exist, including: Layer 3 (IP) schemes, such as IP precedence or the use of the differentiated services code point (DSCP); Layer 2 MAC schemes, such as 802.1P; and the use of an implicit characteristic of the data itself, such as a defined port range.

Cisco uses all of the above, with the Layer 3 schemes providing the unique ability to classify from end-to-end. By default, Cisco IP Phones use an IP precedence level of 5 to provide *critical* priority to voice traffic. This marking evolves to Expedited Forwarding (EF) when using DSCP.

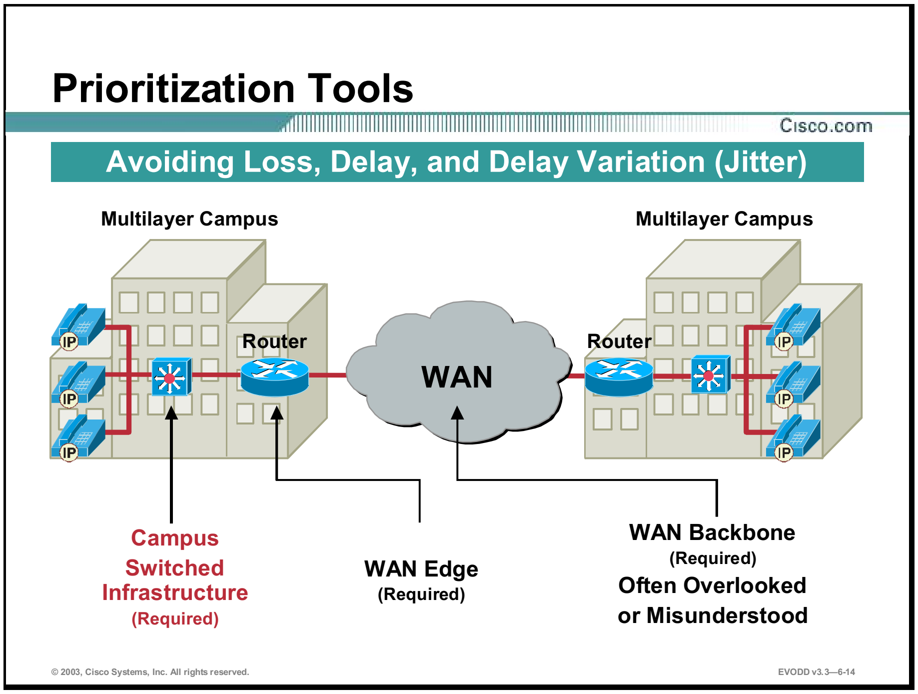
Queuing

Queuing schemes are also numerous and the chosen scheme will vary, depending on the traffic offered to the network and the wide-area media being used. For multiservice traffic provisioned over an IP WAN, the recommendation for low-speed links is to use LLQ, a combination of priority queuing for voice, and class-based weighted fair queuing (CBWFQ) for the data traffic.

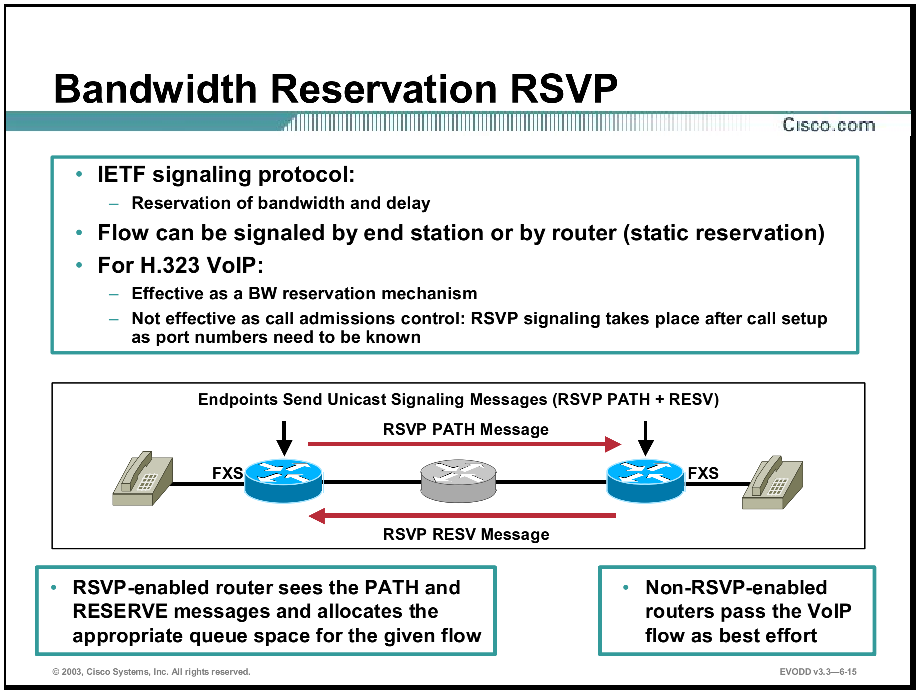
Voice is a predictable and well-behaved type of traffic. The network places voice in a priority queue to be transmitted with low delay and jitter. The standard audio range of User Datagram Protocol (UDP) ports 16384 through 32767 identify voice as Real-Time Transport Protocol (RTP) traffic. Data is further subdivided using CBWFQ to offer differing levels of service to other classes of traffic, such as Systems Network Architecture (SNA) traffic.

Prioritization Tools

This topic details the features of available QoS prioritization tools. Prioritization tools protect voice traffic from data.



Many problems within a campus can be fixed by adding bandwidth. If you use Gb Ethernet links and 10-Gb Ethernet links, you are probably not overloading the infrastructure in the campus. However, a challenge for bandwidth exists in the WAN.

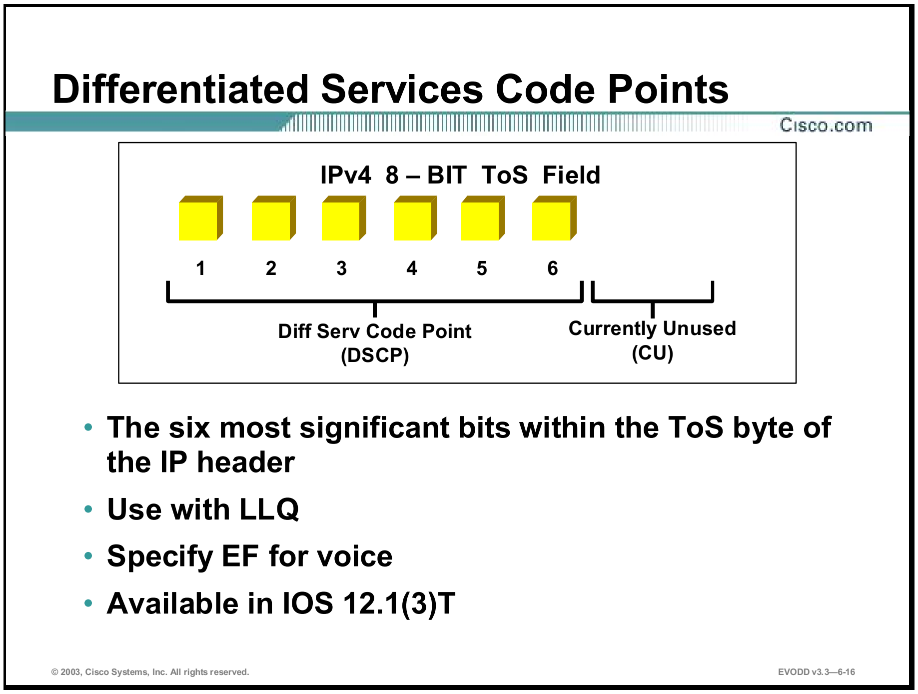


The Resource Reservation Protocol (RSVP) provides dynamic bandwidth and latency guarantees to delay-sensitive IP traffic. RSVP allows client applications to request certain bandwidth and service parameters for their data streams.

RSVP allows the network to identify that a particular workstation is requesting a specific level of bandwidth from each network element on the path.

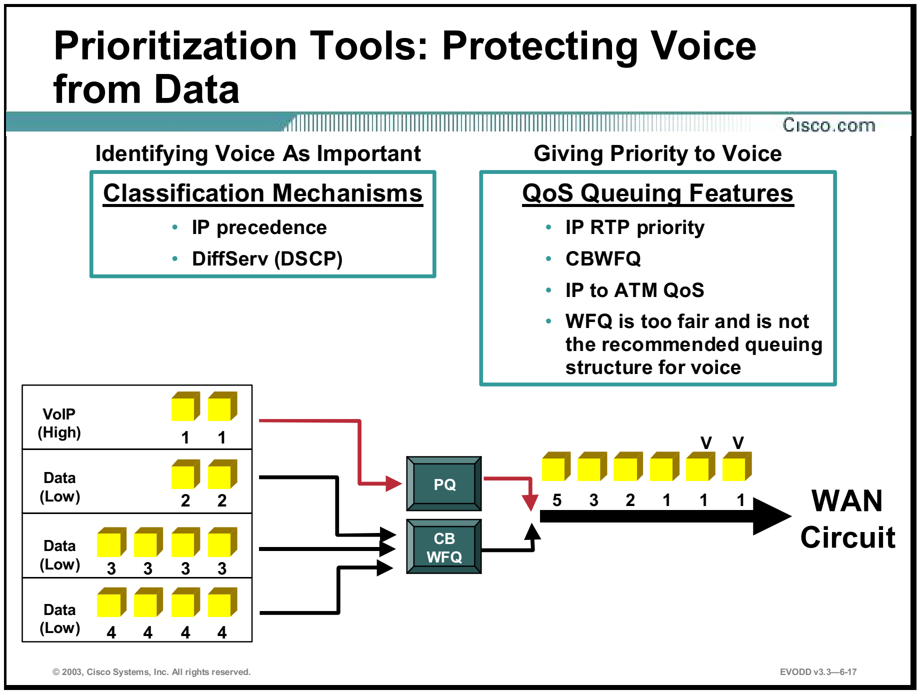
These network elements reserve the requested bandwidth, using priority and queuing mechanisms, if the bandwidth is available. When the server receives the approval, bandwidth is reserved across the entire path.

RSVP works at Layer 3 and is transparent to the Layer 2 switching path. The network can make reservations on behalf of the client. In the figure shown here, the Cisco router with telephones directly attached to Foreign Exchange Station (FXS) ports can act as an RSVP client and initiate the reservation request in VoIP networks.



Cisco IOS Release 12.1 (3) T is the first Cisco IOS software that classifies packets based on their DSCP/IP-precedence values within the type of service (ToS) field in the IP header. You can set a new DSCP/IP precedence value within each class. DSCP allows more classifications than IP precedence. Specifically, DSCP supports 64 values, and IP precedence supports 8 values. Even though DSCP supports 64 possible values, Cisco recommends that you use specific predefined values. These values are called per-hop behaviors (PHBs), because they influence how each router will treat the packet at each hop along the path from the source to the destination. An example of a PHB is EF, which is the PHB that would be applied to latency-sensitive traffic such as voice.

This packet classification feature enables the classification and modification of packets based on the DSCP/IP precedence values in the IP header. The feature allows the queuing of packets based on incoming DSCP/IP precedence values and the coloring of them to different values when moving to the downstream network node.



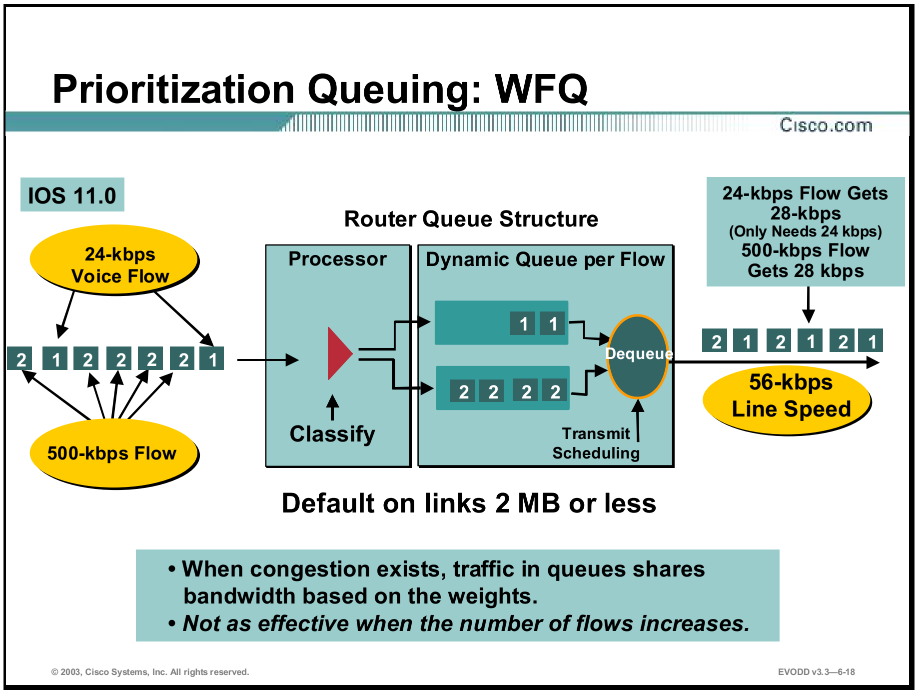
Queuing mechanisms have evolved in this order: weighted fair queuing (WFQ); Class-based WFQ (CBWFQ); priority queuing-weighted fair queuing (PQ-WFQ, now IP RTP Priority); and the latest tool, PQ/CBWFQ (known as Low Latency Queuing or LLQ). For better performance, you should classify traffic as close to its source as possible in a way that ensures end-to-end classification. This classification requires a Layer 3 classification scheme (ToS) since Layer 2 (class of service [CoS]) only functions within a single network.

For prioritization, Cisco IOS Release 12.0 (5) T introduced the command **ip rtp priority**, which is primarily for leased lines. In Cisco IOS Release 12.0 (7) T, the support for this command expanded to Frame Relay. This command allows a priority queue (PQ) for voice traffic. The entrance criterion to the PQ is the use of UDP in a defined port range. Initially this technique served its purpose, but as Cisco IOS evolved, its limitations became evident.

In the evolution of Cisco IOS, IP RTP priority was the voice and data network queuing mechanism of choice as of 12.0(5)T on leased lines. Based on the classification IP precedence, WFQ assigns a particular weight to a flow, gives it preferential treatment, and allows it on the wire ahead of other traffic types. WFQ is not the best solution because it is weighted and fair. This mechanism is comparable to all of the passengers on an airline flight having first-class tickets, even though the first-class cabin only seats half of the passengers on the plane.

IP RTP priority provided a PQ, an exhausted queue that always puts voice traffic at the front of the queue. You can monitor the amount of bandwidth allowed into that queue. The major issue with IP RTP priority is that the classification scheme uses a UDP port range, which is not a very granular way to specify an entrance criterion to a queue.

The figure shown here illustrates a router serving a PQ (voice) before other data queues, which have their own priority mechanism.



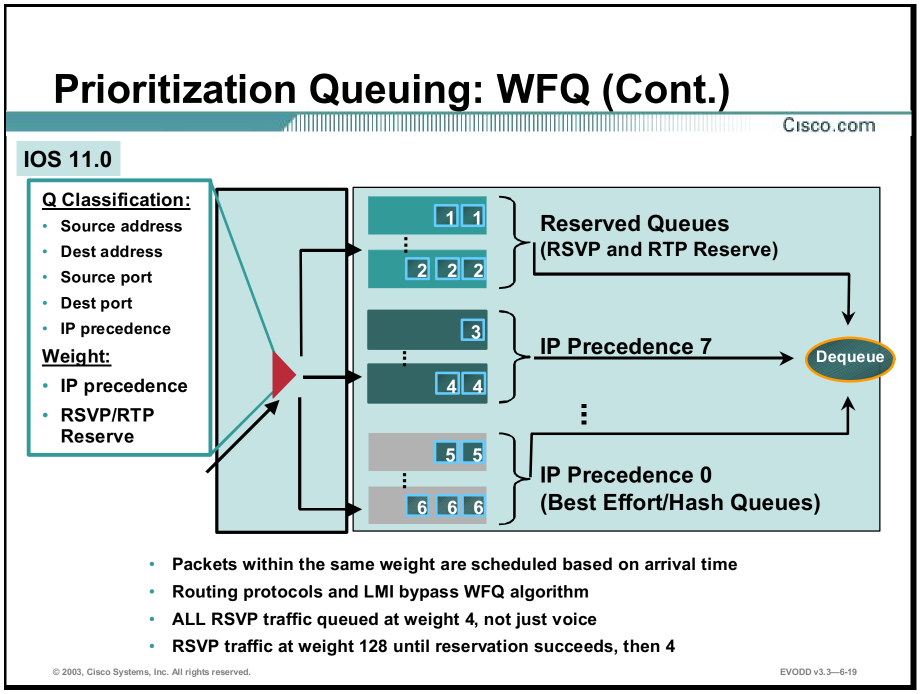
WFQ gives high priority to sessions that are delay-sensitive (such as voice traffic), while ensuring that other applications get fair treatment.

For example, voice traffic jumps right to the front of the queue in the Cisco network, which also services SNA and HTTP traffic. This prioritization works very well, because some applications do not require significant bandwidth as long as they meet their delay requirements. However, some voice applications have strict delay requirements.

A sophisticated Cisco IOS algorithm looks at the size and frequency of packets to determine whether a specific session has heavy or light traffic flow. Then, the algorithm treats the respective queues of each session accordingly. WFQ has the added benefits of being self-configuring and dynamic. Interfaces running E1 speeds or below, turn on WFQ by default.

Flow-based WFQ applies weights to traffic to classify it into conversations and determine how much bandwidth each conversation is allowed relative to other conversations. These weights and traffic classifications are dependent on and limited to the eight IP precedence levels.

CBWFQ specifies the exact amount of bandwidth allocated for a specific class of traffic. Taking into account the available bandwidth on the interface, CBWFQ can configure up to 64 classes and can control distribution among them. This is not the case with flow-based WFQ, because WFQ is not effective enough for assisting voice traffic.



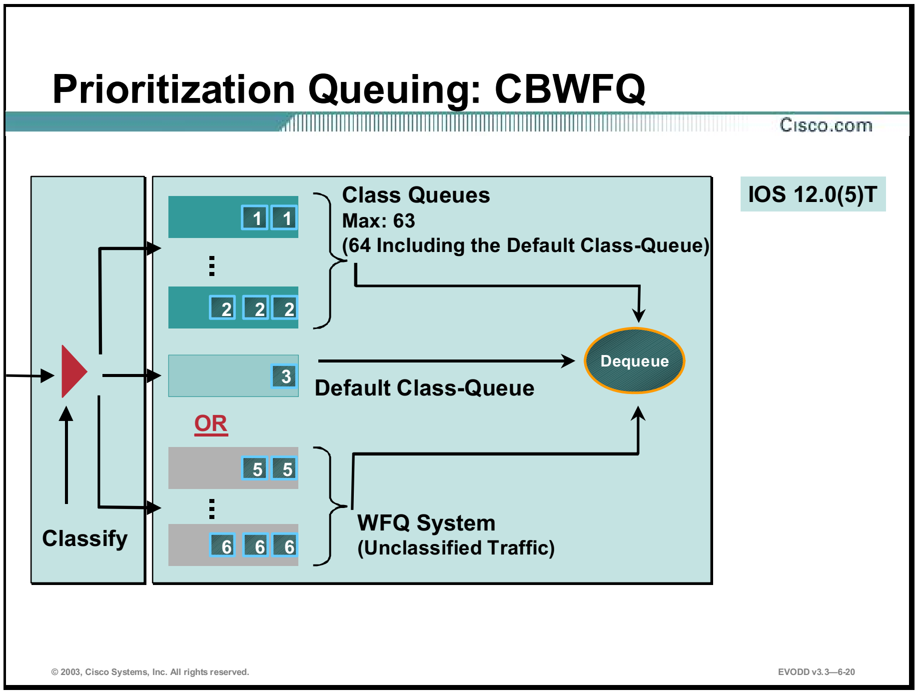
By examining WFQ processes, you can identify the problems associated with using it for voice traffic.

Source and destination address, source and destination ports, and IP precedence setting classify packets first. Either the IP precedence setting or the RSVP/RTP Reserve flag weight these packets.

WFQ first serves the queues in the order of reserved queues and then by IP precedence field. As the precedence value increases, the algorithm allocates more bandwidth to that conversation to make sure that it takes precedence when congestion occurs. WFQ assigns a weight to each flow, which determines the transmit order for queued packets. In this scheme, WFQ serves lower weights first. IP precedence serves as a divisor to this weighting factor. For example, WFQ assigns traffic with an IP precedence field value of 7 a lower weight than traffic with an IP precedence field value of 3; thus, the traffic with the IP precedence field value of 7 has priority in the transmit order.

A problem arises due to WFQ serving all flows. When the number of queues becomes large, even the higher priority traffic begins to experience unacceptable latency. CBWFQ solves this problem by fixing bandwidth for voice queues and exhaustively queuing them ahead of remaining queues.

Cisco recommends using CBWFQ instead of WFQ in your design for data traffic. For VoIP, use LLQ if available; otherwise, you can use IP RTP priority.



CBWFQ extends the standard WFQ functionality to provide support for user-defined traffic classes. Define traffic classes based on match criteria, including protocols, access control lists (ACLs), 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 CBWFQ directs traffic belonging to a class to the queue of that class.

Once you define a class according to its match criteria, you can assign characteristics. To characterize a class, assign bandwidth, weight, and a maximum packet limit. The bandwidth assigned to a class is the minimum bandwidth delivered to the class during congestion.

You must also specify the queue limit for a class, which is the maximum number of packets allowed to accumulate in the queue of that class. Packets belonging to a class are subject to the bandwidth and queue limits that characterize the class.

After a queue has reached its configured queue limit, adding packets to the class causes tail drop or packet drop, depending on how you configure class policy.

CBWFQ classes use tail drop unless you explicitly configure a policy for a class to use Weighted Random Early Detection (WRED) to drop packets as a means of avoiding congestion. If CBWFQ classes use WRED packet drop instead of tail drop for one or more classes comprising a policy map, you should not configure WRED for the interface to which you attach that service policy.

If you configure a default class, CBWFQ treats all unclassified traffic as traffic belonging to the default class. If you do not configure any default class, then by default the traffic that does not match any of the configured classes is flow-classified and given best-effort treatment. Once CBWFQ classifies a packet, all of the standard mechanisms that differentiate service among the classes apply.

Flow classification is standard WFQ treatment. WFQ classifies packets with the same source IP address, destination IP address, source Transmission Control Protocol (TCP) or UDP port, or destination TCP or UDP port as belonging to the same flow. WFQ allocates an equal share of bandwidth to each flow. Flow-based WFQ is also called fair queuing, because all flows are equally weighted.

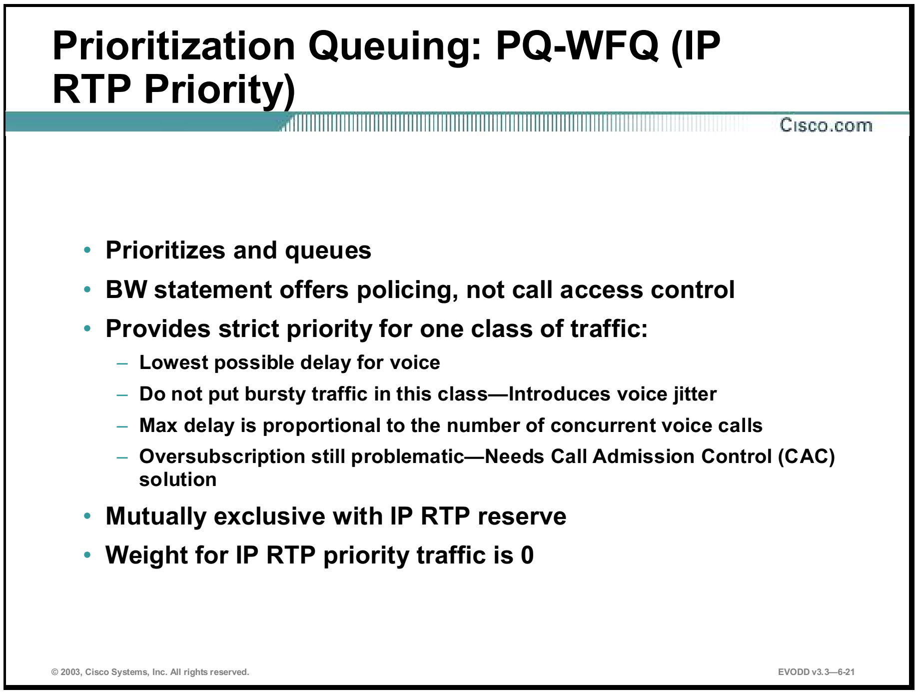
For CBWFQ, which extends the standard WFQ, the weight specified for the class becomes the weight of each packet that meets the match criteria of the class. CBWFQ classifies packets that arrive at the output interface according to the match criteria filters you define, and then it assigns each one the appropriate weight. CBWFQ derives the weight for a packet belonging to a specific class from the bandwidth assigned to the class. You can configure the weight for a class.

After CBWFQ assigns a weight to a packet, it sends the packet to the appropriate class queue. CBWFQ uses the weights assigned to the queued packets to ensure that the class queue is serviced fairly.

To configure a class policy or CBWFQ:

- Define traffic classes to specify the classification policy (class maps). This process determines how packets differentiate from one another.
- Associate policies, or class characteristics, with each traffic class (policy maps). This process applies policy configuration to packets belonging to one of the classes previously defined through a class map. For this process, configure a policy map that specifies the policy for each traffic class.
- Attach policies to interfaces (service policies). This process requires that you associate an existing policy map, or service policy, with an interface so that the particular set of map policies apply to that interface.

Reference More information on CBWFQ can be found at: http://www.cisco.com/univercd/cc/td/doc/product/software/ios120/120newft/120t/120t5/cbwfq.htm.



The following are support limitations for PQ-WFQ (IP RTP priority):

- Subinterfaces
- 75xx-VIP, 12xxx, and 85xx platforms
- Bandwidth-on-demand ISDN

However, the 2600, 3600, 7200, and 7500 Route/Switch Processor (RSP) platforms do support IP RTP priority.

Cisco IOS coverage is as follows:

- Cisco IOS Release 12.0 (5) T or later for Voice over Frame Relay (VoFR)
- Cisco IOS Release 12.0 (5) T or later for VoIP/Multilink PPP (MLP)
- Cisco IOS Release 12.0 (6) T or later for VoIP/Frame Relay

IP RTP priority provides a strict PQ scheme for delay-sensitive data such as voice. RTP port numbers can identify voice traffic, and the **ip rtp priority** command configures the PQ in which the voice is classified. Therefore, IP RTP priority services voice as strict priority in preference to other non-voice traffic.

This feature extends and improves on the functionality offered by the IP RTP Reserve feature. IP RTP priority allows a specified range of UDP/RTP ports and guarantees their voice traffic strict priority service over any other queues or classes using the same output interface, up to a configured bandwidth. Strict priority means that if packets exist in the priority queue and the PQ bandwidth has not been exceeded, PQ-WFQ sends them before packets in other queues.

Cisco recommends that the **ip rtp priority** command be used instead of the **ip rtp reserve** command for voice configurations.

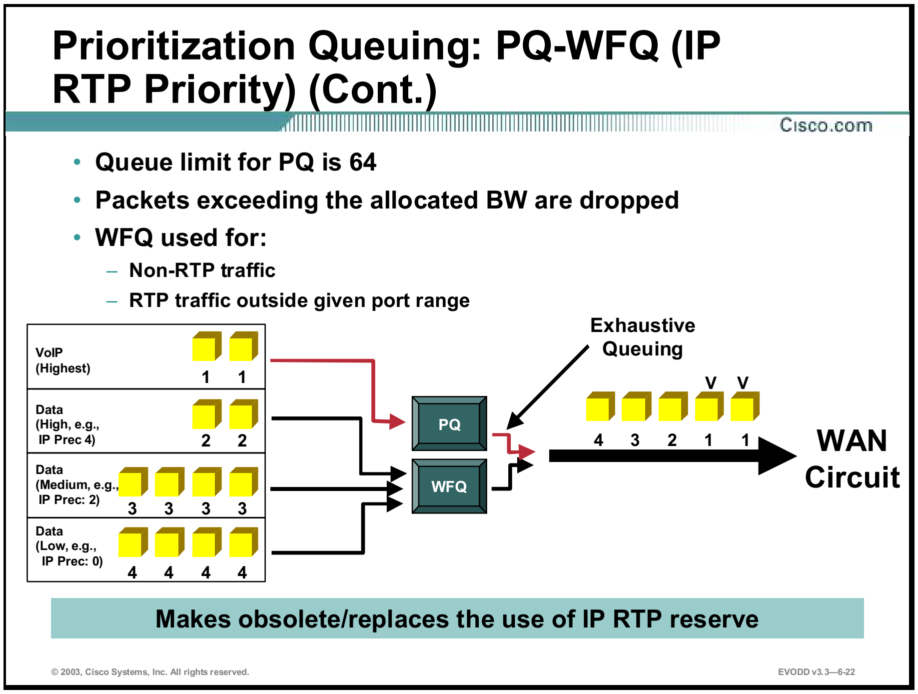
The IP RTP priority feature does not require that the port of a voice call be known. This feature allows the ability to identify a range of ports whose traffic is put into the priority queue. The entire voice port range, 16384 to 32767, can be specified to ensure that all voice traffic is given strict priority service. IP RTP priority is especially useful on slow-speed links of less than 1.544 Mbps.

You can use the IP RTP priority feature in conjunction with either WFQ or CBWFQ on the same outgoing interface. IP RTP priority guarantees strict priority over other CBWFQ classes or WFQ flows for traffic matching the range of ports specified for the priority queue. IP RTP priority always services voice packets in the priority queue first unless the configured bandwidth is exceeded. The voice traffic is then policed.

When used in conjunction with WFQ, the **ip rtp priority** command provides strict priority to voice. WFQ scheduling is applied to the remaining queues.

When used in conjunction with CBWFQ, the **ip rtp priority** command provides strict priority to voice. You can use CBWFQ to set up classes for other types of traffic (such as SNA) that need dedicated bandwidth and need to be treated better than best-effort and not as strict priority. CBWFQ services the non-voice traffic fairly based on the weights assigned to the packets in the queue. CBWFQ can also support flow-based WFQ within the default CBWFQ class, if you configure it.

You should also configure the Link Fragmentation and Interleaving (LFI) feature on lower-speed interfaces, because voice packets are small in size and the interface may have large packets going out. When you enable the LFI, it breaks up the large data packets, so that the small voice packets can be interleaved between the data fragments that make up a large data packet. With LFI, a voice packet will not have to wait while a large packet is sent; it can send the voice packet in a shorter amount of time.



IP RTP priority identifies voice on the top queue and dequeues it in a PQ-like fashion, giving voice strict priority over data. WFQ, depending on the number of flows, treats each flow equally based on precedence. IP RTP priority gives voice utmost priority no matter how many flows exist.

To understand the behavior of the IP RTP priority feature and properly use it, you must consider its admission control and policing characteristics. When using the **ip rtp priority** command to configure the PQ for voice, specify a strict bandwidth limitation. IP RTP priority guarantees this amount of bandwidth to voice traffic enqueued in the PQ. This is the case whether you use the IP RTP priority feature with CBWFQ or WFQ.

The IP RTP priority closely polices use of bandwidth for the PQ every second, ensuring that the allocated amount is not exceeded in the event of congestion. IP RTP priority prohibits transmission of additional packets once the allocated bandwidth is consumed. If IP RTP priority discovers that the configured amount of bandwidth is exceeded, it drops packets—an event that voice traffic tolerates poorly. (Enable debugging to watch for this condition.) Close policing allows for fair treatment of other data packets enqueued in other CBWFQ or WFQ queues. To avoid packet drop, allocate to the priority queue the most optimum amount of bandwidth, taking into consideration the type of coder-decoder (codec) used and interface characteristics. IP RTP priority will not allow traffic beyond the allocated amount.

For safety, allocate to the priority queue slightly more than the known required amount of bandwidth. For example, suppose that you allocate 24 kbps of bandwidth (the default amount for VoIP) to the priority queue. This allocation seems safe because transmission of voice packets occurs at a constant bit rate.

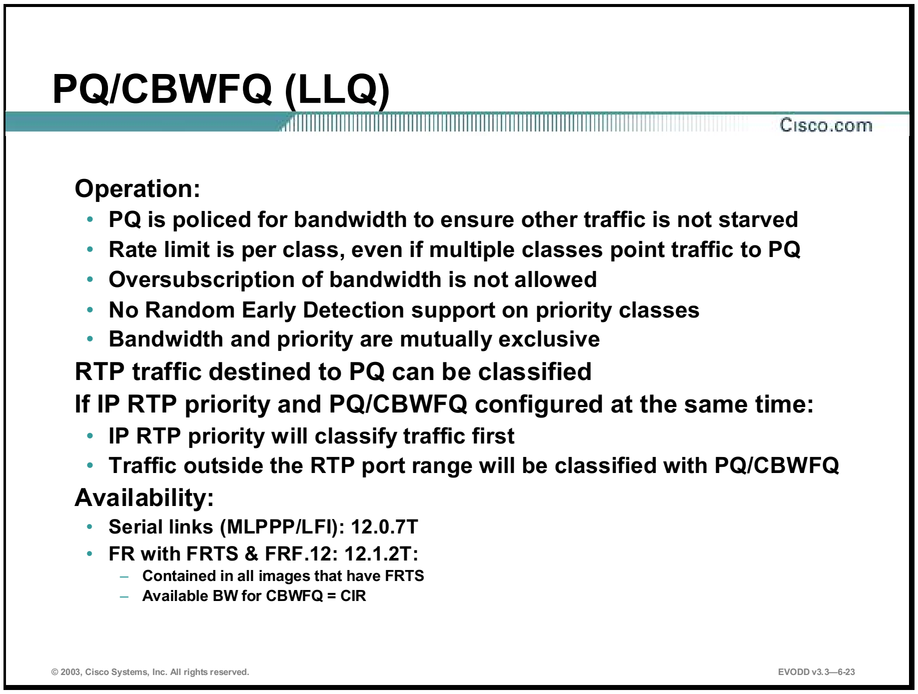
Because the network and the router, or switch, can use some of the bandwidth and introduce jitter and delay, you should allocate slightly more than the required amount of bandwidth (such as 25 kbps) to ensure constancy and availability.

The IP RTP priority admission control policy does take RTP header compression into account. When you configure the bandwidth parameter of the **ip rtp priority** command, only configure the bandwidth of the compressed call. For example, if a G.729 voice call requires 24 kbps uncompressed bandwidth but only 12 kbps compressed bandwidth, you only need to configure a bandwidth of 12 kbps.

The bandwidth management feature of IP RTP priority stipulates that the sum of all bandwidth allocation for voice and data flows on the interface is not to exceed 75 percent of the total available bandwidth. It uses the remaining 25 percent of bandwidth for other overhead, including the Layer 2 header and routing traffic.

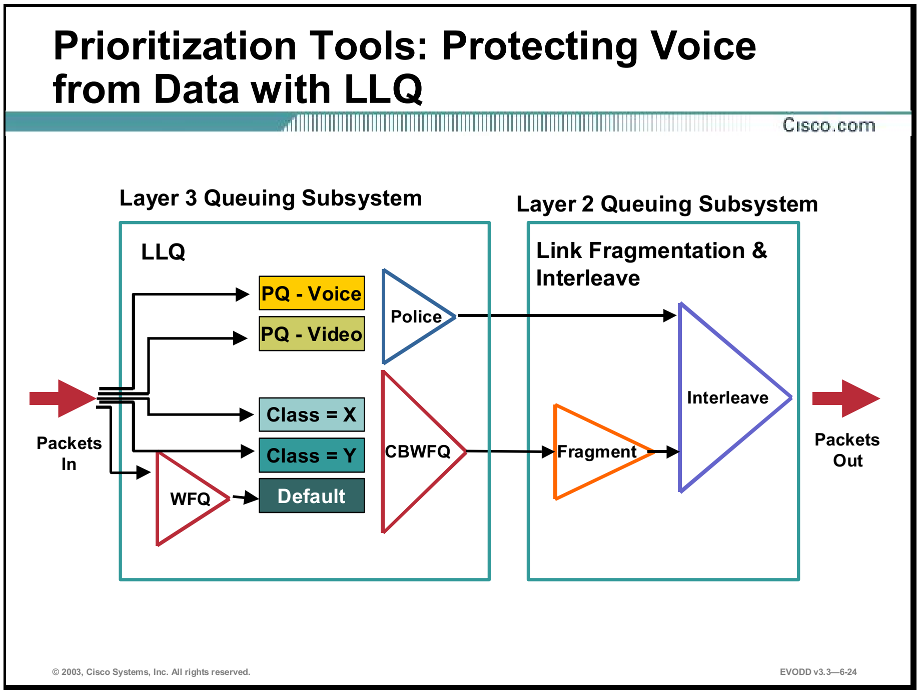
Typically, 75 percent of available bandwidth to all classes and flows is apportioned. This apportioned bandwidth includes the CBWFQ voice class assigned to the PQ or the WFQ voice flow using the PQ. Bandwidth allocation for voice packets accounts for the payload plus the IP, RTP, and UDP headers, but not the Layer 2 header. The Layer 2 header and best-effort traffic, among other overhead, use the remaining 25 percent of bandwidth. The CBWFQ default class, for instance, is takes bandwidth from the remaining 25 percent. Allowing 25 percent bandwidth for other overhead is conservative and safe. On a PPP link, for example, overhead for Layer 2 headers assumes 4 kbps.

If you know the amount of bandwidth required for additional overhead on a link, you can override the 75 percent maximum allocation for the bandwidth sum allocated to all classes or flows. This change is useful for aggressive circumstances, such as when you want to give voice traffic as much bandwidth as possible. The IP RTP priority feature includes a command called **max-reserved-bandwidth** that you can use to override the 75 percent limitation. If you override the fixed amount of bandwidth, exercise caution and ensure enough remaining bandwidth to support best effort and control traffic.



Cisco IOS Release 12.0 (7) T introduced a new queuing scheme called LLQ for leased lines and for ATM, and Cisco IOS Release 12.1 (2) T introduced LLQ for Frame Relay. LLQ is a slight modification to CBWFQ. You can still configure classes the same way, but with LLQ, one or more of those classes can have PQ.

The entrance criteria to these queues is much more granular than with IP RTP priority, because they do not limit priority packets to a UDP port range. If you have classified all the traffic at Layer 3 using the DSCP, or IP precedence at the periphery, you may now use those markings to place packets into classes using class maps.



The Layer 3 subsystem defines multiple classes (up to 64), and one or more have PQ behavior. You can specify a PQ for voice and the entrance criteria for voice-related traffic. You could even specify a second PQ for interactive video. You can define additional queues based on the classification schemes to provide CBWFQ treatment.

You might define classes for specific application traffic. For example, you could define Class X for SNA and Class Y for Oracle. You could place non-classified traffic in the default queue. The **class-default** class is limited to WFQ behavior.

Traffic classified in a PQ interleaves directly from the output. PQ traffic cannot be fragmented. There are ramifications to this behavior, specifically with video. Voice is in small fragments by default. If you specify a video queue as well, the minimum link over which you can send voice, video, and data may be fairly large.

The traffic for the other classes travels through the CBWFQ subsystem and can be fragmented before it leaves the line.

Cisco recommends using this tool on the WAN. If you configure IP RTP priority as well as CBWFQ, IP RTP priority takes precedence. The use of IP RTP priority should decline over time.

PQ/CBWFQ Issues

Cisco.com

- **Priority classes (video) cannot be fragmented for FRF.12**
- **Priority classes (video) should not be fragmented for MLPPP**
- **Not possible to interleave fragments from different classes**

  Hence, minimum bandwidth recommended for voice, video and data = 768K
- **Leased lines:**
  - **Voice + data** — **LFI is possible**
  - **Voice + video + data** — **768K min bandwidth**
- **Frame Relay:**
  - **Voice + data** — **LFI is possible**
  - **Voice + video + data** — **768K min bandwidth**
- **ATM:**
  - **No LFI technique available** — **(2 PVCs)**
  - **Voice + video + data** — **768K min bandwidth**
- **ATM/Frame Relay:**
  - **No LFI technique available**
  - **Voice + video + data** — **768K min bandwidth**

EVODD v3.3—6-25

The design recommendations related to the use of LLQ are shown here.

You cannot interleave fragments from multiple classes. The PQs bypass the fragmenting logic. Therefore, do not put a 1500-byte frame in a PQ on a slow speed link (such as less than 768 kbps), because the frame will not be fragmented and will delay other PQ traffic.

To run voice, video, and data over a single link, the minimum recommended link bandwidth is 768 kbps to accommodate the video without causing excessive serialization delay for the voice traffic.

LLQ Benefits

Cisco.com

- Consistent configuration and operation across all media types:
  - ATM
  - Frame Relay
  - Leased lines
- Entrance criteria to a class can be defined by an ACL:
  - Not limited to UDP ports as it is with IP RTP priority
  - Use of IP RTP priority should be phased out
  - Ensure trust boundary is defined to assure simple classification and entry to a queue

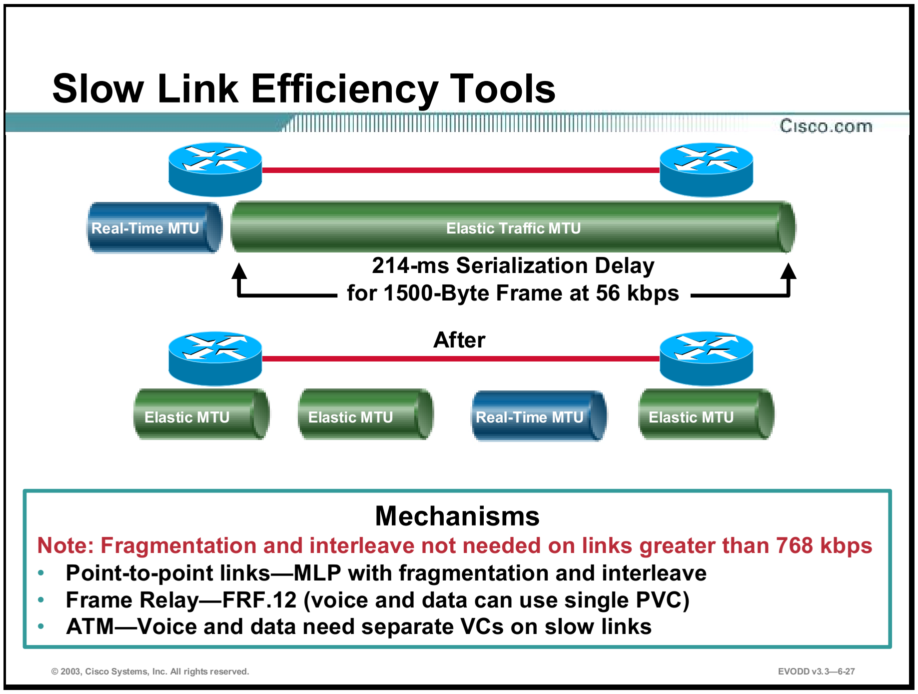
EVODD v3.3—6-26

One benefit of LLQ is having a consistent configuration across all media types, regardless of what media you use. Previously, you had to complete configurations for MLP in one case and for Frame Relay Forum implementation agreement for Voice over IP over Frame Relay (VoIPoFR) (FRF.12) in another. LLQ provides a consistent queuing mechanism.

With LLQ, the entrance criteria to a class can be as granular as needed. LLQ does not have simple UDP port limits as IP RTP priority does. If Cisco had not changed the port range feature, future applications may have taken advantage of it, knowing that when they hit the Cisco infrastructure they would get preferential treatment.

Slow Link Efficiency Tools

This topic focuses on methods of maximizing the efficiency of voice transfer on slower speed links. In a WAN, QoS features such as queuing, slow link efficiency tools, and traffic shaping work together to provide a seamless, efficient, wide-area transport of data, voice, and video.

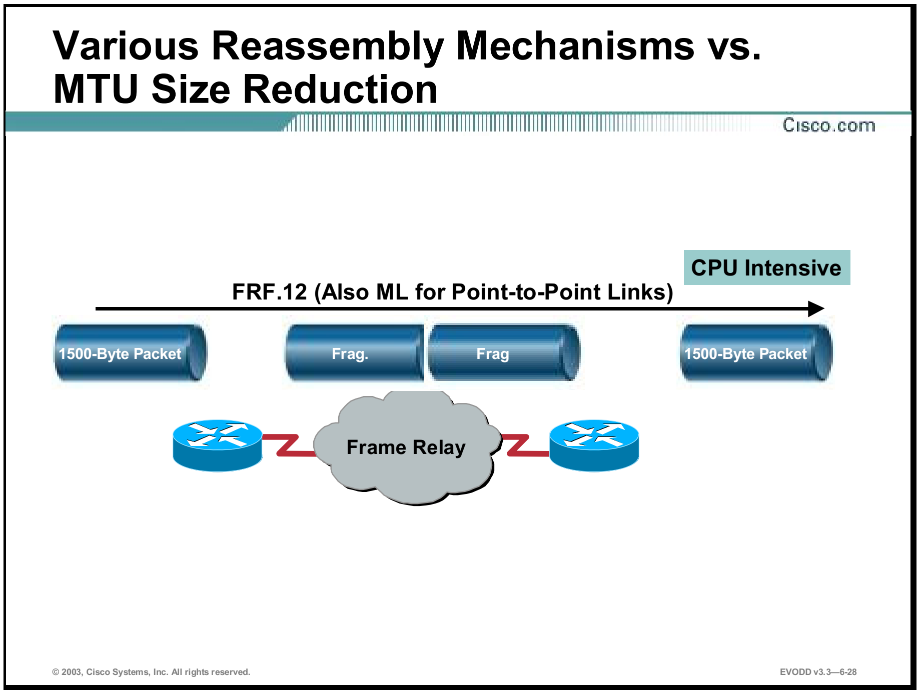


On slow links, when a large packet has to clock out on a serial link, any voice packets behind it must wait until the larger packet is fully transmitted on the line. For example, on a 56 kbps circuit, a 1500-byte packet takes 214 ms to fully clock out. The design goal is a 150-ms delay from mouth to ear, one-way end-to-end, so 214 ms far exceeds the recommended delay budget.

Fragmentation and interleaving tools can fix this problem. These tools fragment the large packet and interleave the voice packets among the pieces so that the voice packets do not suffer significant delay.

Following are some of the more popular LFI tools used for various media types:

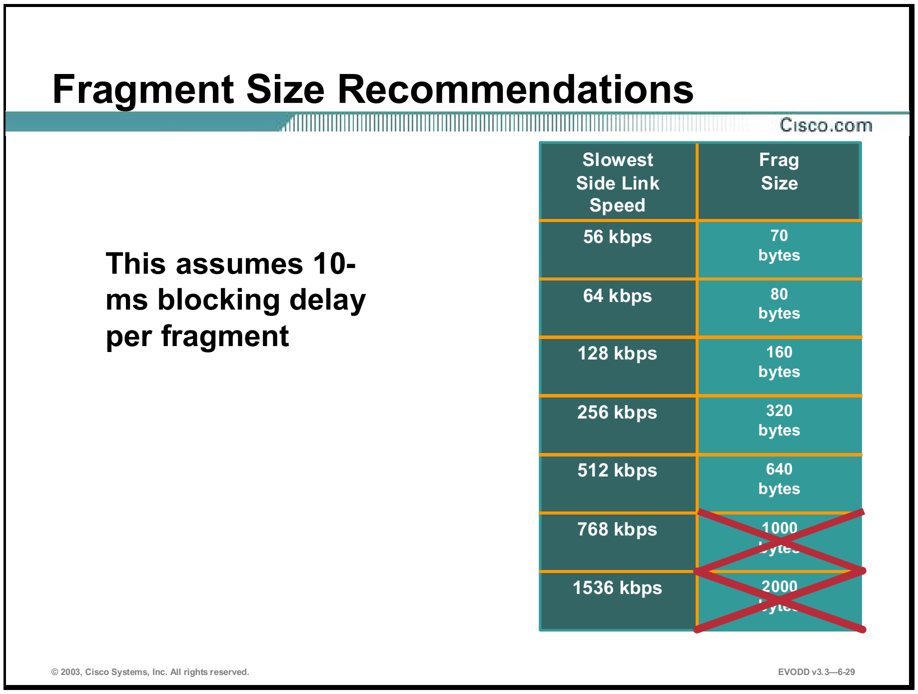
- For point-to-point interfaces, you can use MLP with fragmentation and interleaving. MLP is supported in Cisco IOS Release 11.3.
- For VoIPoFR, FRF.12 performs LFI functions. For VoFR, Frame Relay Forum implementation agreement for VoFR (FRF.11) performs LFI functions.
- You can use two virtual circuits as a solution for ATM to ATM. ATM to Frame Relay has a number of issues associated with it.



For VoIPoFR, available as of Cisco IOS 12.0 (4) T, FRF.12 performs LFI functions. FRF.12 is the functional equivalent of MLP in that it fragments large frames at Layer 2, interleaves the voice packets, and reassembles the larger frames on the other end of the circuit.

ATM, inherently using its 53-byte cell, also has the ability to interleave. In many cases, ATM has an interface speed higher than 768 kbps and may not require interleaving. However, you may need some sort of link efficiency mechanisms if low-speed permanent virtual circuits (PVCs) exist. If you take a 1500-byte frame and turn it into cells during the ATM segmentation and reassembly (SAR) process, the cells must remain in sequence throughout the life of that transmission. In some cases, you may need two PVCs.

Note	For many of these designs, Cisco recommends that you work with your local systems engineer in environments dealing with LFI for ATM.

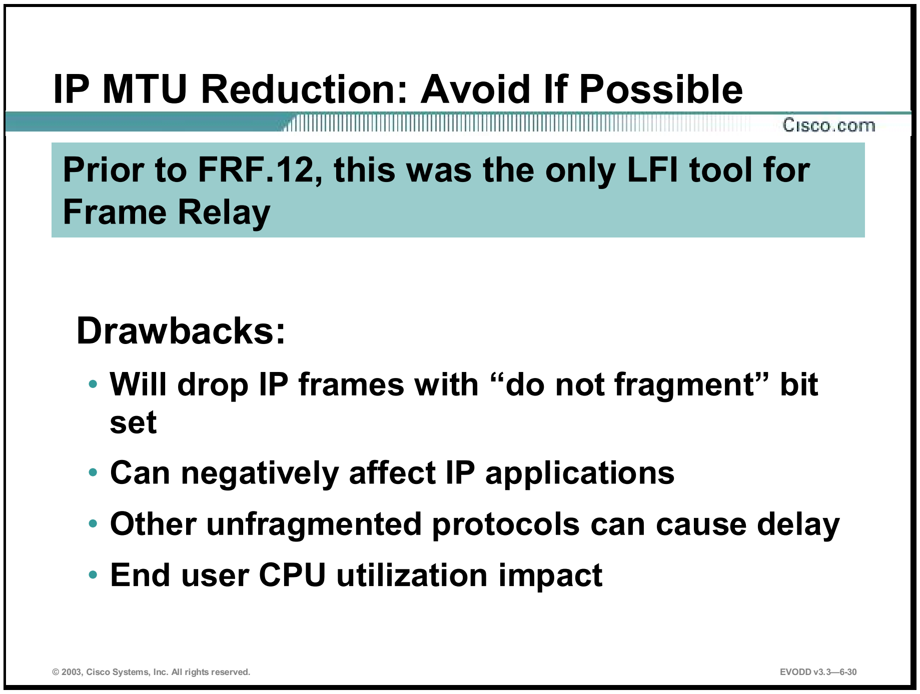


To make a fragment for both VoFR and VoIP, some Cisco design guides suggest setting a fragment size for data to ensure that if voice gets caught behind a given fragment, the voice packet is delayed no more than 10 ms.

The figure shows the link speeds from 56 kbps through T1 and the associated fragment size that will result in no more than a 10-ms blocking delay.

In an environment with a Frame Relay network and a T1 at a central site that feeds many remote sites, you should base fragmentation at the central site on the speeds at the remote sites instead of speeds at the central site. This consideration is commonly overlooked, because at the far end of that circuit, a real 128-kbps circuit may exist, and a large packet there would result in adverse quality.

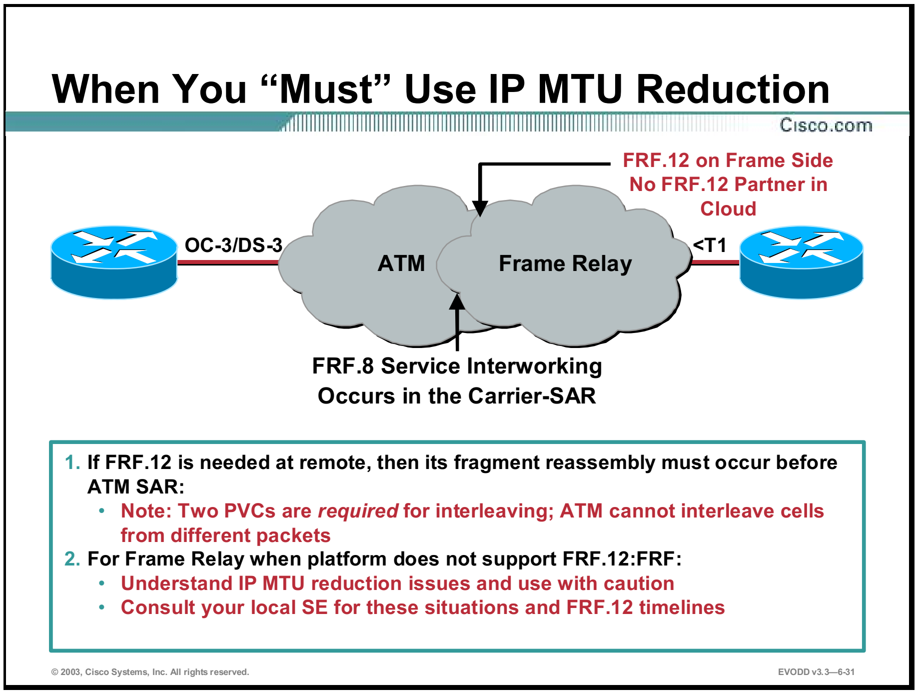
Note For Frame Relay PVCs, you should set the fragment size according to the speed of the PVC. For example, set the fragment size on a 128-kbps PVC on a T1 to 160-byte.



Avoid using IP maximum transmission unit (MTU) size reduction. Prior to the Cisco IOS 12.0 (4) T code train, this was the only way of providing fragmentation and interleaving on Frame Relay interfaces.

Using IP MTU size reduction can negatively affect quality, because if an IP packet has its “do not fragment bit” set, it will be dropped.

When packets are fragmented in this manner, they remain as fragments until they reach the ultimate endpoint server. The endpoint must reassemble the fragments. The processing overhead associated with reassembly can have an adverse effect on performance and throughput.

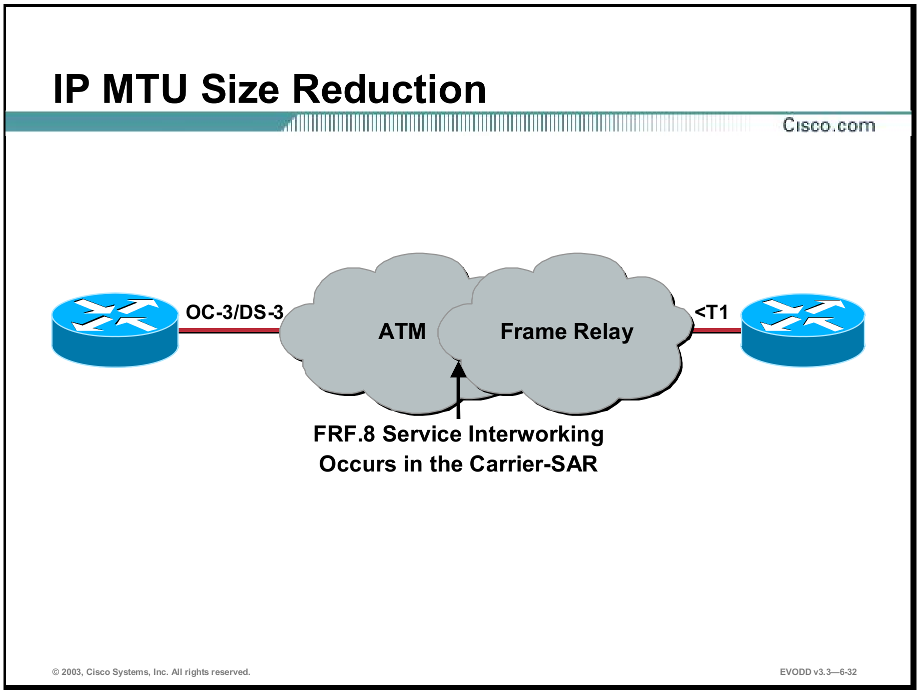


There are certain cases when you must use an IP MTU size reduction.

When converting Frame Relay to ATM, carriers may not reassemble frames in the cloud before they are converted to ATM cells. An option in this environment is to perform IP MTU size reduction to keep the frames from being fragmented.

The other situation is the state of FRF.12 support. Since there are platforms that do not support FRF.12, you must know which products are involved and the image releases used. Currently, FRF.12 is supported on the following platforms: 805, 1600, 1700, 2500, 2600, 3600, 4500, 4700, MC3810, 7200, and 7500 series routers. When possible, use FRF.12. If an existing platform does not support FRF.12, use IP MTU size reduction. When specifying platforms in a design, specify platforms that support FRF.12 where needed.

Note If you need to run IP MTU size reduction, Cisco recommends that you use the availability of your local systems engineer, who can set the proper design guidelines to ensure a successful deployment.



Remember the following warnings if your design includes IP MTU size reductions:

- IP MTU size reductions will drop packets with the “do not fragment bit” set
- IP MTU size reductions are very inefficient on CPU resources—end points must reassemble fragments

Note G.729 VoIP consumes 42.4 kbps when transported over ATM at 50 packets per second (pps). G.711 VoIP consumes 106 kbps when transported over ATM at 50 pps.

Capacity Planning Considerations

Cisco.com

Payload bandwidth requirements for various codecs

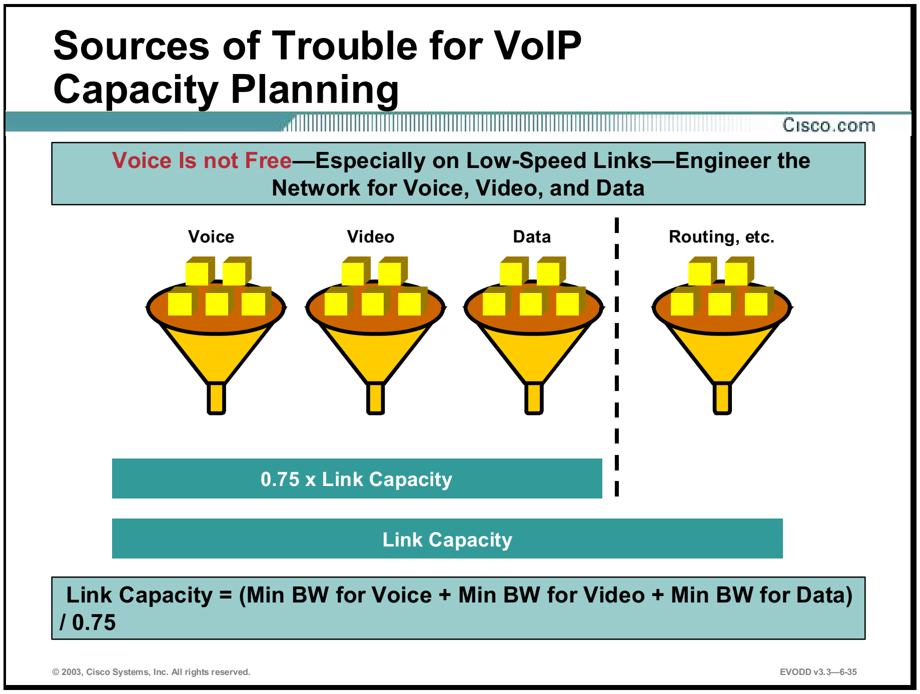
Encoding/Compression	Resulting Bit Rate
G.711 PCM A-Law/mu-Law	64 kbps (DS-0)
G.729 CS-ACELP	8 kbps
G.723.1 CELP	6.3/5.3 kbps

EVODD v3.3—6-34

Deploying IP Phones requires the same kind of capacity planning as the tie-line replacement, although the paradigm is slightly different. If you are using G.711, you need a greater amount of bandwidth. However, you can lower equipment costs, and other problems (such as voice mail integration) are easier to resolve. Remember to watch your overhead.

The following are common acronyms involved with capacity planning:

- Pulse code modulation (PCM)
- Adaptive differential PCM (ADPCM)
- Code excited linear prediction compression (CELP)
- Algebraic code excited linear prediction (ACELP)
- Conjugate Structure Algebraic Code Excited Linear Prediction (CS-ACELP)
- Low-delay CELP (LDCELP)



Before you place voice and video on a network, ensure that adequate bandwidth exists for all required applications. To do this, add the minimum bandwidth requirements for each major application (for example, voice, video, and data). This sum then represents the minimum bandwidth requirement for any given link and should only constitute approximately 75 percent of the total bandwidth for the link. This 75-percent rule assumes that some bandwidth is required for overhead traffic such as routing and Layer 2 keepalive messages, as well as for additional applications such as e-mail and HTTP traffic.

Include the IP Overhead

Cisco.com

Each VoIP packet contains 40 bytes of IP overhead

Per-call BW consumption depends on:

- **IP header size**
- **Voice payload size**
- **Packets per second**

Not Including Link Layer Header or RTP Header Compression

Cisco router at G.711	**=**	**160-byte voice payload at 50 pps (80 kbps)**
Cisco router at G.729	**=**	**20-byte payload at 50 pps (24 kbps)**
Cisco IP phone at G.711	**=**	**240-byte payload at 33 pps (74.6 kbps)**
Cisco IP phone at G.723.1	**=**	**24-byte payload at 33 pps (17 kbps)**

Link layer header sizes vary per media.

EVODD v3.3—6-36

You must consider some additional planning to run compressed voice across the WAN.

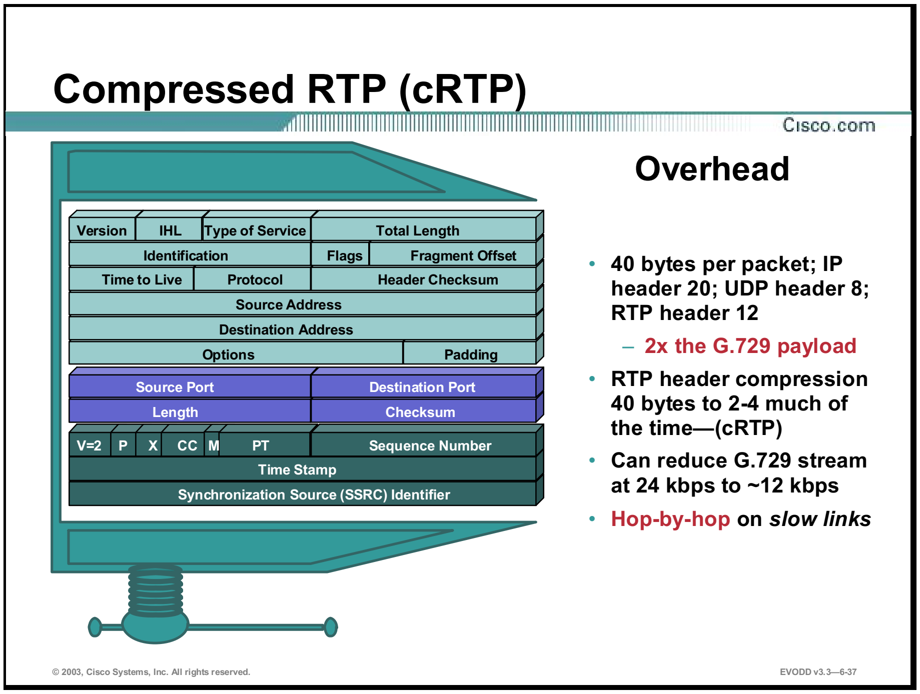
Compressed voice, such as G.729, only consumes 8 kbps from a voice payload perspective, so it is an 8:1 savings from the traditional standard 64 kbps PCM transport model.

Every packet will have 40 bytes of IP overhead, so the actual 8 kbps does not reflect the effective end bandwidth consumed on a WAN.

The payload header and the pps at which the router or IP Phone transmits will change depending on the packet size. For example, a Cisco router running G.711 with a 160-byte payload at 50 pps consumes 80 kbps.

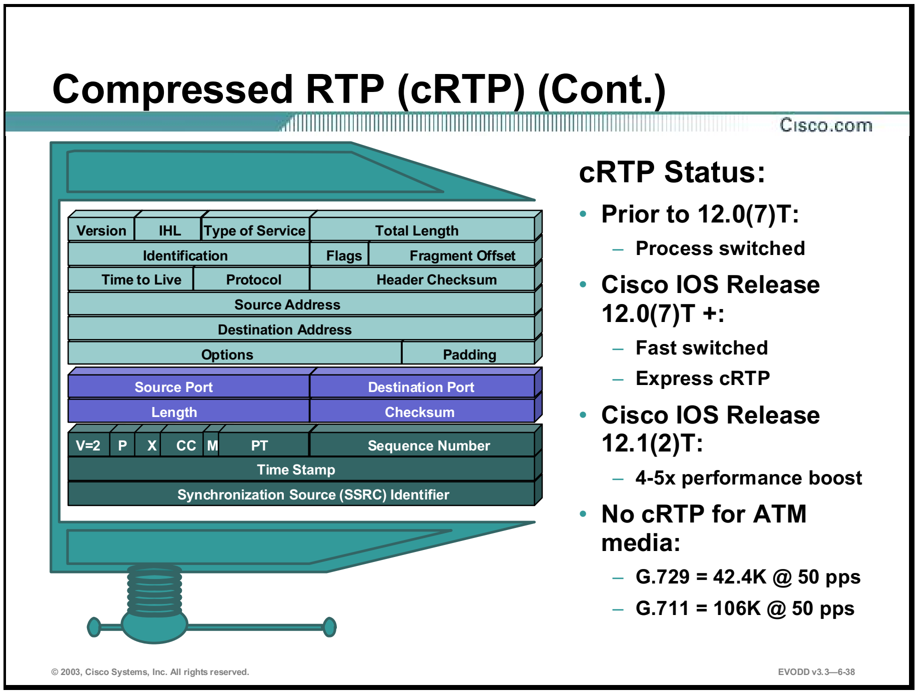
The G.729 codec with 8 kbps of voice payload consumes 24 kbps of WAN bandwidth. This amount varies per media, such as Frame Relay or MLP. Header sizes, even Ethernet, will boost the overall consumed bandwidth.

Cisco recommends that you run your design at 50 pps, or about 20 ms per packet, to ensure good quality voice. Many service providers are going to a 10-second sample range, so the bandwidth for handling 100 pps is not an issue for them. For those needing more bandwidth, adjusting the payload size is a viable solution. The default of 50 pps is a good starting point.



RTP header compression is a tool that can take a 40-byte header and reduce it to 2-byte or 4-byte most of the time. RTP header compression is also called cRTP. It performs on a link-by-link basis and will reduce a G.729 voice stream from 24 kbps down to approximately 12 kbps.

While this approach is called a compression technique, the header is not actually compressed. Rather, a copy of an RTP header is cached in the routers at each end of a link. Then, the new (compressed) header carries an integer called the session context identifier (CID) to identify which flow the packet belongs to. A UDP *checksum* may also be sent in the compressed header.



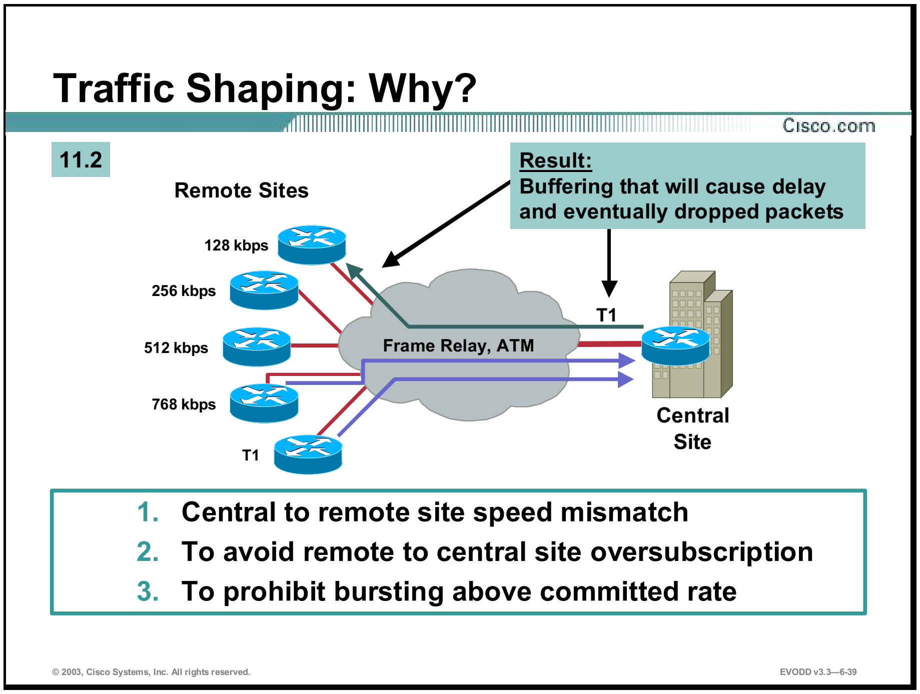
Cisco IOS Release 12.0 (7) T moved cRTP from the process-switching path into the fast-switching path within the Cisco IOS software, nearly doubling throughput. An algorithm implemented in Cisco IOS Release 12.1 (2) T provides a 400-500 percent improvement in the number of cRTP packets sent in a period of time.

ATM media does not currently support cRTP, but that may change in future Cisco IOS software releases. Use these numbers for planning on ATM:

- For G.729, 42.4 kbps at 50 pps
- For G.711, 106 kbps at 50 pps

Traffic-Shaping Tools

This topic examines the components of traffic shaping. Traffic shaping is the third tool in the QoS arena and is probably the most misunderstood.

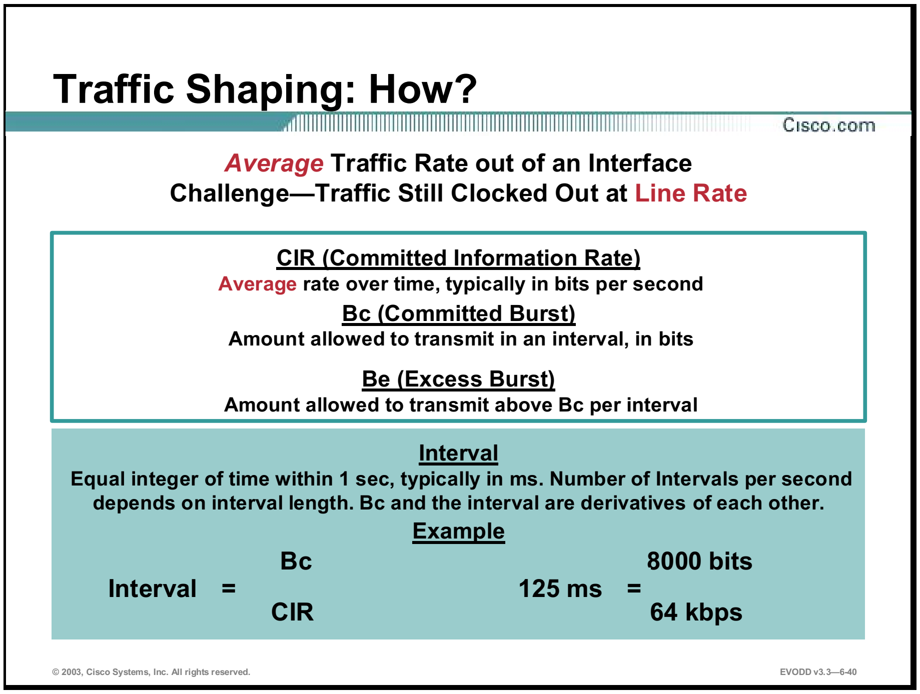


Traffic shaping applies on an interface basis, and works with a variety of Layer 2 technologies (for example, Frame Relay, ATM, Switched Multimegabit Data Service [SMDS], and Ethernet). Use an access list to select the traffic. You can set up traffic shaping to adapt dynamically to available bandwidth on Frame Relay subinterfaces using backward explicit congestion notification (BECN) integration. Typically, you use traffic shaping for three reasons:

- When transmitting from a central site router that may contain a T1, such a router will transmit at line rate if it can. In the figure shown here, the central site will transmit towards the 128-kbps remote site at a T1 rate. Voice and data are on the same PVC traveling at a T1 rate towards a 128-kbps line rate. You will find a bottleneck in either the Frame Relay cloud or ATM cloud, and buffering causes delay and dropped packets. Perform traffic shaping from this central site so that traffic heading toward the 128-kbps remote site will leave that central site at no greater than 128 kbps.

- You can ensure that simultaneous bursts from many remote sites will not congest the central site. For data transmission only, a common practice was to have the aggregate burst rate of all the remote sites possibly exceed what the T1 could handle at the central site. You could expect that data would not transmit at the same time. If you did have a delay or a dropped packet, you may not have known about it. Typically, your applications would retransmit. Voice traffic has caused this paradigm to change, because the remote routers can oversubscribe the T1 at any given time. Traffic shaping ensures that remote sites cannot oversubscribe your central site.

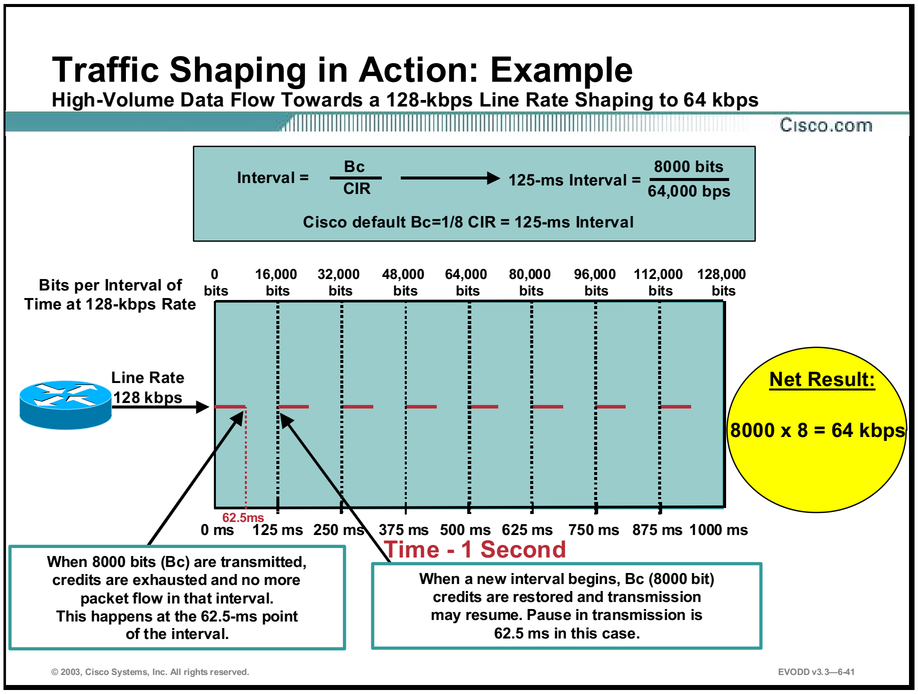
- Traffic shaping prohibits bursting above committed rates, or shaping to the committed information rate (CIR). Customers expect bursting to line rate and then getting guaranteed CIR. In a voice environment, this paradigm changes. You do not want voice up in a best-effort domain because the voice packets may be dropped, which results in choppy conversation.



Traffic shaping over a Frame Relay network is often misunderstood.

The CIR is an average rate over time, typically bits per second (bps). Traffic exits an interface at the clocked line rate, not at the CIR. Typically, the committed burst (Bc) is the number of bits that are sent in a time interval.

Calculate the time interval by dividing the CIR into the Bc. For example, a CIR of 64 kbps with a Bc of 8,000 bits equates to a time interval of 125 ms.



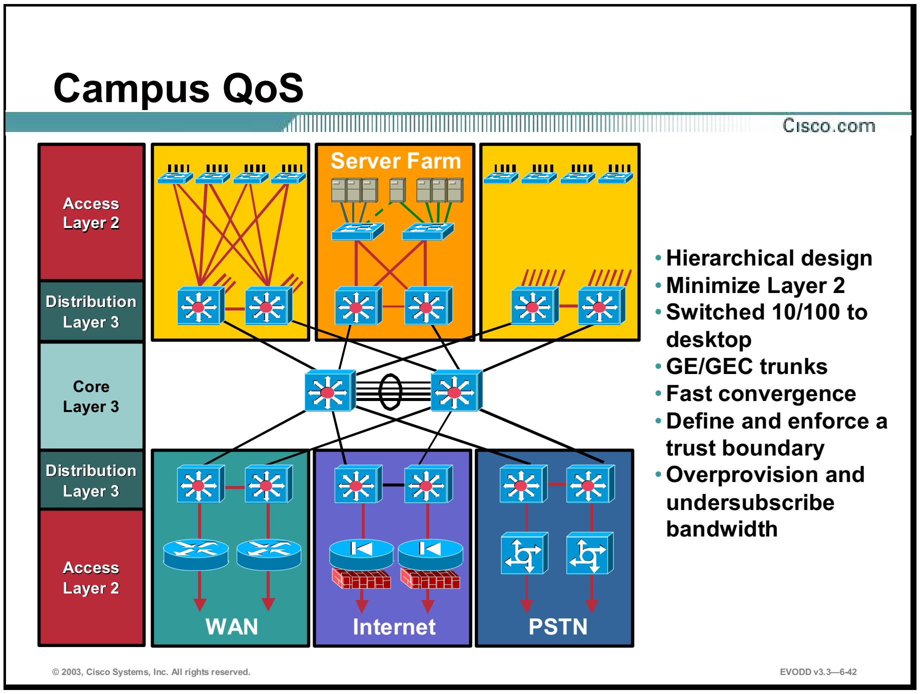
An example of traffic shaping is shown here. A high-volume data flow is sent towards a 128- kbps circuit, shaping to a CIR of 64 kbps. This results in a 125-ms interval, which is the default value for Cisco routers.

This is how the data plays out on the 128-kbps line. When Bc (8000 bits) transmits, it exhausts the credits for that interval, and no more packet flow can occur in that interval. This happens at the 62.5-ms point of the interval. When a new interval begins, it restores Bc credits, and transmission may resume. The pause in transmission is at 62.5 ms in this case.

The 125-ms interval means that there are 8 intervals in a second. One second will transmit the Bc eight times, resulting in a 64-kbps transfer, which represents the CIR.

Prioritization

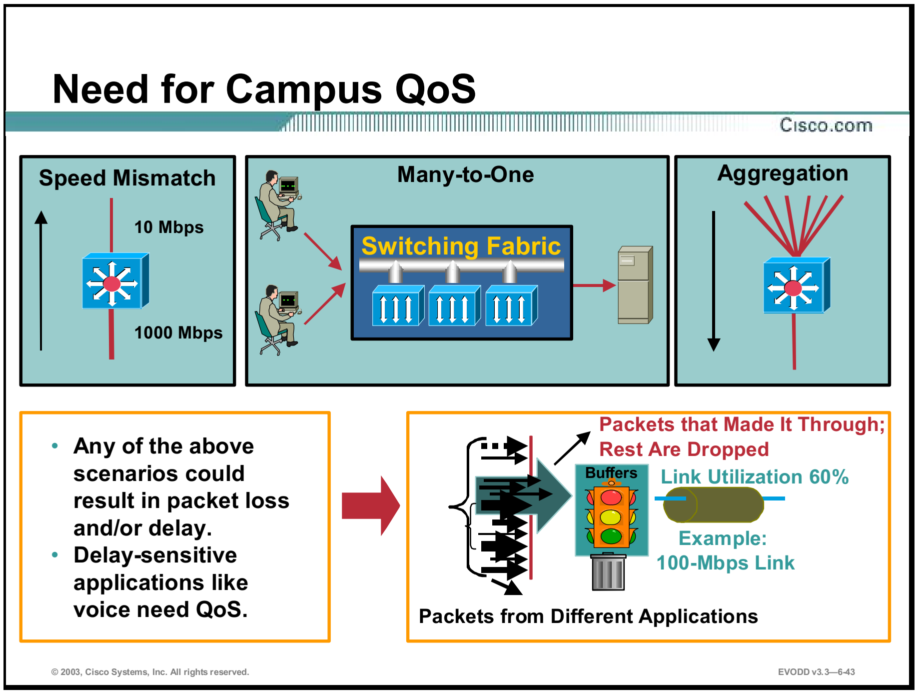
This topic details how you must enable the campus infrastructure with QoS, even though the WAN is typically the focus of QoS tools.



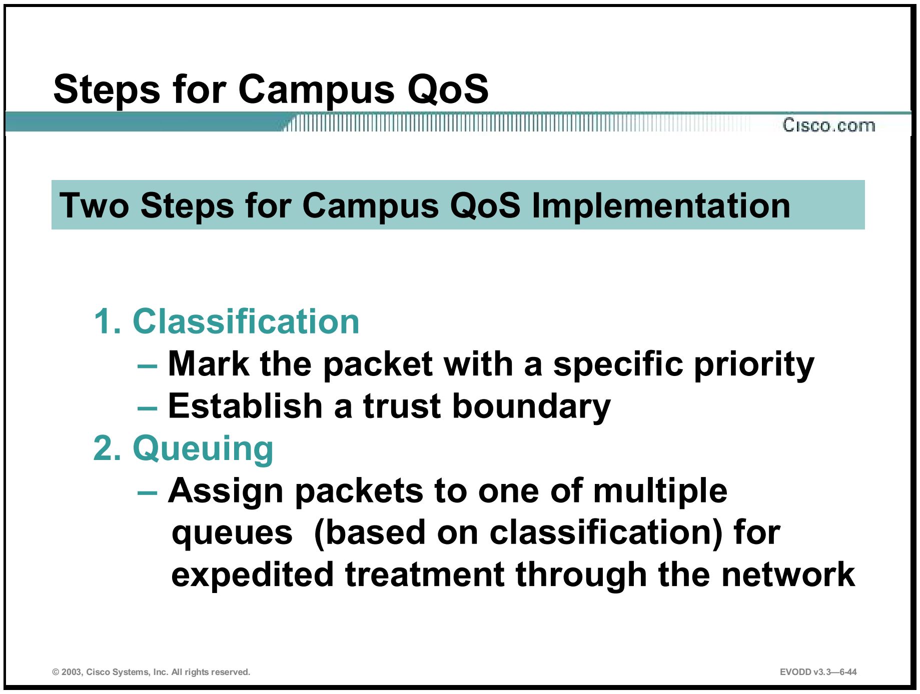
You would typically build campus networks of today on a switched infrastructure in the wiring closet (such as the access layer) and Layer 3 switched at the distribution layer. This approach lends itself well to Voice over Data applications.

Because shared media (such as hubs) does not support IP telephony, 10/100-switched Ethernet access to the desktop is a requirement. Layer 2 switches eliminate the problem of collisions experienced with shared media hubs.

Another important factor is the establishment of a trusted boundary within a network. Because of the port, you usually view the PC, or attached device, as something to trust. With the addition of voice, video, and data, you must define this trust boundary.



One reason that you need QoS in the campus is to accommodate speed mismatches going from Gb Ethernet to Fast Ethernet. Many-to-one means having many things coming down to a common link and oversubscribing it. For example, if you have a Catalyst 6000 with 384 kbps Fast Ethernet ports on it, you can potentially oversubscribe the uplink. These are reasons why you must identify voice and perhaps treat it differently from data.



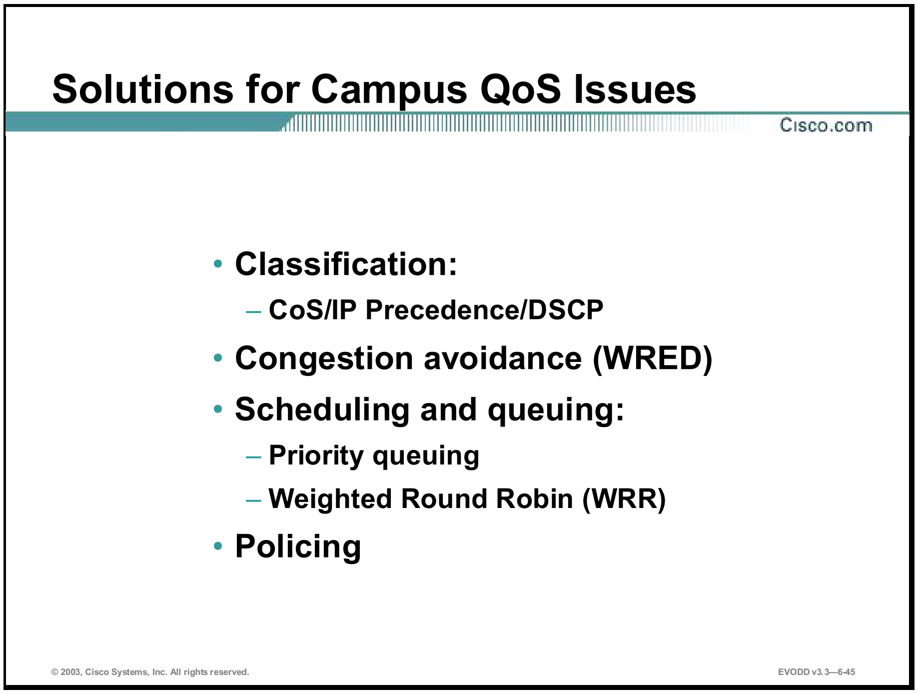
The two steps required for campus QoS are classification and queuing.

Classification

The need to distinguish between voice, video, and data traffic is critical. You must mark traffic and establish a trust boundary. After you have classified the traffic, use these criteria to give it preferential treatment in the WAN. If you classify traffic at the edge of the WAN, it can use this classification mechanism as the entrance criteria to a priority queue for the WAN.

Queuing

The second step is queuing, or assigning packets to one of multiple queues. The campus employs queuing as its classification technique for expedited treatment through the network.



Available classification schemes include: CoS at Layer 2, IP precedence at Layer 3, or DSCP and Layer 3.

The preferred schemes are at Layer 3. However, you cannot always classify at Layer 3. The benefit of classifying at Layer 3 is that classification potentially follows a packet from the source to the destination, regardless of how many hops it crosses. This is true as long as reclassification does not occur.

A congestion avoidance scheme, such as WRED, does not help voice directly. Because voice uses the UDP, dropping a packet is not helpful. However, WRED improves handling TCP traffic by allowing you to hold it back, which helps voice indirectly.

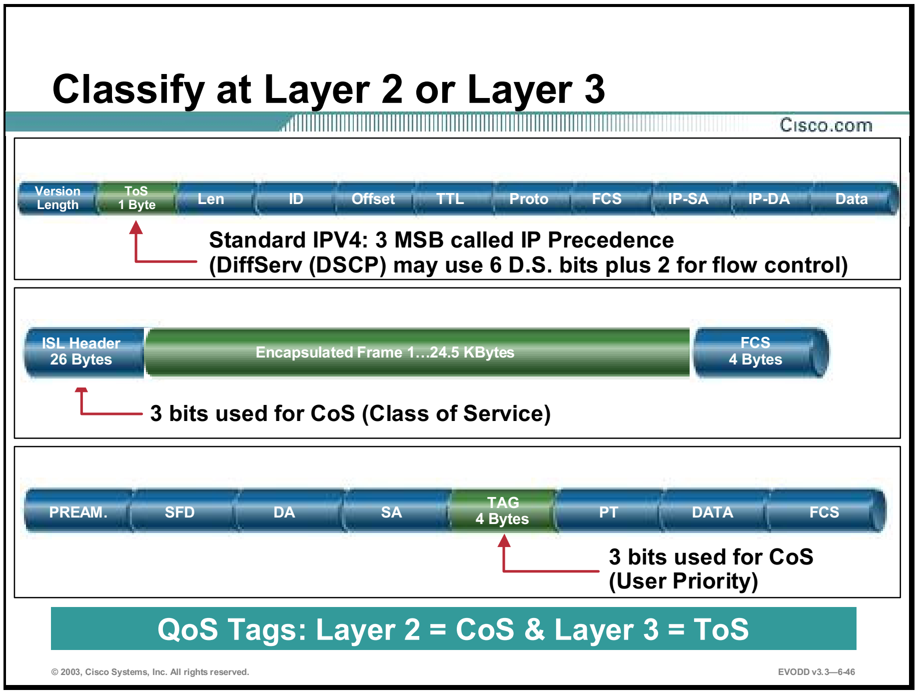
Queuing and scheduling are also issues. The current preferred queuing mechanism is LLQ. LLQ, which is PQ-CBWFQ, is a rigid traffic prioritization scheme. For example, if packet A has a higher priority than packet B, packet A always goes through the interface before packet B.

When you define the QoS property of an interface as PQ, it automatically creates four queues on the interface: high, medium, and low. The interface then defines PQ policies and places packets in these queues. Any unclassified packets go in the normal queue, so that the voice queue can empty before attempting to deliver other classes of traffic. However, this approach can lead to "protocol starvation." So, you can opt to use less strict queuing approaches, such as LLQ or IP RTP priority.

Weighted Round Robin (WRR) scheduling is used on Cisco Catalyst 8500 series switch routers (Layer 3 switches) on egress ports to manage the queuing and sending of packets. WRR places a packet in one of four queues based on IP precedence, from which it derives a delay priority. With WRR, each queue gets a weight. When congestion occurs on the port, the queue uses the weight to give weighted priority to high-priority traffic without depriving low-priority traffic.

The weights provide the queues with an implied bandwidth for the traffic on the queue; the higher the weight, the greater the implied bandwidth. The queues do not have specific bandwidths assigned, and when the port is not congested, it treats all queues equally.

Policing is another QoS tool. QoS Policy Manager (QPM) allows QoS policies to be defined at the enterprise level using a graphical user interface (GUI). For example, QPM can define policies for groups of devices rather than for one device at a time. QPM can also create policies that apply to applications or groups of hosts more easily than using Cisco IOS commands.



You can accomplish the goal of protecting voice traffic from data traffic by classifying voice traffic as high priority and then allowing it to travel in the network before low-priority traffic. You can classify at Layer 2 or at Layer 3 as follows:

- At Layer 2 (CoS) using the 3 bits in the 802.1p field, which is part of the 802.1Q tag or using 3 bits in the Inter-Switch Link (ISL) header

Note	Support for Layer 2 marking comes from trunk links, such as either IEEE 802.1Q or ISL.

- At Layer 3 using either 3 bits (IP precedence) or 6 bits (DSCP) in the ToS byte of the IP header

Classification is the first step toward achieving QoS. The concept of trust is important and integral to deploying QoS. After the end devices have set CoS or ToS values, the switch has the option of trusting them. If the switch trusts the values, it does not need to do any reclassification. If it does not trust the values, it must perform reclassification for the appropriate QoS.

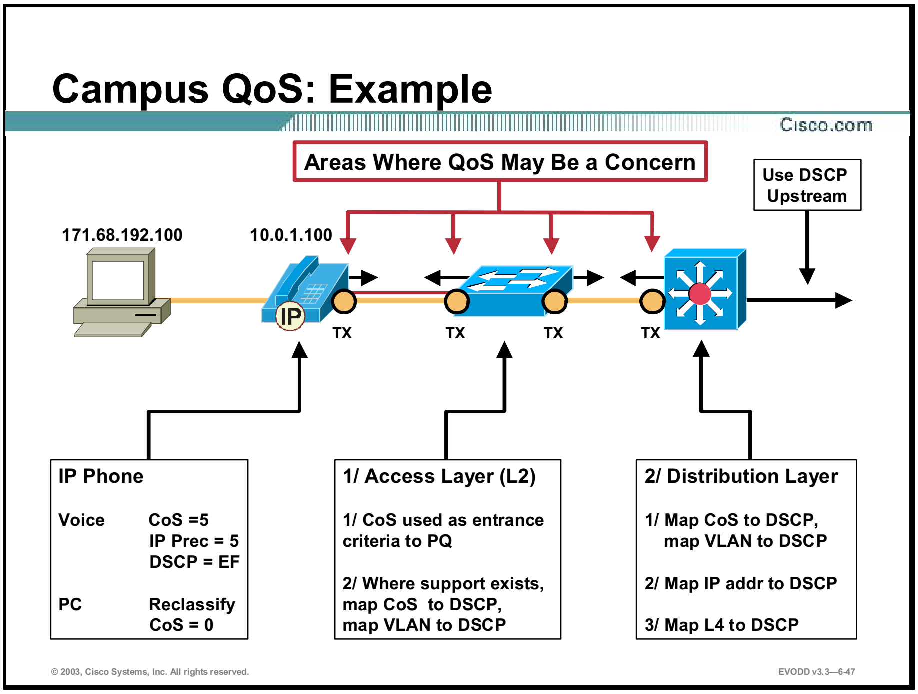
The notion of trusting—or not trusting—forms the basis for the trust boundary. You should classify as close to the source as possible. If the end device can perform this function, then the trust boundary for the network is at the access layer in the wiring closet. If the device cannot perform this function, or the wiring closet switch does not trust the classification done by the end device, the trust boundary may shift. How this shift happens depends on the capabilities of the switch in the wiring closet. If the switch can reclassify the packets, the trust boundary remains in the wiring closet. If the switch cannot perform this function, the task falls to other devices in the network going toward the backbone. If this occurs, you should perform reclassification at the distribution layer. This means that the trust boundary shifted to the distribution layer, which probably has a high-end switch with features to support this function. You should not perform this task in the core of the network if possible.

You should maintain the trust boundary in the wiring closet. If necessary, move it down to the distribution layer on a case-by-case basis, but avoid moving it down to the core of the network. This conforms to the general guidelines to keep the trust boundary as close to the source as possible.

This discussion assumes that you are using a three-tier network model, which is a proven scalable architecture. If the network is small and the logical functions of the distribution layer and core layer are in the same device, the trust boundary can reside in the core layer if it has to move from the wiring closet.

Using the 802.1p bits within the 802.1Q tag provides the desired QoS results at Layer 2. When traffic has to cross a Layer 3 boundary, you must implement these mechanisms using Layer 3 parameters. An example is the three IP precedence bits or the six DSCP bits. Traffic crosses a Layer 3 boundary when Layer 3 switches or routers route packets between subnets. Traffic also crosses a Layer 3 boundary when packets need to leave the campus network and go onto the WAN through edge routers. Layer 2 markings are not preserved in this case. You need Layer 3 marking to achieve the desired level of QoS. All of the QoS techniques employed by the routers (including the very important WAN QoS) rely on Layer 3 classification.

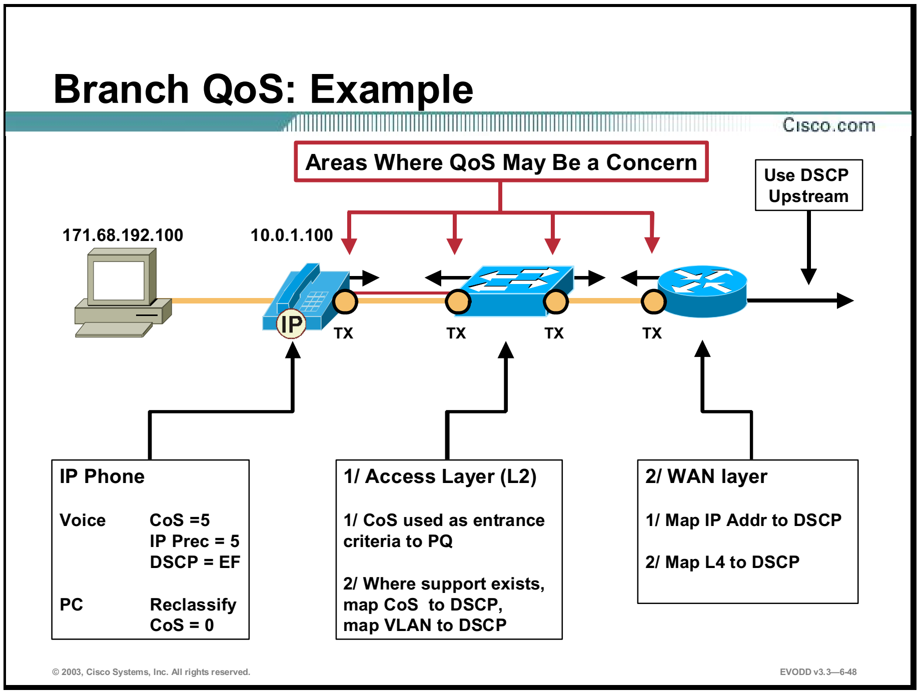
You can achieve Layer 3 classification and marking by using the appropriate platforms in the campus. The IP Phones present packets to the switch with CoS and ToS equaling 5. Packets preserve this Layer 3 classification even if they travel all the way through to the WAN edge router where the Layer 2 header is removed. If the trust boundary is at the source (IP Phone), voice traffic has the IP precedence bits set to 5, and the network devices will give it the appropriate treatment. WAN routers can use this classification to employ many of the queuing techniques. If the trust boundary is not at the source, the device reclassifying the packets should do so at Layer 3 before they can cross a Layer 3 boundary.



The figure shows a typical campus deployment with an access switch (Layer 2 device). In this example, the IP Phone classifies all of the telephony packets. This classification occurs for both the control plane and the bearer plane. For the control plane, the Real-Time Transport Control Protocol (RTCP) traffic goes to the Cisco CallManager (CCM); for the bearer plane, CoS equals 5 and classifies the actual RTP stream as such. IP precedence also equals 5, and you can use it as the Layer 3 classification for the RTP stream itself. If the PC marks IP precedence, the phone cannot do anything, so the packet will simply flow through the device. It can only reclassify CoS at Layer 2.

Assuming that the first switch connected is a Catalyst 3500 series switch, and the PC does not support 802.1Q, the switch can reclassify the traffic in the native VLAN to any CoS value. The 3500 would have two queues on egress: a PQ for CoS 4-7 and a low-priority queue for CoS 0-3.

Prioritizing voice on egress from the switch achieves QoS at Layer 2 in that device. Assuming that the switch is a Catalyst 6500 series switch with a Policy Feature Card (PFC), CoS would match IP precedence or DSCP. Anything that comes in as IP precedence 5, or DSCP EF (which maps to IP precedence 5) remains a 5. Anything that comes in from the PC with a zero maps to DSCP zero. You can use these values as the entrance criteria to a PQ on egress from that switch.



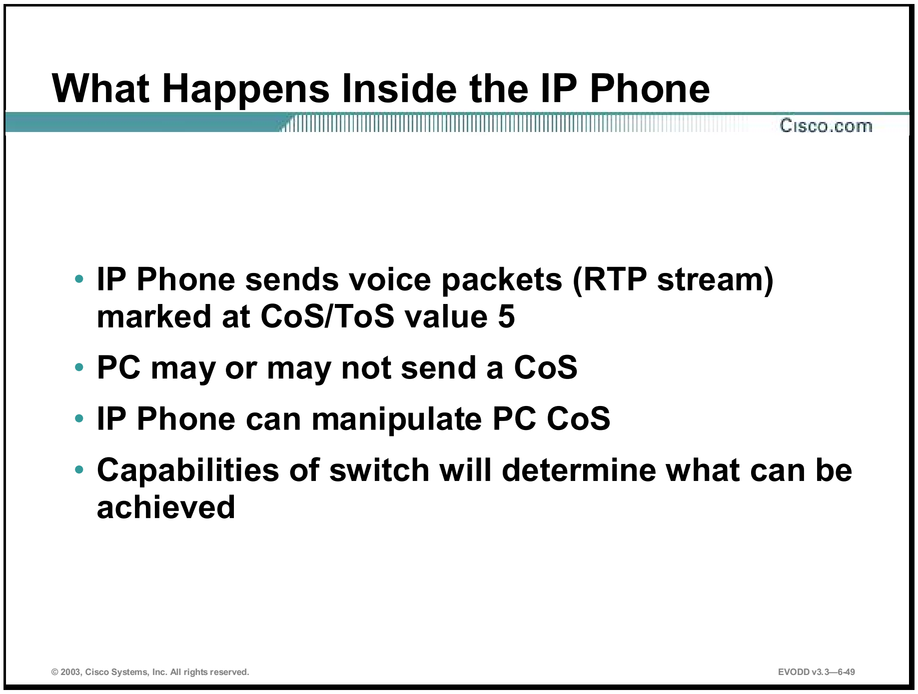
The other end of the conversation may reside in a branch office. That branch office could have any Ethernet switch except a hub. In this case, the switch cannot classify traffic.

If the switch could classify traffic, the process is exactly the same as before. A Catalyst 3500 can reclassify CoS and use it as the entrance criteria to a PQ.

At the WAN edge router, you can use policy-based routing to map an IP address to a DSCP value (assuming the telephones have different addresses than the PC). You can establish the trust boundary at the edge, eliminating the need to classify further into the network. You achieve classification based on the Layer 3 criteria, and the network can use it to decide how to handle a packet.

Inside the IP Phone

This topic addresses how the Cisco IP Phone works with Cisco Catalyst switches to either *trust* or *not trust* the QoS marking of a PC.



Only use the extended CoS value if you set the extended QoS configuration to *untrusted*.

In this case, the mini-switch inside the IP Phone will retag the frame coming from its access port with the specified CoS value.

If you set the extension of the QoS to trusted, the IP Phone will accept the CoS that the PC set and pass it to the switch. If the CoS value of the PC is higher than the CoS of the voice packets, the voice quality and the signaling messages between IP Phone and CCM may not function.

Summary of QoS Capabilities on the Catalyst Switch Family

Cisco.com

Platform	Ability to Trust	Classify CoS	Classify ToS	Avoid Conges-tion (WRED)	Priority Queue	Multiple Queues	Manage Conges-tion (WRR)	Policing
Catalyst 6000	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Catalyst 5000	No	Yes	Yes	Yes	No	Yes	No	No
Catalyst 4000	No	Yes	No	No	No	No	No	No
Catalyst 3500	Yes	Yes	No	No	Yes	Yes	No	No

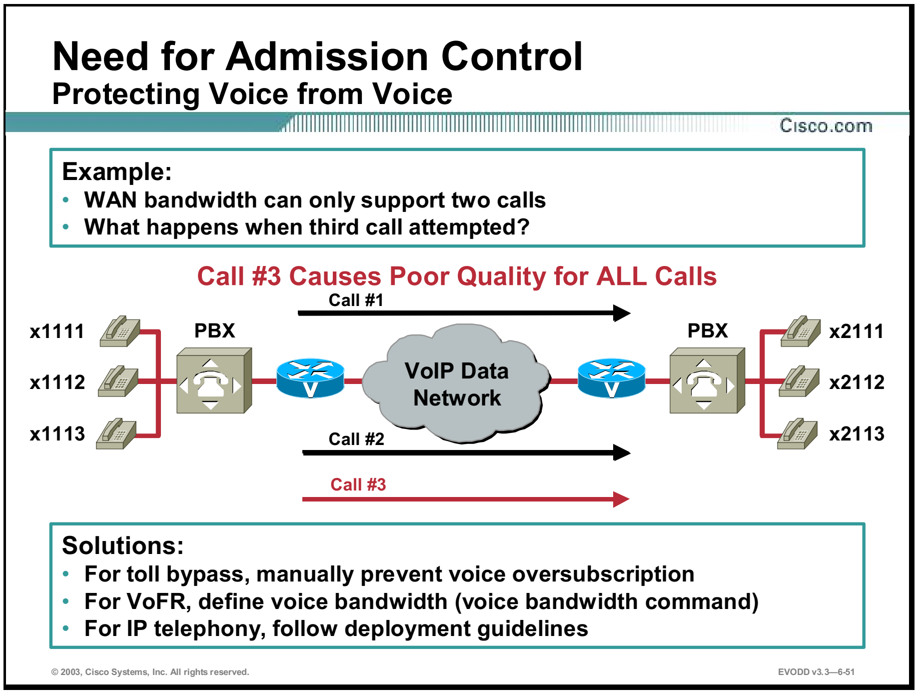
EVODD v3.3—6-50

Cisco recommends that you complete the following for QoS deployment:

- Create a trust boundary at the network edge in the wiring closet. Make ports trusted on the wiring closet switch where you attached IP Phones.
- Reclassify ToS at the edge if you cannot trust devices.
- Shift the trust boundary to the distribution layer and reclassify ToS there if you cannot reclassify it at the edge.
- Use a PQ if possible for delay-sensitive traffic.
- Use QoS ACLs for more granular classification of packets using Layer 4 information.
- Use policing, if necessary, to limit traffic for individual flows as well as aggregate flows.
- Classify traffic going to the WAN edge at Layer 3 so that the router can reference Layer 3 markings for advanced WAN queuing mechanisms.
- Use a WAN edge router as the classifier for very small remote site networks where you would not want to use a Layer 3-capable switch.

Admission Control

This topic details how admission control strategies and schemes protect voice traffic from other voice traffic.

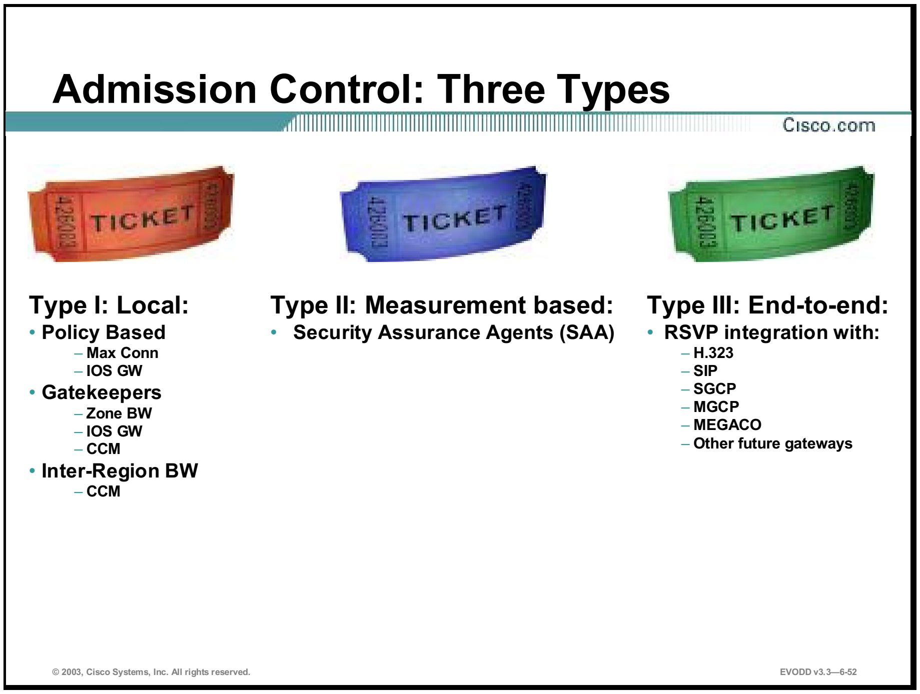


If a third call is added to and sent through a WAN link that mathematically can handle only two calls, all three calls will have their sound quality affected.

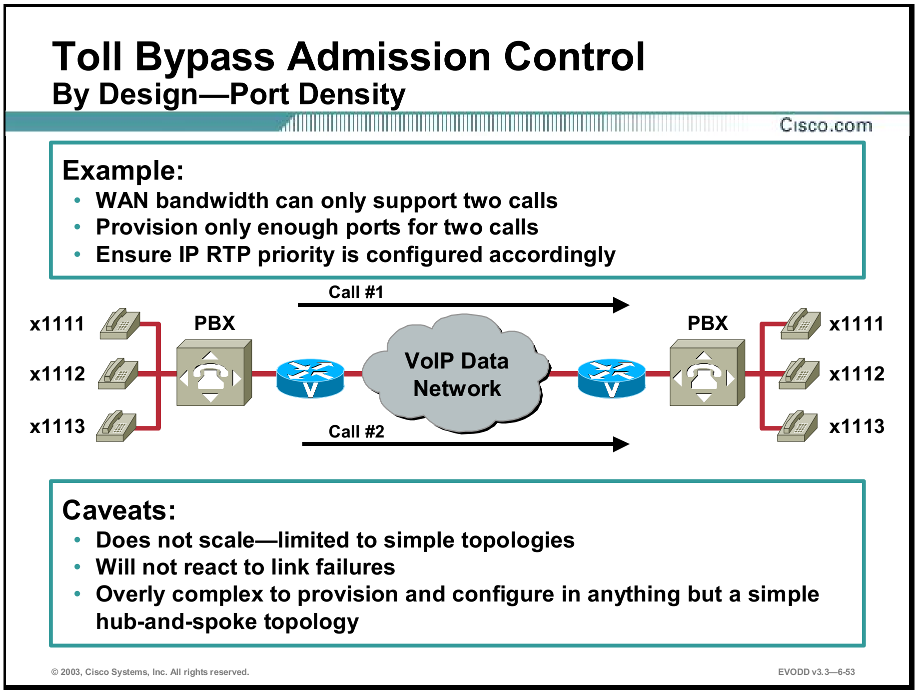
In the toll bypass environment of today, a manual method of admission control solves this problem. If a WAN link can handle only two calls, only two trunks out of the PBX are allowed to connect to the router. The PBX acts as the gating factor. If both trunks are busy, the PBX routes the call over the Public Switched Telephone Network (PSTN).

When deploying VoFR, the Cisco IOS **voice bandwidth** command defines the maximum bandwidth allocated for a voice channel over the PVC. When the voice channel reaches this bandwidth, the router rejects additional call attempts.

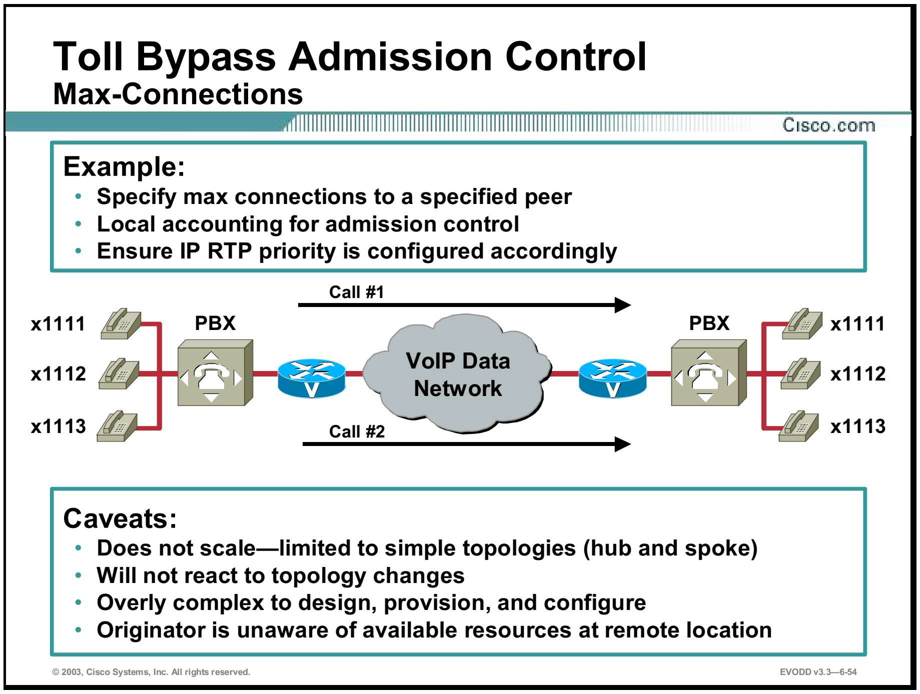
CCM uses a different admission control approach for end-to-end IP telephony. Shown here are a finite number of trunks leaving a PBX that telephones can use, no matter how many telephones are behind the PBX. However, in an IP telephony environment, an unknown number of telephones may have an open IP path to the WAN. Therefore, CCM, for example, can use an external gatekeeper to permit or deny access to the VoIP WAN.



You can break down Call Admission Control (CAC) schemes into three primary types: local, measurement-based, and end-to-end. A simple local CAC approach is to limit the number of physical trunks from a PBX to a router. Measurement-based CAC tools send Service Assurance Agent (SAA) probes out into the network to measure network conditions. End-to-end CAC uses RSVP to reserve an amount of bandwidth from the source router to the destination router for the duration of the call.



A local CAC tool is shown here. A PBX may have two physical connections to the voice-enabled router. After the PBX begins to use those physical paths, it denies future call attempts. This approach works with toll bypass Voice over Data networks. This CAC method is not suitable for a network with IP phones.



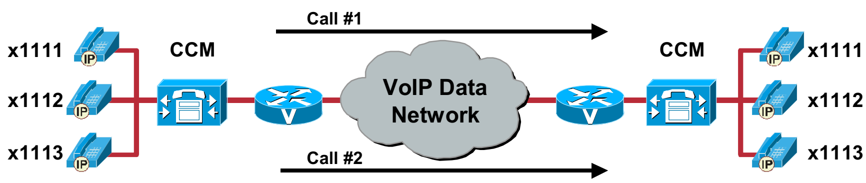
Another way to achieve admissions control is on a per-dial-peer basis using a concept called max connections.

Use the **max-conn** command in dial-peer configuration mode to limit the number of connections for a dial-peer. For example, you may have a dial-peer on the router that points to a PBX and indicates that the maximum connections are 12. As soon as the 12th call connects, the 13th call attempt bumps down to the next dial-peer and perhaps goes out the PSTN.

Toll Bypass Admission Control
Gatekeepers

Cisco.com

Example:
- Zone bandwidth used for admission control
- Ensure IP RTP priority is configured accordingly

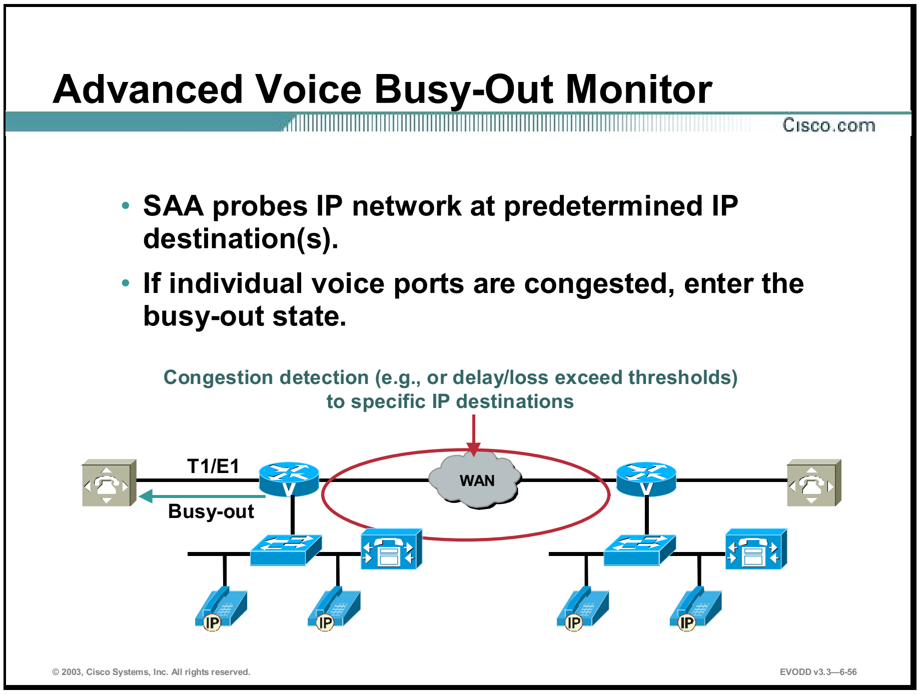


Caveats:
- Does not scale—limited to simple topologies (hub and spoke)
- Will not react to topology changes
- Overly complex to design, provision, and configure
- Adjust math if a single codec is used everywhere

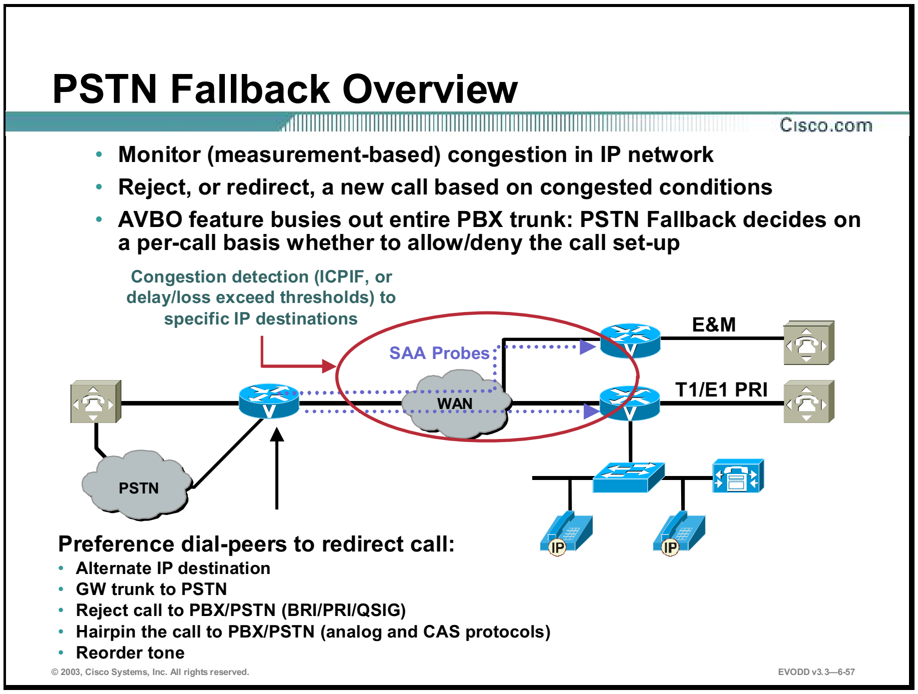
 EVODD v3.3—6-55

You can use a Cisco IOS gatekeeper and configure it as a zone. From a Cisco IOS point of view, each call uses a certain amount of bandwidth depending on the codec. You can define the number of calls per zone. The gatekeeper tracks how much bandwidth each call uses. This allows the gatekeeper to determine when to reject calls in order to maintain expectable voice quality.

The gatekeeper is not cognizant of the network topology. If the topology changes, it cannot adjust; therefore, it is limited to hub-and-spoke topologies.



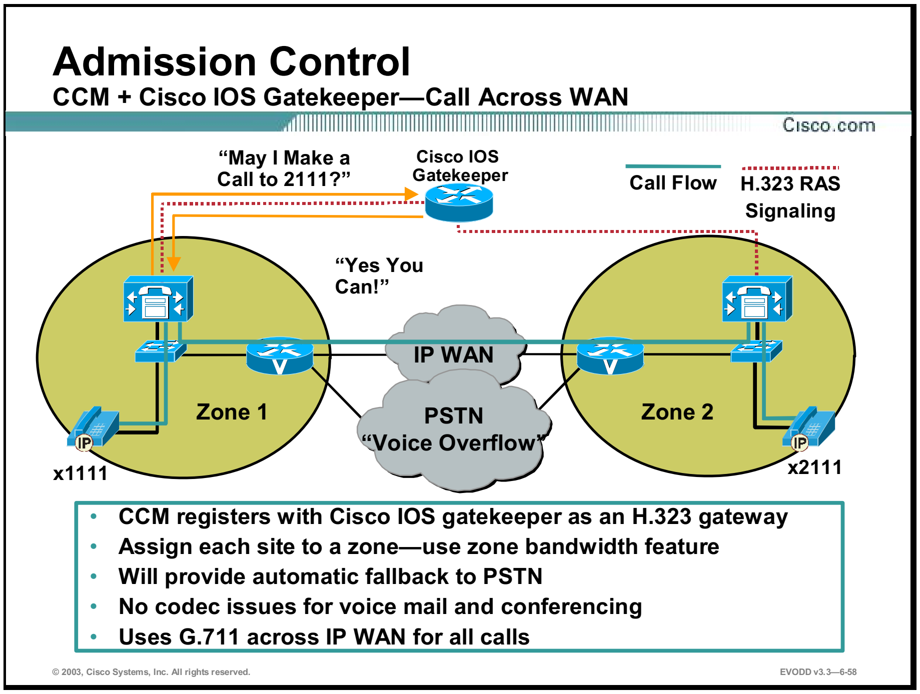
The advanced voice busyout (AVBO) monitor is a measurement-based CAC tool that can trigger SAA probes to one or more remote IP destinations. These probes return information (such as explicit loss or delay values) that can trigger a busyout of the connection to the PBX. AVBO can perform PBX trunk, or individual voice ports, busyout based on the current conditions of the IP network, as shown in the figure.



Like AVBO, PSTN fallback is a measurement-based CAC tool. PSTN fallback does not busyout trunks. Instead, the CAC decision is triggered only when you attempt a call setup.

PSTN fallback, based on SAA probes, has all the benefits and drawbacks of measurement-based techniques. PSTN fallback is unusually flexible because it can make CAC decisions based on any type of IP network, including the Internet. IP networks carry the SAA probe packet as just another IP packet. Therefore, the customer backbone network can contain one or more service provider networks, or the Internet, or any combination. The only requirement is that the destination device (the owner of the IP address the probe is sent to) must support SAA responder functionality.

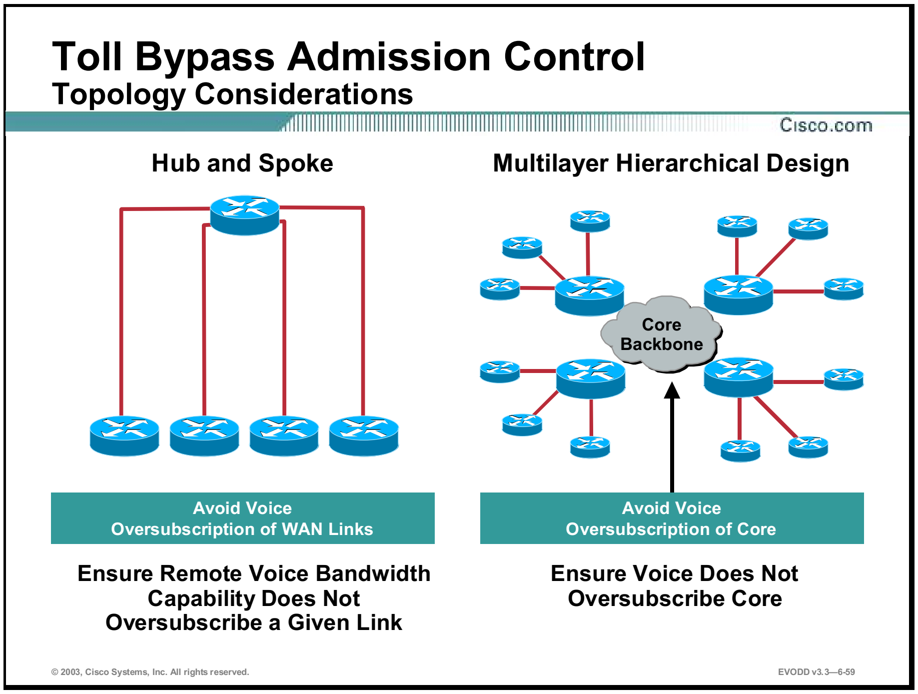
The customer network at the destination site should include the destination device (with a service provider network in between). Therefore, you cannot use PSTN fallback directly to IP Phones, but you can use it indirectly if IP Phones are behind a Cisco IOS router that supports the SAA responder. The destination device does not need to support the PSTN fallback feature; only the SAA probe responder is needed.



The most widely deployable admission control option is to have CCMs at each location that are networked together via H.323.

In that instance, you can use the zone bandwidth feature to limit the number of calls. This feature is a call accounting mechanism between regions.

If the IP WAN can no longer accommodate any calls, it will automatically fall back and use the PSTN.

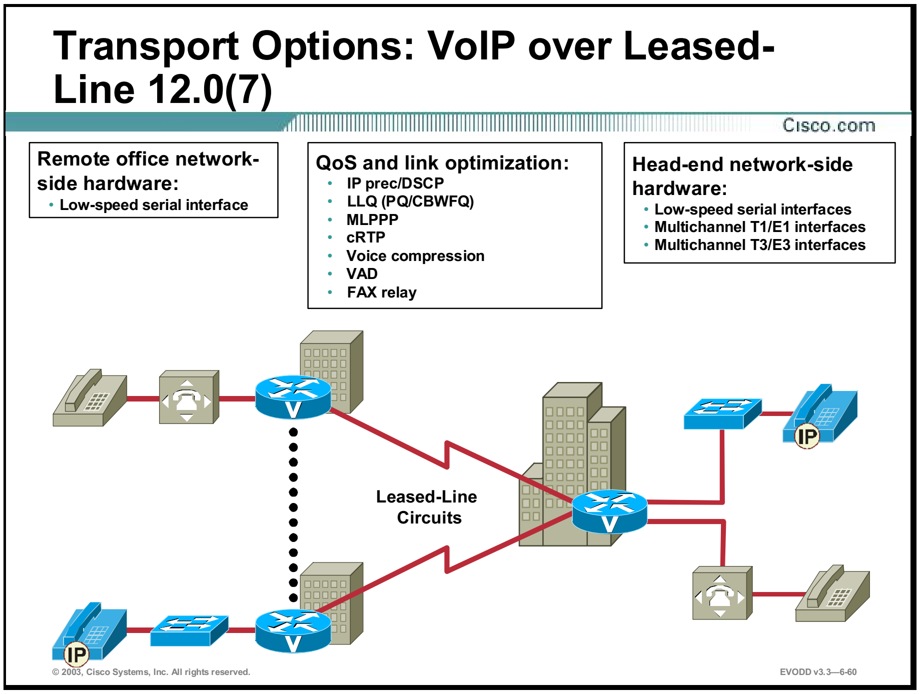


All of the admission control schemes available, regardless of whether they are for toll bypass or for a Cisco IP Telephony network, are limited to hub-and-spoke topologies with the exception of RSVP. They are not cognizant of the network topology.

A hierarchical network uses the same mechanisms only if they do not oversubscribe the backbone. This design goal is very difficult to accomplish in a WAN. Therefore, you are limited to hub-and-spoke topologies from a CAC mechanism perspective, because you can define only local policies in most cases.

Transport Options

This topic suggests appropriate QoS tools for use in various WAN topologies.

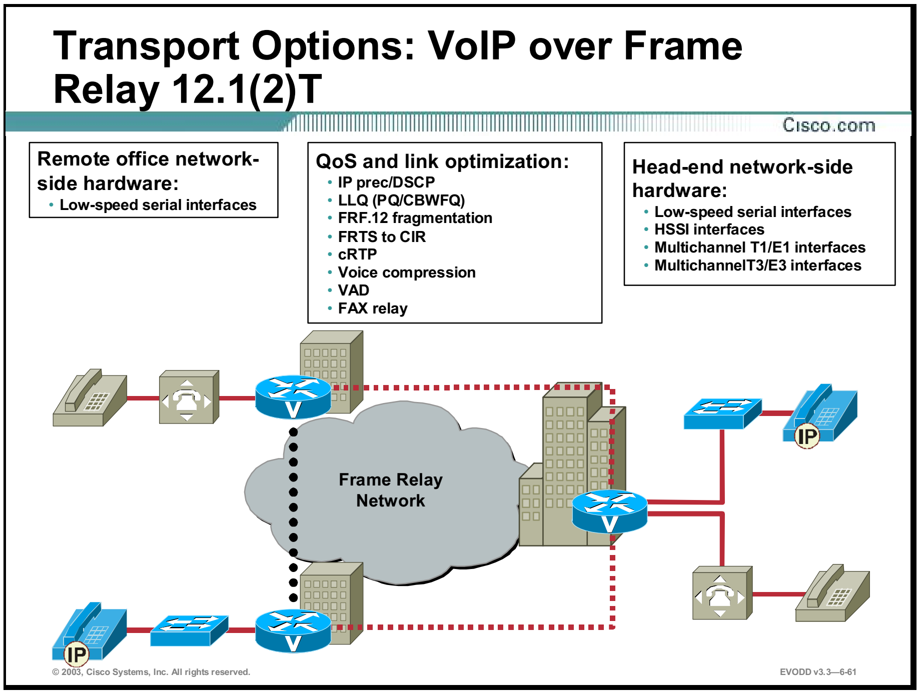


Cisco recommends the following QoS and link optimization services available through Cisco IOS for VoIP over leased-line circuits:

- DSCP, the standardized method of differentiating traffic via the IP ToS field, replaces the IP precedence method, which provided less granularity than DSCP.
- LLQ provides a strict PQ method for voice traffic and guarantees an amount of bandwidth for other defined data classes.
- MLP is utilized for PPP circuits running at speeds less than 768 kbps. For circuits greater than 768 kbps, High-Level Data Link Control (HDLC) encapsulation can be utilized, because it is more efficient than PPP. LFI is not necessary for higher speed circuits running at speeds greater than 768 kbps.
- cRTP compresses the RTP, UDP, and IP headers from 40-byte to 2-4-byte. This compression reduces overhead on the circuit.
- Cisco supports a wide range of voice compression algorithms, including G.729, G.723, and G.711. You should base your choice of a compression algorithm upon the voice quality you desire and the amount of bandwidth available for voice transport.

- Voice activity detection optimizes link utilization by not sending packets if there is no voice traffic.
- You can also use fax relay services to provide optimization for fax traffic.

At the head end, you can use a variety of serial interfaces. Individual circuits can terminate on standard Cisco low-speed serial interfaces. The carrier can also aggregate the circuits and bring them in on a multichannel T1/E1 or T3/E3 interface to lower overall costs and ease maintenance.



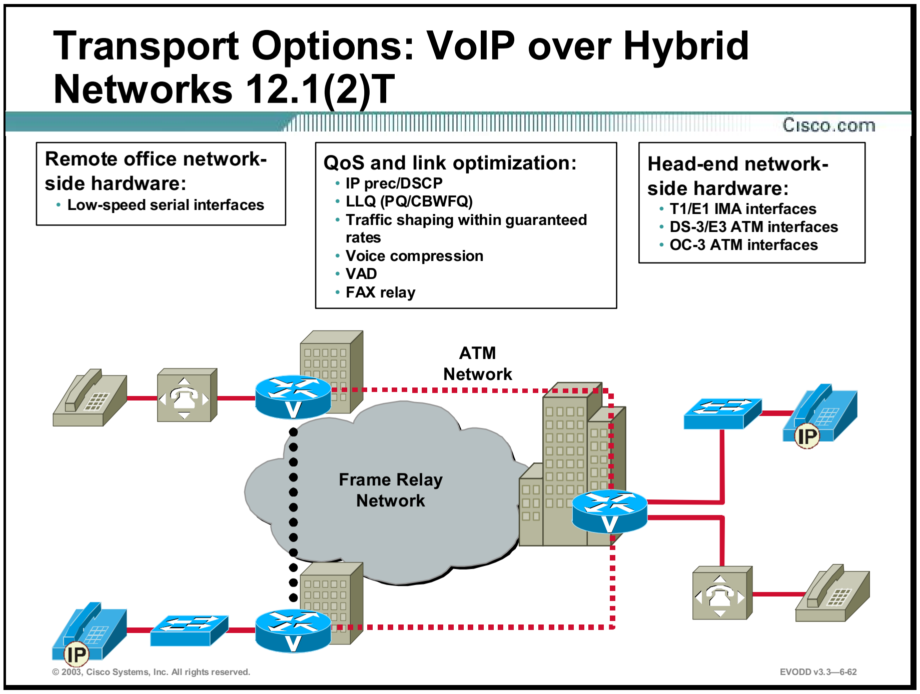
The network-side hardware for a remote office connecting into a Frame Relay cloud is typically a standard Cisco low-speed serial interface.

Cisco recommends the following QoS and link optimization services available through Cisco IOS for VoIPoFR circuits:

- DSCP, the standardized method of differentiating traffic via the IP ToS field, replaces the IP precedence method, which provided less granularity than DSCP.
- LLQ provides a strict PQ method for voice traffic and guarantees an amount of bandwidth for other defined data classes.
- You should implement standards-based FRF.12 fragmentation as your LFI method on low-speed Frame Relay circuits.
- In order to guarantee voice quality, set Frame Relay traffic shaping to the CIR. Also, do not oversubscribe at the head-end.
- cRTP compresses the RTP, UDP, and IP headers from 40-byte to 2-4-byte. This compression reduces overhead on the circuit.
- Cisco supports a wide range of voice compression algorithms, including G.729, G.723, and G.711. You should base your choice of a compression algorithm upon the voice quality you desire and the amount of bandwidth available for voice transport.
- Voice activity detection optimizes link utilization by not sending packets when there is no voice traffic.

- You can also use fax relay services to provide optimization for fax traffic.

At the head end, you can use a variety of serial interfaces. Individual circuits can terminate on standard Cisco low-speed serial interfaces. The carrier can also aggregate the circuits and bring them in on a T1/E1 or T3/E3 interface to lower overall costs and ease maintenance.



In this solution, the remote side consists of Frame Relay service, while the head-end side consists of ATM service. The service provider provides the FRF.8.

The network-side hardware for the remote office connecting into a Frame Relay cloud is typically a standard Cisco low-speed serial interface.

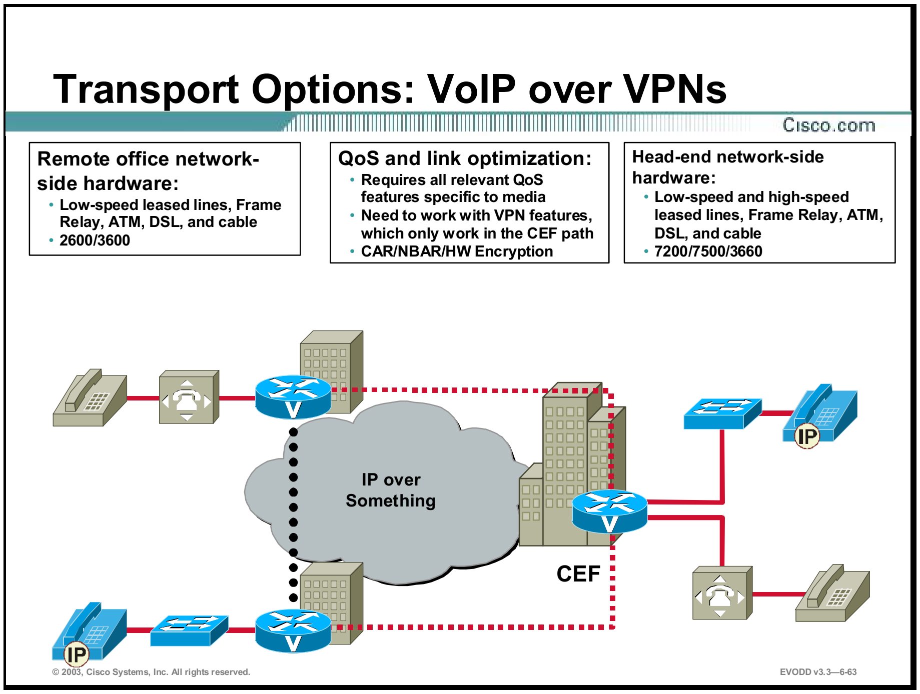
Cisco recommends the following QoS and link optimization services available through Cisco IOS for VoIP over a hybrid network:

- DSCP, the standardized method of differentiating traffic via the IP ToS field, replaces the IP precedence method, which provided less granularity than DSCP.
- LLQ provides a strict PQ method for voice traffic and guarantees an amount of bandwidth for other defined data classes.
- For PVC speeds of less than 768 kbps, Cisco recommends separate PVCs for voice and data traffic. Although the ATM SAR function fragments all traffic to 53-byte cells, no interleaving function exists that guarantees prioritization of voice cells over data cells within a single PVC.
- To guarantee voice quality, set traffic shaping to the guaranteed rate of the Frame Relay or ATM service. Also, do not oversubscribe at the head end.
- Cisco supports a wide range of voice compression algorithms, including G.729, G.723, and G.711. You should base your choice of a compression algorithm upon the voice quality you desire and the amount of bandwidth available for voice transport.

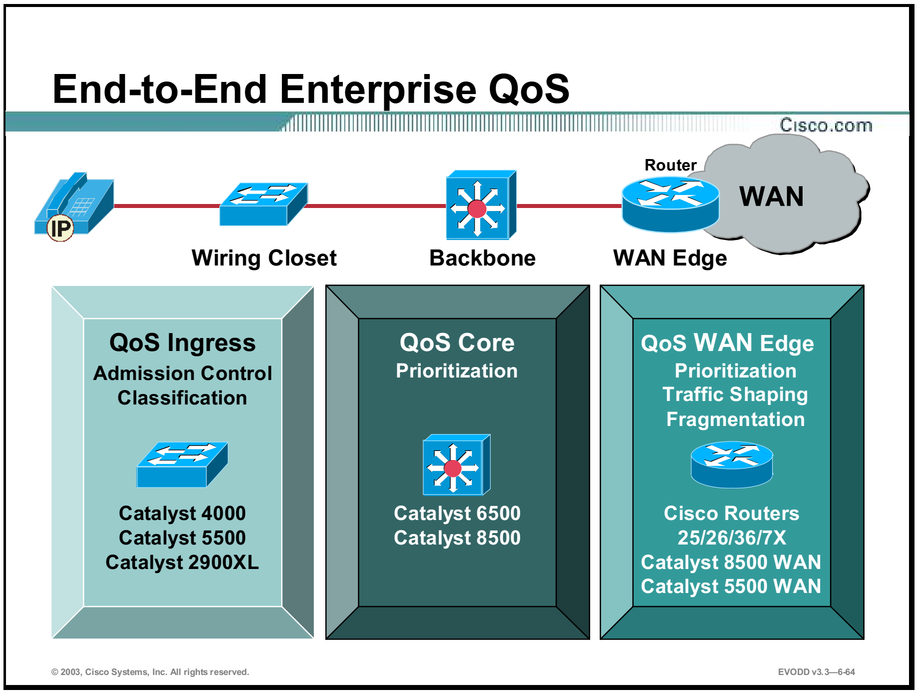
- Voice activity detection optimizes link utilization by not sending packets when there is no voice traffic.

- You can also use fax relay services to provide optimization for fax traffic.

At the head end, you can use a variety of interfaces. Individual circuits and/or PVCs can terminate on digital signal level 1 (DS-1) ATM interfaces. The carrier can also aggregate the circuits or PVCs and bring them in on a DS-3 or Optical Carrier 3 (OC-3) ATM interface to lower overall costs and ease maintenance.



Most of the QoS techniques do not work in the Cisco Express Forwarding (CEF) path. Therefore, installing VoIP over a Virtual Private Network (VPN) is a challenge. Depending on the local route media, VPNs cannot run with the QoS features they need.



QoS requires attention throughout the enterprise to ensure a QoS-enabled infrastructure over which voice traffic can run.

In the wiring closet at the campus, use classification tools to classify traffic so that the infrastructure treats voice with priority throughout its travel. Admissions control also occurs here so that voice does not oversubscribe its allocated bandwidth.

In the QoS core, LLQ now supplants queuing strategies like CBWFQ and IP RTP priority to ensure that the infrastructure handles voice with preference above other data traffic.

At the WAN edge, you must prioritize traffic. This prioritization sometimes includes traffic shaping and fragmenting.

Lay the Foundation for a Converged Network

Cisco.com

Wide-Area Network:

- QoS-enable the WAN
- Install voice-capable gateways
- Toll bypass
- Capacity planning for bandwidth:
  - Video
  - Voice
  - Data

Local-Area Network:

- Switched infrastructure is a must
- Switched to desktop is a must
- QoS-enable the LAN
- Large site hierarchical design
- Network design for capacity and fast convergence at Layer 2 and 3

 EVODD v3.3—6-65

Laying the foundation for the converged network is the first step, and it has two components:

- **LAN:** You can deploy immediately at the LAN. Switched infrastructure and switched connections to every desktop is a requirement. You need to QoS-enable the LAN where appropriate. If you want to deploy on a large site, you must adopt a hierarchical design that has capacity at both Layer 2 and Layer 3 and allows for fast convergence. This requires fast converging routing protocols such as Open Shortest Path First (OSPF) or Enhanced Interior Gateway Routing Protocol (EIGRP). Features such as the Hot Standby Router Protocol (HSRP) are a must. If not, voice will not tolerate the delay of the Routing Information Protocol (RIP), for example, which can take several minutes to converge. If possible, convergence should take less than 1 second.

- **WAN:** You must QoS-enable the WAN. You must put all of the tools in place and install voice-capable gateways to access the PSTN. You could use the current Cisco router for the gateway, or you may require an upgrade of existing equipment.

Toll bypass connections are the first step in IP telephony. With toll bypass, routers interconnect existing PBXs and prepare the infrastructure for eventual end-to-end IP telephony. Capacity planning is the most important part of this process.

IOS Support for Queuing by Media

	Leased Lines	VoFR	VoIPoFR	VoATM	VoIPoATM
CBWFQ	12.1(5)T	12.1(5)T	12.1(5)T	12.1(5)T	12.1(5)T
IPRTP	12.1(5)T	12.1(5)T	12.0(7)T	Not Supported	Not Supported
LLQ	12.0(7)T	12.0(7)T	12.1(2)T Future Support	12.0(7)T	12.0(7)T

Does your hardware support the feature?
Is there enough DRAM/Flash memory installed to support it?

EVODD v3.3—6-66

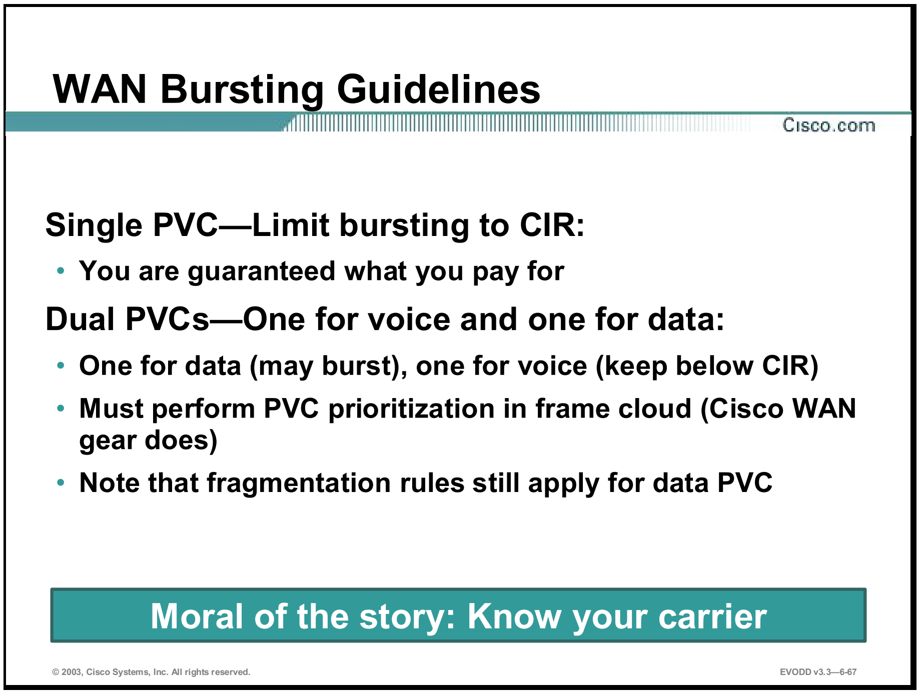
The mechanisms required to actively administer voice on a data network are available in various versions of Cisco IOS. These include:

- WFQ is available in Cisco IOS Release 11.0
- IP precedence is available in Cisco IOS Release 11.0
- Bandwidth Reservation Protocol is available in Cisco IOS Release 11.2
- MLP fragmentation is available in Cisco IOS Release 11.3
- DSCP support is available with Cisco IOS Release 12.1 (3) T

A number of features are scattered over Cisco IOS software. You should audit the existing hardware platforms on the network, including whether the hardware supports the features you need.

Certain ATM adapters, for example, may not support certain prioritization needs because the chip set will not support them. You must install adequate DRAM and Flash memory in the router to support the new release.

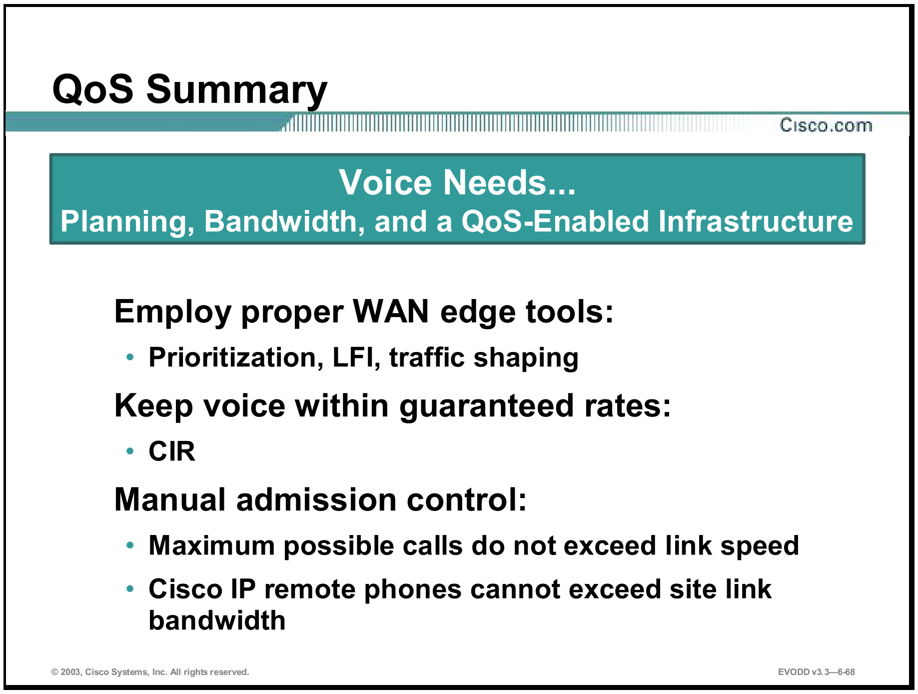
You should audit the Cisco IOS platforms across the network before deploying the voice. Therefore, you can prepare to deploy the features necessary for a converged voice and data network.



If you are running a single PVC in the WAN, whether it is Frame Relay or ATM, limit the bursting on that PVC to the CIR or sustainable cell rate (SCR). Keep voice within guaranteed domains to ensure success in the WAN.

Bursting for data is possible. If you run two PVCs, particularly for Frame Relay, you can use one for the voice maintained under CIR, and the data PVC can burst to the line rates.

A PVC running data only on a slow-speed clock rate interface still needs fragmentation, because the data and voice PVCs still use single serial interface. Know the characteristics, oversubscription rates, and guarantees of your carrier to correctly configure traffic shaping.



To summarize QoS, remember that voice is not free. Voice takes planning, bandwidth (which may cost money), and a QoS-enabled infrastructure. That infrastructure may require upgrades to hardware and possibly software.

Laying the proper foundation enables lower cost of ownership as these latency-sensitive applications evolve.

At the WAN edge, use prioritization tools, slow link tools, and traffic-shaping tools. Also, take precautions to protect voice from voice using the range of Cisco CAC tools.

Summary

This topic summarizes the key points discussed in this lesson.

Summary

Cisco.com

- To ensure voice quality, the network must minimize loss, delay, delay variation, and echo.
- As voice and data leave the router, you have three areas of concern: prioritization, slow-speed links, and traffic shaping.
- Prioritization protects voice from data by ensuring that the voice packets are out in front of the data packets.
- Slow link efficiency tools fix the problem of large packets having to clock on a serial link.
- Traffic shaping prevents site speed mismatch, to avoid remote-to-central site oversubscription, and to prohibit bursting above a committed rate.
- You must use QoS tools throughout the WAN and the LAN.
- Create a trust boundary at the network edge in the wiring closet. Make ports trusted on the wiring closet switch where you attached IP Phones.
- Admission control tools act as an external gatekeeper to permit or deny access to the VoIP WAN.
- There are several key prerequisites to designing a Voice over Data design. These include: ensuring that a voice-enabled infrastructure is in place; QoS mechanisms in place that will prioritize voice traffic and ensure its timely delivery ahead of other data; classify traffic and choose a queuing scheme; protect voice traffic from data traffic.

 EVODD v3.3—6-69

Next Steps

After completing this lesson, go to:

- Comprehensive Design Strategies lesson

References

For additional information, refer to these resources:

- Voice Quality: http://www.cisco.com/warp/public/788/voice-qos/voice-qos.shtml
- Cisco IP Telephony QoS Design Guide: ttp://www.cisco.com/univercd/cc/td/doc/product/voice/ip_tele/avvidqos/index.htm

Laboratory Exercise: QoS

The laboratory exercises are designed to reinforce concepts discussed throughout the course. This laboratory exercise considers quality of service approaches for varied Voice over Data topologies.

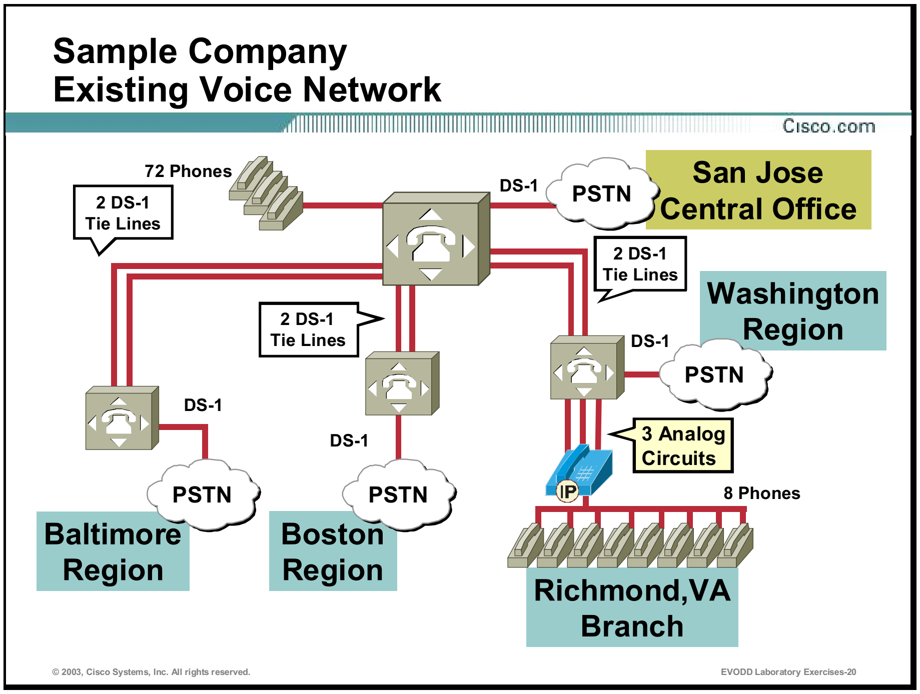
Exercise Objective

In this exercise, you will design a solution to maintain voice quality while merging voice traffic onto an existing data network.

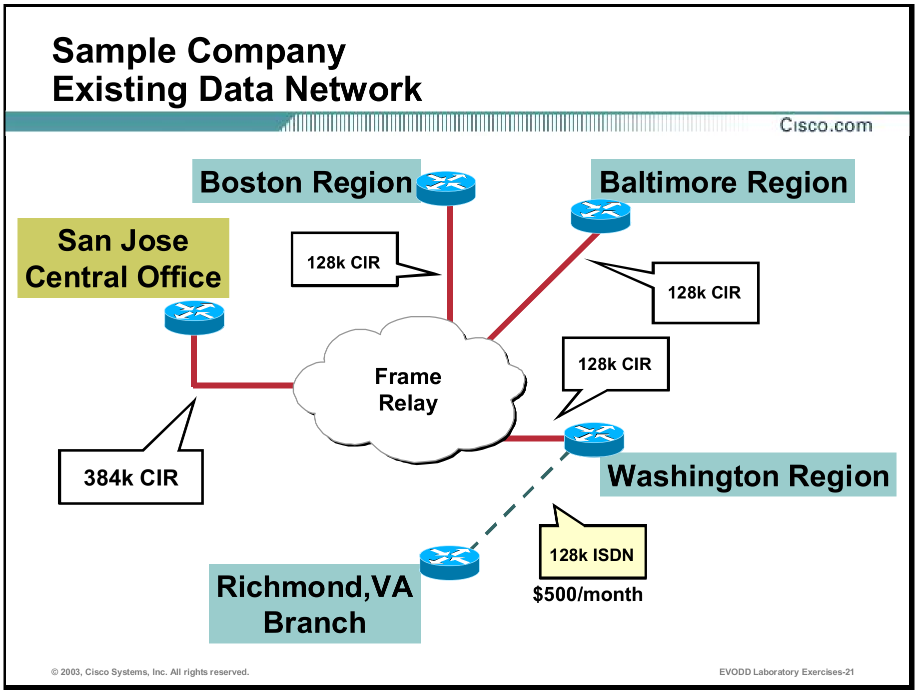
After completing this exercise, you will be able to:

- Analyze an existing voice network
- Design a solution to implement voice and data traffic on network while maintaining voice quality

Exercise Procedure



- The existing voice network for this enterprise is as follows:
  - Central office and three regional offices have PBXs
  - Each of the four PBXs is connected to the PSTN via a DS-1
  - Each regional office is connected to the central office by 2 DS-1 circuits provisioned as 48 DS-0s
  - The Washington region is connected to the Richmond branch keyset by three DS-0 analog circuits
  - The Richmond branch also has a connection to the PSTN by a fractional DS-1 configured as four DS-0s
- The staffing of the locations and the total number of telephones is as follows:
  - Central office - 65/72
  - Washington - 25/30
  - Baltimore and Boston - 20/25 each
  - Richmond - 8/8



The customer has an existing Frame Relay network used for data only. They have the following links:

- From the central office in San Jose to the Frame Relay WAN—384-kbps CIR
- From each of the three regional offices—Boston, Baltimore, and Washington—to the Frame Relay WAN—128-kbps CIR
- Between the Washington regional office and its Richmond, Virginia, branch—128-kbps ISDN link

Assume the ISDN link costs $500 per month, which is a flat rate.

Practice

The company wants to merge its voice traffic onto an existing data network. Determine the interfaces that require QoS tools and identify the QoS tools that need to be implemented. Use the QoS chart to record your solutions.

Table: QoS Chart

Tools	Leased Lines	VoIPoFR	VoFR
Classification			
Prioritization			
Slow-Speed Link Efficiency			
Traffic Shaping Required			
Admission Control Required			

Lesson Review

This practice exercise reviews what you have learned in this lesson.

Q1) What is the queuing strategy that Cisco recommends for voice traffic?

A) WFQ

B) CBWFQ

C) IP RTP priority

D) LLQ

Q2) Which two of the following are marking types? (Choose two)

A) IP precedence

B) WFQ

C) LLQ

D) DSCP

Q3) How many ToS bits are used by DSCP?

A) 3

B) 6

C) 8

D) 16

Q4) What is the serialization delay for a 1,500-byte frame exiting a serial interface running at 56 kbps?

A) 10 ms

B) 20 ms

C) 150 ms

D) 214 ms

Q5) When configuring traffic shaping, what is the timing interval if the CIR = 64 kbps and the Bc = 8,000 bits?

A) 20 ms

B) 125 ms

C) 150 ms

D) 214 ms

Q6) What are the two steps for campus QoS implementation? (Choose two.)

A) Classification

B) Marking

C) Queuing

D) Policing

Q7) A Cisco IP Phone marks voice frames with what CoS value?

A) 2

B) 5

C) 7

D) 64

Q8) What type of admission control tools uses SAA probes?

A) local

B) measurement-based

C) end-to-end

D) remote

Q9) What LFI tool would be appropriate for a PPP network?

A) FRF.11 Annex C

B) FRF.12

C) MLP

D) CRTP

Comprehensive Design Strategies

Overview

Numerous design strategies have been presented in this course. This lesson takes what you have learned in the course and uses that information in a comprehensive design case study.

Importance

To design a successful migration from a legacy PBX-based telephony system to an IP telephony system, you cannot view the tools presented in this course in isolation. Rather, the tools, strategies, and design guidelines presented to you throughout this course must be used together to help you produce a comprehensive solution.

Objectives

Upon completing this lesson, you will be able to:

- Describe legacy PBX and messaging features to be migrated
- Identify design requirements based on currently installed customer equipment
- Describe migration stages and strategies
- Select a WAN topology for a Voice over Data network, given the characteristics of the existing network
- Select LAN components for a Voice over Data network, given the characteristics of the existing LAN infrastructure
- Identify the quality of service tools to use in the Voice over Data network, given the WAN and LAN characteristics

- Identify the number of Cisco Call Managers and Unity servers required to support a specified number of users

Learner Skills and Knowledge

To fully benefit from this lesson, you must have these prerequisite skills and knowledge:

- Cisco Voice over Data design recommendations as presented throughout the preceding modules in this course
- *Cisco IP Telephony* (CIPT) course

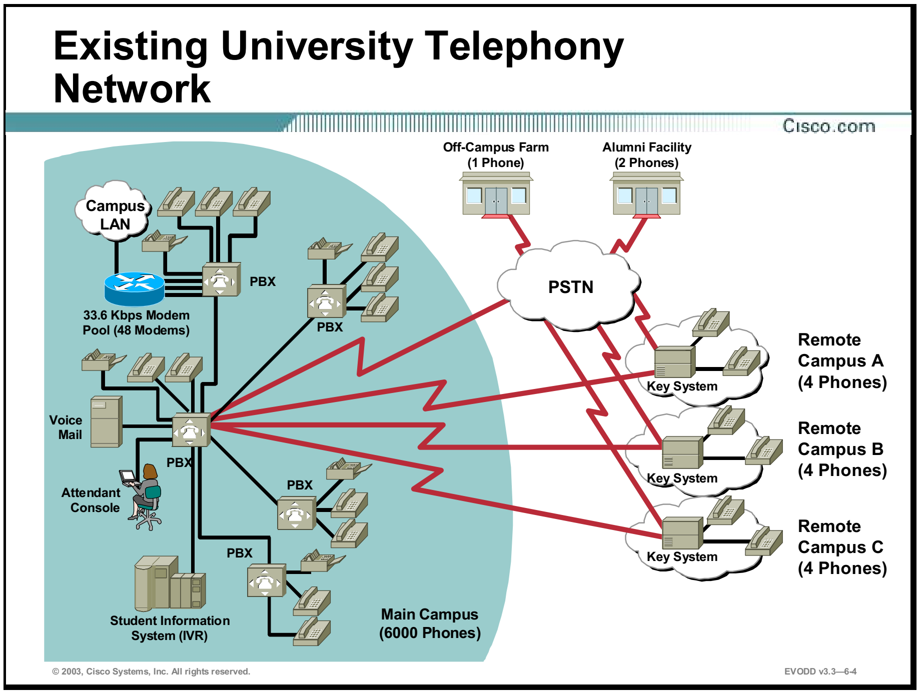
Outline

This lesson includes these topics:

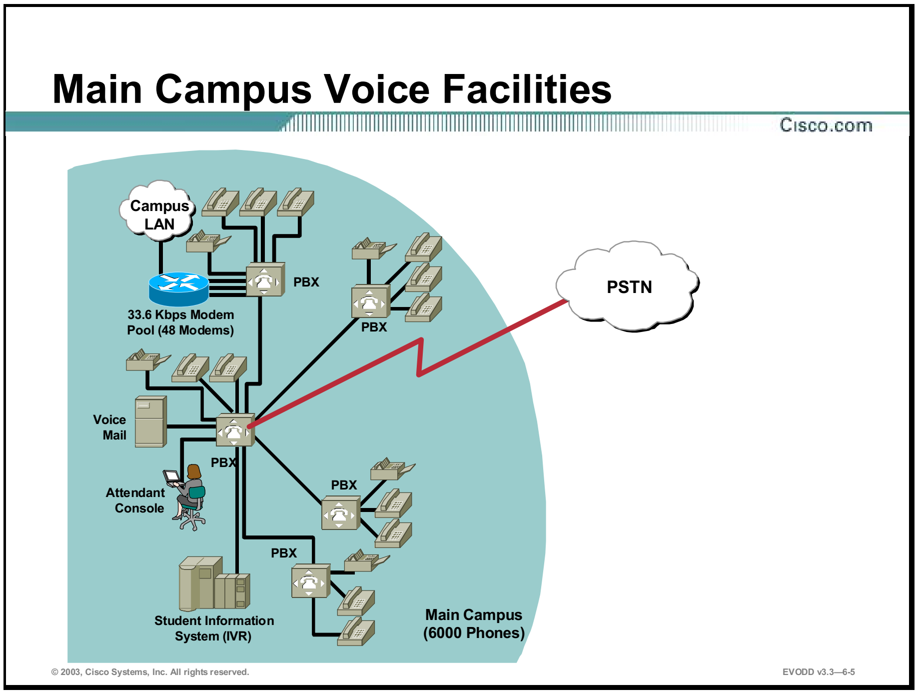
- Overview
- Determination of Current Features
- Identification of Feature Replacements
- Migration Stages and Strategies
- WAN Considerations
- LAN Considerations
- Quality of Service Considerations
- Cisco CallManager and Unity Scaling Considerations
- Summary
- Lesson Review

Determination of Current Features

This topic identifies voice features currently in place in the case study scenario. Before you start on your Voice over Data migration design, the first step is to determine the voice features in place today that require migration to the Voice over Data environment.



This case study examines the migration from a legacy PBX-based telephony system to an IP telephony system at a university. The facilities span six geographical locations. This topic details the telephony facilities at each of these locations.



The main campus location has five PBXs spread throughout the campus. One of the five PBXs is designated as the control cabinet. The other four PBXs feed back to the control cabinet for call processing. This PBX design is called a dispersed control model. The PBX uses nine T-1 circuits (for example, 214 voice channels) to connect to the Public Switched Telephone Network (PSTN).

These five PBX nodes service approximately 6,000 analog telephones. Each telephone allows you to connect to the nearest PBX node using existing station wire that has been in place for over 25 years. In addition to analog telephones, each PBX supports other plain old telephone service (POTS) devices, such as fax machines and conference speakerphones.

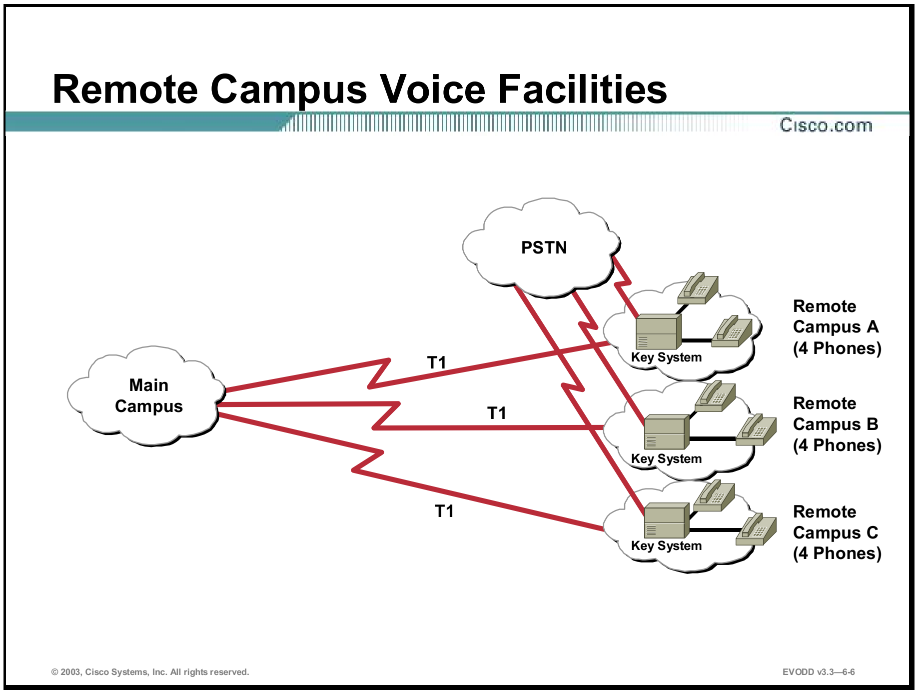
The campus operator has an attendant console that is directly connected to the PBX via a 25-pair cable. This particular PBX model requires that you have an attendant console directly attached to the PBX control cabinet. Therefore, the campus operator has to be located in the same building as the control cabinet. In the past, this limitation has caused scheduling issues for operator shift changes. You would like to have attendant consoles in two buildings, but this has not been possible due to PBX limitations.

Five years ago, the university added a voice mail system. The voice mail system connects to the PBX via 24 physical POTS connections from the voice-mail server to PBX station cards.

A student can register for classes, check grades, and verify housing information via the Student Information System. This system runs on a third-party interactive voice response (IVR) system. The IVR system connects to the PBX via 24 physical POTS connections and to the university mainframe via an Ethernet connection.

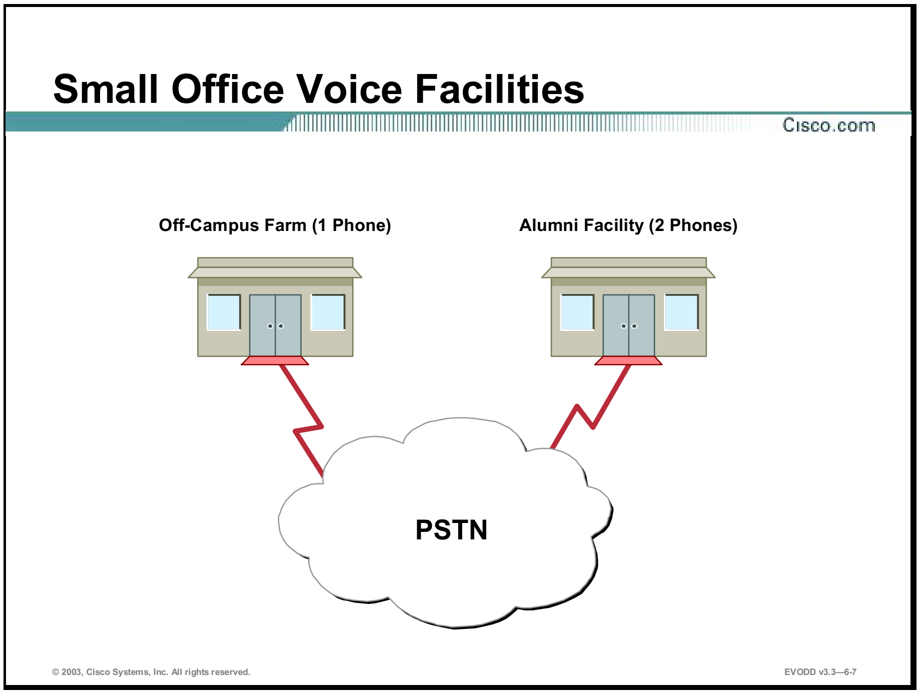
Off-campus students can dial into the university and access UNIX hosts—for Computer Science and Computer Information System classes—via a 48-port modem pool. The modem pool consists of a Cisco Systems 3640 router with three 16-port analog modem modules.

Because the Cisco 3640 connects to its local PBX node via 48 physical analog connections, the maximum modem speed is 33.6 kbps.



There are three remote campus locations located in surrounding counties. These remote campuses are located in rented office space and serve nontraditional students in those remote areas. Because of the small size of these remote campuses, each site has only four POTS connections, one of which also functions as a fax connection.

A key system services the telephones at each remote location. For toll-bypass purposes, each key system has a dedicated T-1 circuit connecting back to the main campus. For local calls, you have two POTS connections from each key system to the local PSTN central office.

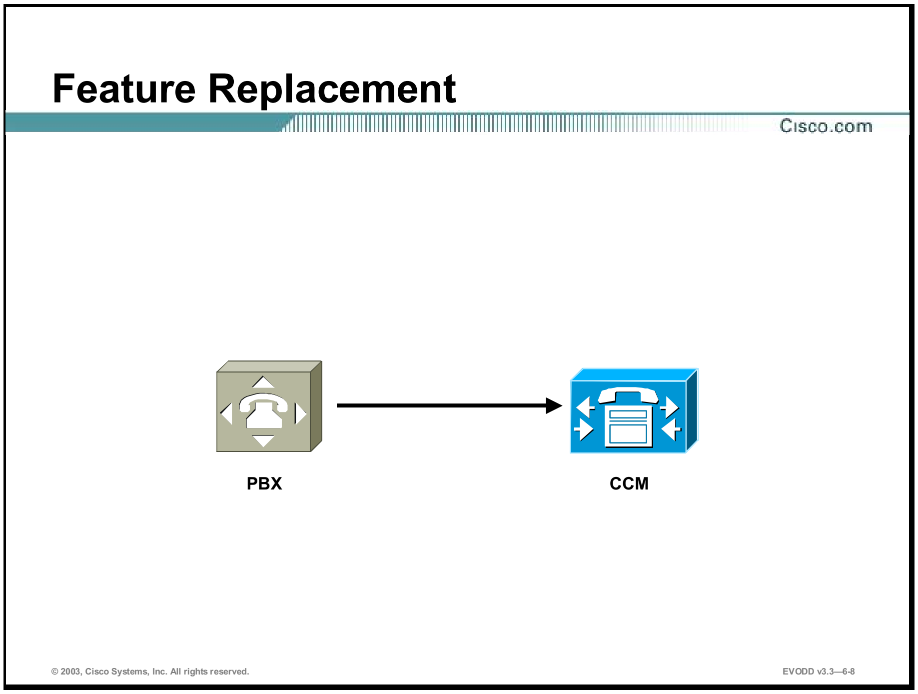


There are two small offices located off-campus. One of the small offices is located at a farm maintained by the university. This farm is located approximately 10 miles from the main campus. Because there is a single telephone located at the farm office, it does not connect directly back to the university. Instead, the telephone has its own POTS connection back to the local central office (CO).

The second off-campus office houses the university alumni office, which also serves as a reception area for visiting alumni. This is an older home, located approximately 1 mile from the main campus, which was donated to the university. The alumni facility has two POTS telephones, one of which doubles as a fax machine. Because of the small number of devices, the telephony service is provided by two POTS connections going back to the local CO.

Identification of Feature Replacements

This topic maps existing PBX features to their Cisco IP telephony counterparts. With current PBX features defined, the next step is to identify Cisco solutions to replace or enhance existing features.



The following pages elaborate on the *Feature Replacements* table, which suggests possible Cisco IP telephony replacements for the existing PBX features. You must be aware that migrations of this scale happen over time. Therefore, you must have interim solutions for some features, such as voice mail.

Table: Feature Replacements

Existing PBX Feature	IP Telephony Replacement Feature
Main Campus - 9 T-1s from PBX to CO.	Catalyst 6500 with two WS-X6608-T1 Modules
Main Campus - 6,000 Telephones (including fax)	Cisco CallManager 3.3 Cluster (call processing IP Phones (replacement for analog telephones) Catalyst 6500s (with in-line power capable 48-port Ethernet modules) Catalyst 3550s (with in-line power and FXS[1] ports for remaining POTS devices such as fax machines)
Main Campus - Operator Console	Cisco CallManager 3.3 Attendant Console
Main Campus - Voice Mail	Short Term - Catalyst 6500 with WS-X6624-FXS module running SMDI[2] protocol to CCM Long Term - Cisco Unity
Main Campus - Student Information System	Cisco IP IVR
Main Campus - 48-port 33.6 kbps Modem Pool	Cisco AS5300 with MICA[3] Modems (56 kbps support)
Remote Campus – 4 Telephones (including fax)	Cisco 1700 router (with SRST[4] support, FXS port for fax support, and FXO[5] port for PSTN support)
Small Office - 1 or 2 Telephones	Cisco 1700 router (with 2-port FXS VIC[6] and Ethernet WIC[7] connection to DSL[8])

1 FXS = Foreign Exchange Station

2 SMDI = Simplified Message Desk Interface

3 MICA = Modern ISDN channel aggregation

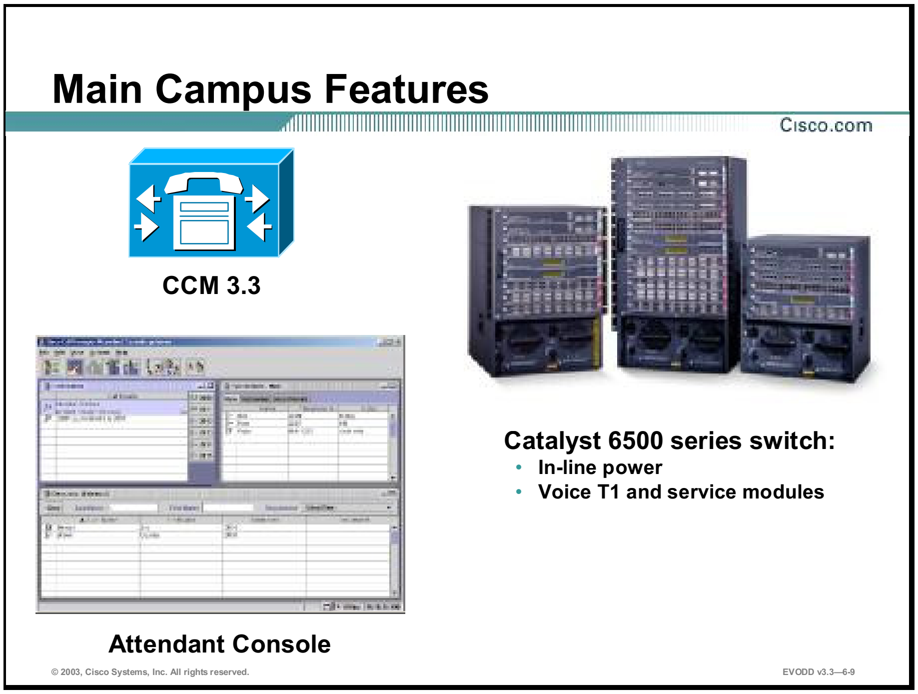
4 SRST = Survivable Remote Site Telephony

5 FXO = Foreign Exchange Office

6 VIC = voice interface card

7 WIC = WAN interface card

8 DSL = Digital Subscriber Line



At the main campus location, there are nine T-1 circuits that interconnect the PBX system to the local CO. After you replace the PBX, these digital circuits need a place to terminate. Under the Cisco Architecture for Voice, Video and Integrated Data (AVVID) umbrella, a Catalyst 6500 series switch equipped with two WS-X6608-T1 modules could accommodate nine T-1 circuits. Each of the WS-X6608-T1 modules supports eight T-1 circuits. The first WS-X6608-T1 module is completely populated with T-1 circuits. You will need the second module to support a single T-1 circuit only. Cisco CallManager (CCM) can use the Domain Specific Parts (DSPs) behind the remaining seven T-1 ports for services, such as transcoding and conferencing.

The existing PBX-based telephony network supports 6,000 telephones, which can be supported with a single CCM cluster. The analog telephones will be replaced with Cisco IP Phones, in phases, until the PBX is removed. However, even after the PBX is removed, you will still need support for legacy POTS devices, such as fax machines and conference room speakerphones.

You can deploy the ATA 188 into wiring closets requiring FXS capability to satisfy this requirement. The ATA 188 supports two FXS ports, which is sufficient for the isolated pockets of POTS devices that will *not* be replaced. Additionally, you can deploy the Catalyst 3550 switches to support 24 inline power ports for the Cisco IP Phones. Depending on the density of IP Phones in a geographical area, Catalyst 6500s with in-line powered Ethernet modules may be deployed to support IP Phones.

Before CCM 3.3, Cisco offered the WebAttendant as a replacement for an existing attendant console. CCM 3.3 provides you with a completely rewritten client portion of the WebAttendant. This new product is the CCM Attendant Console. A single CCM cluster supports 96 CCM Attendant Consoles that can monitor 10,000 lines.

Because you have an existing standalone voice-mail system connected to the PBX via analog POTS connections, this system will be maintained until the PBX is completely replaced. As an interim solution, one of the Catalyst 6500 series switches could connect to ports on the voice-mail system using a WS-X6624-FXS module. You could then have CCM use those FXS ports for voice mail services using the SMDI protocol. After the PBX is replaced, you would migrate this legacy voice mail system to a Cisco Unity solution.

The existing student information system is an IVR system that allows students to access their information on the campus mainframe. The Cisco IP IVR product allows you to replace this legacy IVR system. IP IVR supports Open DataBase Connectivity (ODBC) access to databases, text to speech conversion, and a variety of other features. This product can integrate into the existing environment with a minimum of custom programming and configuration.

The existing campus modem pool is limited in speed to 33.6 kbps, because it uses analog modems. These modems can be upgraded to 56-kbps digital modems by replacing the existing Cisco 3640 router with a Cisco AS5300 equipped with MICA modems.

Remote Campus Features

Cisco.com

Cisco 3700 router:

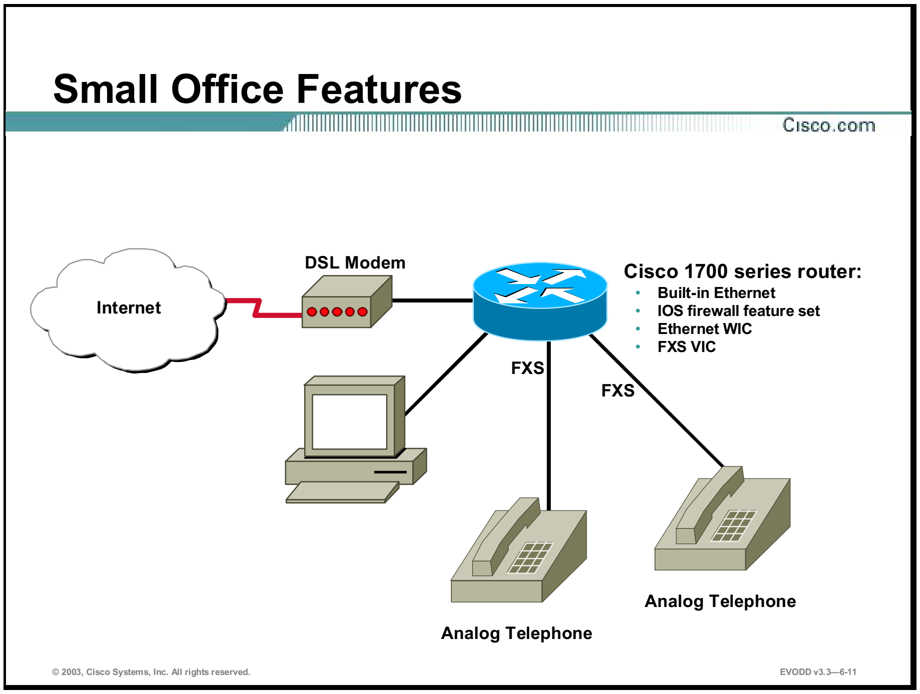
- **SRST**
- **FXS port for fax**
- **FXO ports for PSTN access**

EVODD v3.3—6-10

Currently, each of the remote campus locations has a dedicated T-1 circuit connecting back to the main campus to support its four local telephones. Additionally, there is a second T-1 for data, which is located between the main campus and each remote campus. By replacing each existing remote campus Cisco 1600 series router with a voice-enabled 3700 router, you can eliminate the voice T-1, and save the university the recurring cost of a dedicated T-1 circuit from each remote campus to the main campus.

The voice-enabled 3700 router can continue to support the dedicated data T-1 PPP connection from each remote campus location to the main campus. Additionally, Cisco IP Phones can replace the existing analog telephones by connecting into the LAN. The fax machine at each location can connect into an FXS VIC port in the 3700. In addition, the 3700 can support an FXO VIC, with two FXO ports. These FXO ports can connect to the existing PSTN trunks for local calls and PSTN fallback.

With only four telephones at the remote campus locations, a centralized CCM deployment model is used. The telephones at remote locations register with a centrally located CCM. However, if the WAN connection to the main campus is lost, the IP Phones loose connectivity with any CCM. In the event of a WAN failure, the Cisco 3700 supports SRST, which provides you with basic call-processing tasks in the absence of CCM, allowing the outgoing calls to use the PSTN connections.

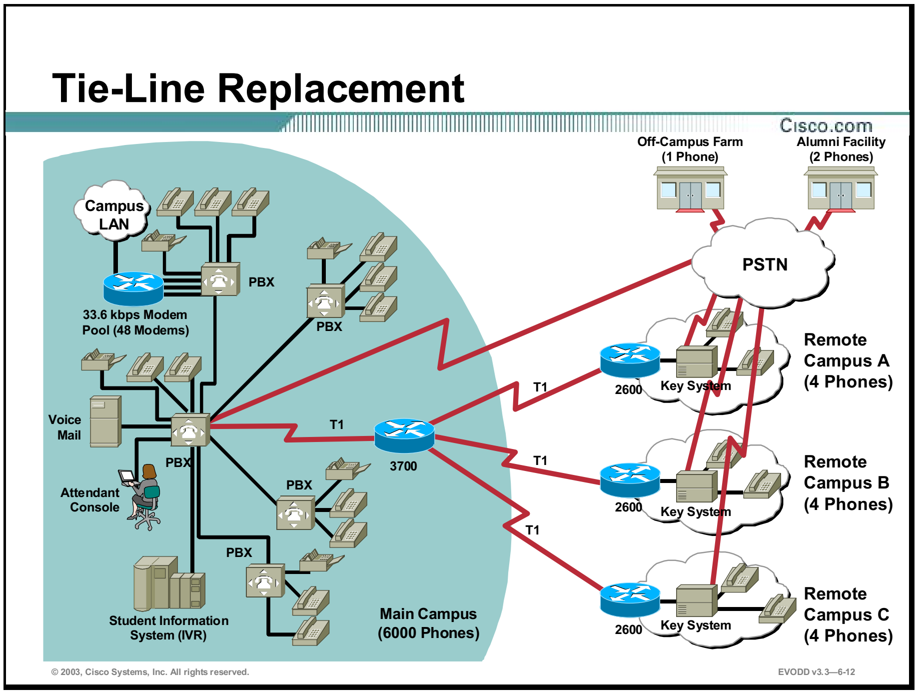


The small offices at the university have only one or two telephones. In addition, these sites use dial-up networking for data connectivity back to the campus. The recurring Internet service provider (ISP) charges and recurring telephone line charges can be eliminated and replaced with a less expensive recurring charge for DSL access.

In the new design, the DSL modem connects into an Ethernet WAN WIC inside a Cisco 1750 series router. The 1750 also has a built-in Ethernet port, which connects to a single PC at the small office. If you require additional PC connections, a low-end Catalyst switch could be installed. For voice support, an FXS VIC is inserted into one of the VWIC 1750 slots.

Migration Stages and Strategies

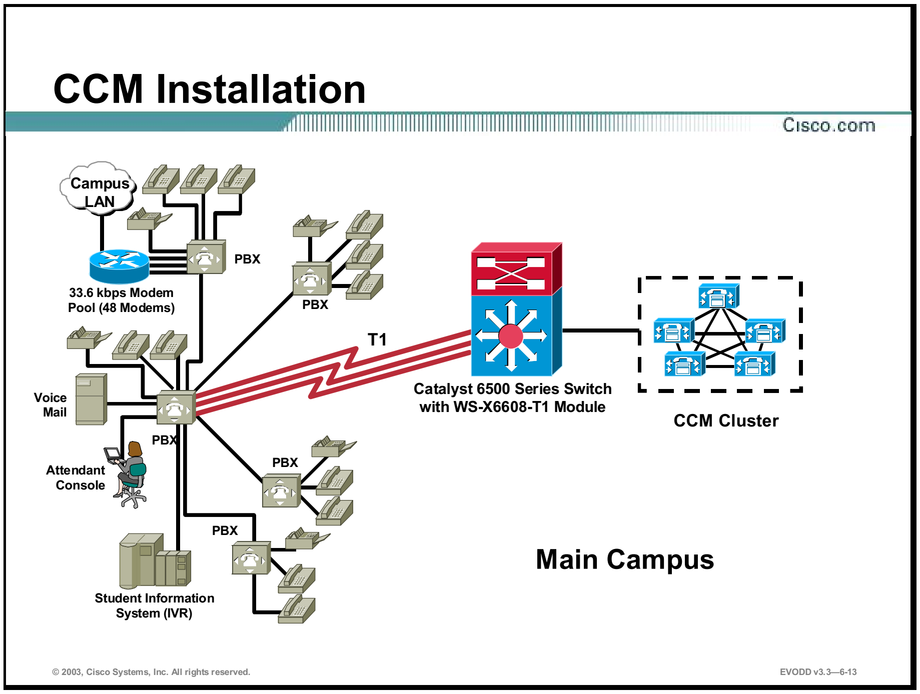
This topic selects stages of migration for the given case study. A migration on the scale of thousands of telephones typically occurs in stages.



There are two T-1 lines between each remote campus site and the main campus. One T-1 is for data traffic, and the other T-1 is for voice traffic. As an initial migration step, you can rapidly realize cost savings by eliminating one of the existing T-1 connections to each remote campus. This step involves placing voice traffic and data traffic on the same physical medium. This introduces voice quality issues, which are addressed later in this lesson.

After you complete this initial step, the PBX at the main campus connects via a T-1 into a Digital T1 Packet Voice Trunk network module. You can purchase this module and place it in a vacant slot in the existing Cisco 3700 router that services the data T-1 connections to the remote campuses. Since the PBXs intercommunicate using the Lucent Digital Crossconnect System + (DCS+) protocol, the routers can be configured to transport this proprietary signaling using frame forwarding.

Each of the remote campuses has data traffic serviced by a Cisco 1600 series router. However, as part of the migration, the 1600 series routers will be replaced with voice-enabled 3700 routers, which are equipped with FXS and FXO VICs. These VICs will be used during the latter stages of the migration to service a local fax machine and provide PSTN fallback.



When you have realized savings by replacing existing tie-lines to remote campuses, the next step is the introduction of CCM into the voice network. You can accomplish this step by attaching multiple CCMs to the existing data network at the university to form a CCM cluster. The number of CCMs selected is discussed later in this lesson.

You will need a gateway for the CCMs to communicate with the existing PBX. A Catalyst 6500 series switch with a WS-X6608-T1 module is used to connect to T1 modules in the PBX. With this gateway in place, your newly attached IP Phones will have a path to telephones attached to the PBX and to outside trunk lines.

As you introduce IP Phones to the topology, the telephones will need power and redundant LAN connections. These issues are addressed later in this lesson.

Dial Plan Migration

Cisco.com

- The PBX "First Digit Table" assigns system-wide mapping of digits to their use.

Digit	Use
0	Attendant/operator console
1	Extensions 1000-1999
2	Extensions 2000-2999
3	Extensions 3000-3999
4	Extensions 4000-4999
5	Extensions 5000-5999
6	Extensions 6000-6999
7	Features (e.g., call pickup)
8	Off-net – trunk/long distance
9	Off-net local (non-toll)

EVODD v3.3—6-14

You can implement PBX dial plans via tables *and* programmed instructions. The digit table shown here provides you with the logic for the highest level of routing decisions. These include:

- **Leading Digit 0:** A leading digit of 0 directs the caller to the university operator. During the migration from the existing PBX to the CCM, the existing PBX attendant console can coexist with a newly installed CCM 3.3 Attendant Console. Because of the proprietary nature of the existing PBX attendant console, it remains in place until completion of the migration. As you install IP Phones, the CCM is configured to direct calls with a leading digit of 0 to the CCM 3.3 Attendant Console.

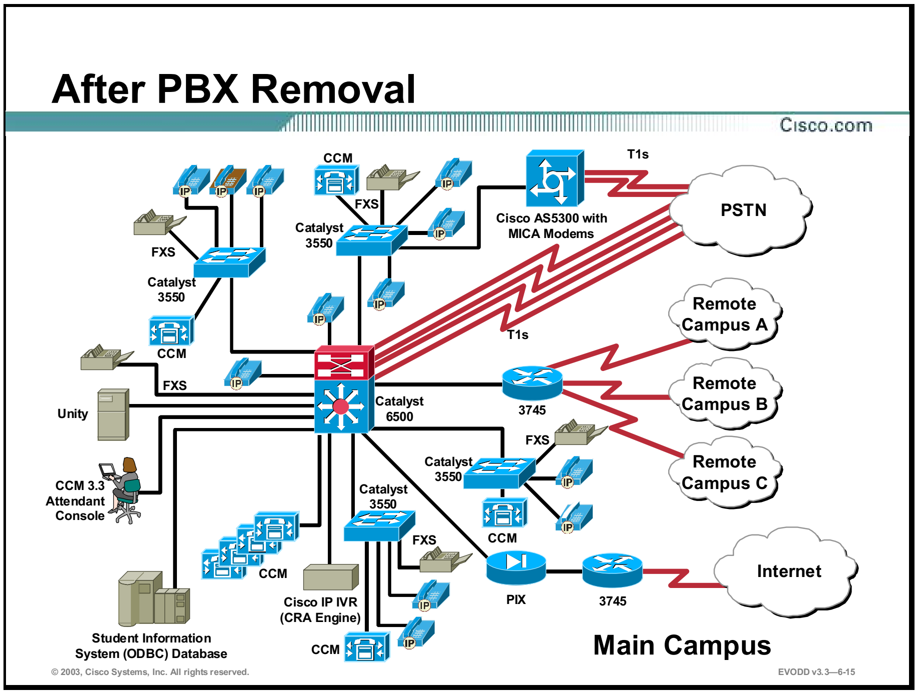
- **Leading Digits 1 – 6:** Numbers with a leading digit of 1 through 6 are extension numbers for local telephones. Cisco recommends that you migrate the local extension numbers in blocks. In this case, blocks of 1000 numbers can be migrated. For example, during the early stages of the migration, extension numbers 1000 through 1999 can be migrated at the same time. When you have verified the functionality and features of the first 1000 extensions, the next block of numbers (for example, 2000 through 2999) is migrated. This process continues until you have migrated all 6000 extension numbers from the PBX to CCM.

- **Leading Digit 7:** Currently, the PBX uses numbers with a leading digit of 7 for PBX features. For example, 72# allows you to pick up a call that is ringing in the call pickup group. Also, 75# allows you to park a call. CCM 3.3 supports these features, which can be configured to transparently match the previous dial plan.

- **Leading Digit 8:** A leading digit of 8 currently forwards callers to a long distance trunk. Therefore, after you press an 8 on your telephone keypad, you will receive a second dial tone. At that point, you can dial the desired long distance number. All university faculty, staff, and students, however, are not allowed to call long distance numbers using the university trunk. You can access 1010XXX numbers to bill calls to a personal credit card. You can configure CCM to send calls with a leading digit of 8 to a digital gateway in a Catalyst 6500 series switch that is connected to the PSTN. To restrict who can place long distance calls, use CCM features, such as calling search spaces and partitions.

- **Leading Digit 9:** A leading digit of 9 currently forwards you to local PBX trunk lines. Therefore, after you press a 9 on your telephone keypad, you receive a second dial tone. At that point, you can dial the desired local number. You can configure the CCM to send calls with a leading digit of 9 to a digital gateway in a Catalyst 6500 series switch that is connected to the local PSTN trunks.

Another dial plan you may consider is Enhanced 911 (E911) service. Because the campus is geographically dispersed, the centralized CCM cluster supports remote facilities. Therefore, you can use the Cisco Emergency Responder to provide accurate address information to emergency response personnel.



After you have completed the dial-plan migration, and the PBX has been removed, the logical topology of the telephony network is shown in the figure. You will note that the diagram is a logical topology. In addition to the specific switches and connections shown, redundant hardware and connections must also exist to provide network resiliency. Specific switch block design is discussed later in this lesson.

After you have replaced analog telephones with Cisco IP Phones, some of the analog devices—such as fax machines and speakerphones—remain, and are interconnected to the infrastructure via FXS ports in Catalyst 3550s or Catalyst 6500s.

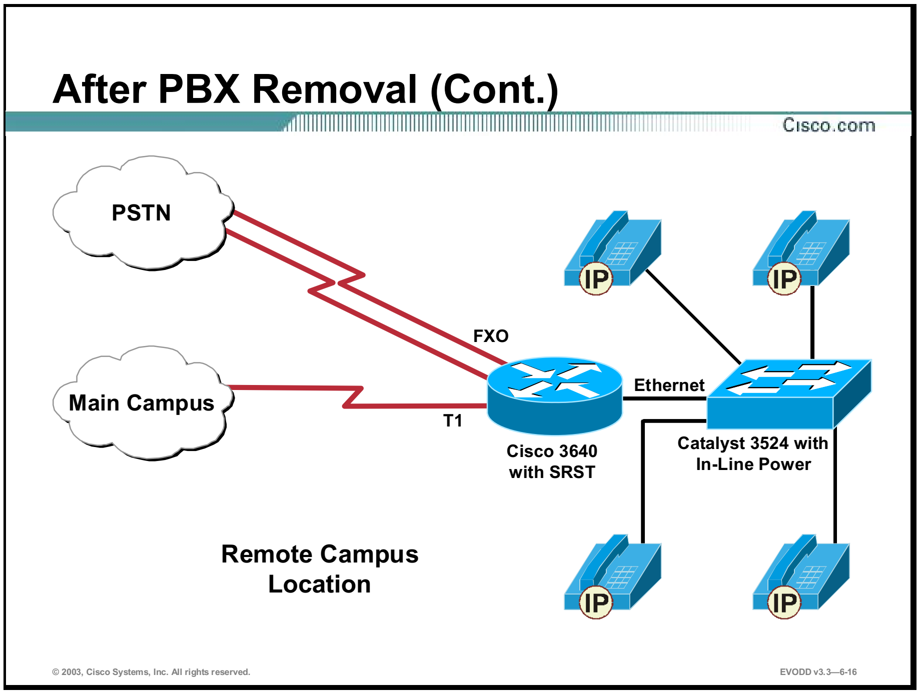
The remote campus connectivity is achieved over existing data T-1 connections to each remote site. At the main campus, the T-1s terminate in a Cisco 3745 router, which connects to a backbone Catalyst 6500 via Fast Ethernet.

The legacy voice-mail system has been replaced with a Cisco Unity messaging system. Cisco Unity provides you with a centralized message store for voice mail, fax messages, and e-mail.

The legacy attendant console, which you attached directly to the control cabinet of the PBX, is replaced with one or more instances of the CCM 3.3 Attendant Console.

The legacy interactive voice response system provided you with a mechanism for students to query the Student Information System database (that is, housing and grade queries). After the migration, this function is provided by the Cisco IP IVR system. Because the Student Information System database is ODBC-based, it can be integrated with IP IVR.

The legacy modem pool is connected at 33.6 kbps only, because it was based on analog modems in a Cisco 3640. This system is replaced with 56-kbps MICA modems in a Cisco AS5300. The AS5300 connects you directly to the PSTN via two T-1 circuits.



After you removed the PBX at the main campus, the remote campus IP Phones register with the CCM cluster at the main campus via the data T-1. You chose the centralized CCM deployment based on the small number of telephones (for example, four) at remote campus locations. The university has three such remote campus locations.

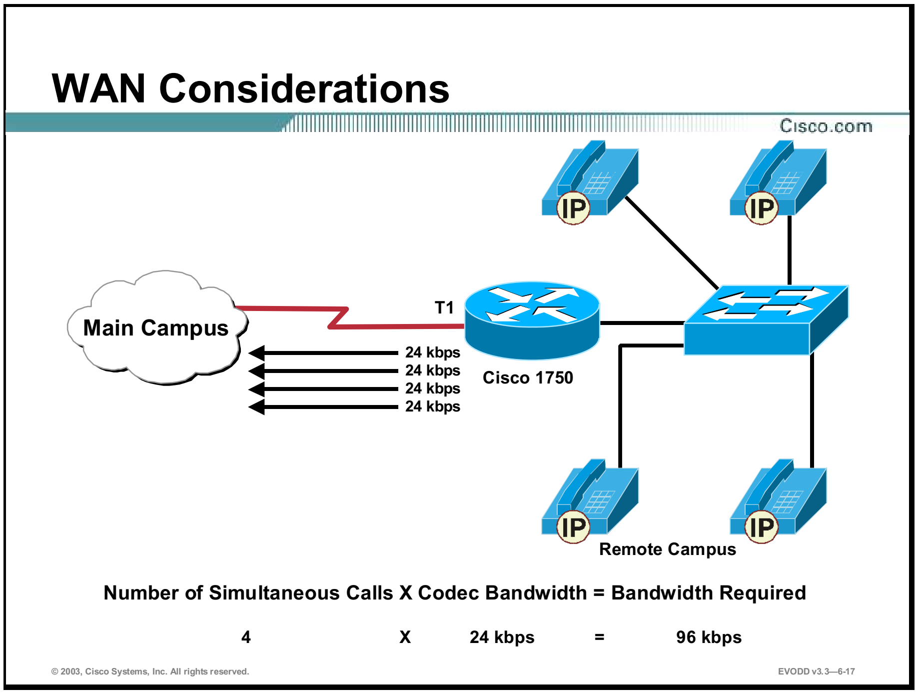
The data T-1 to each remote campus location was previously terminated in a Cisco 1600 series router. You replaced this router with a Cisco 3700 router, which supports the insertion of VICs and the SRST feature.

You also connected a Catalyst 3524 to the remote campus 3700 router, and Cisco IP Phones connect, via Fast Ethernet, to the Catalyst 3524, which provides in-line power for the IP Phones.

In the event of a WAN failure, your remote IP Phones would loose connectivity with the centralized CCM cluster. Therefore, you can configure the 3700 router to redirect off-net calls to the PSTN using a 2-port FXO VIC.

WAN Considerations

This topic addresses issues presented when accommodating voice and data traffic over a single medium. During the migration process, voice traffic is sent over the WAN link that previously transported data traffic only.



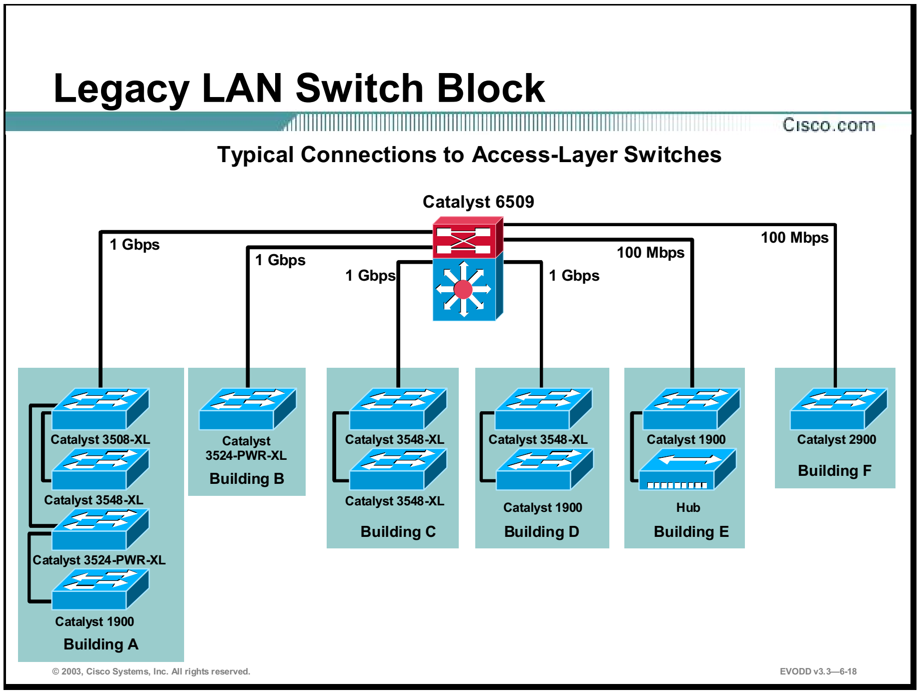
In addition to having a limited number of telephones, the remote campus locations also have a limited number of PCs. The number of PCs at the three remote campus locations range from five to ten. The bandwidth utilization on the previous 1600 series router showed that the 5-minute traffic average during peak conditions was approximately 45 percent (for example, .45 X 1.544 Mbps = 695 kbps).

Each of the remote campuses has four telephones. Because you chose second-generation Cisco IP Phones (Cisco 7960G IP Phones), G.729 was chosen as the codec to use across the WAN. When you use G.729 compression over a PPP encapsulated circuit, each call is measured, consuming approximately 24 kbps of bandwidth, including overhead. The voice bandwidth required to support four simultaneous G.729 calls is 96 kbps (for example, 4 calls X 24 kbps per call = 96 kbps).

Based on the current voice and data bandwidth demands (for example, 695 kbps [data] + 96 kbps [voice] = 791 kbps [peak]), it is determined that the WAN bandwidth capacity does not need to be increased beyond the current 1.544 Mbps T-1 circuit capacity. In the future, you can use compression mechanisms to increase effective bandwidth.

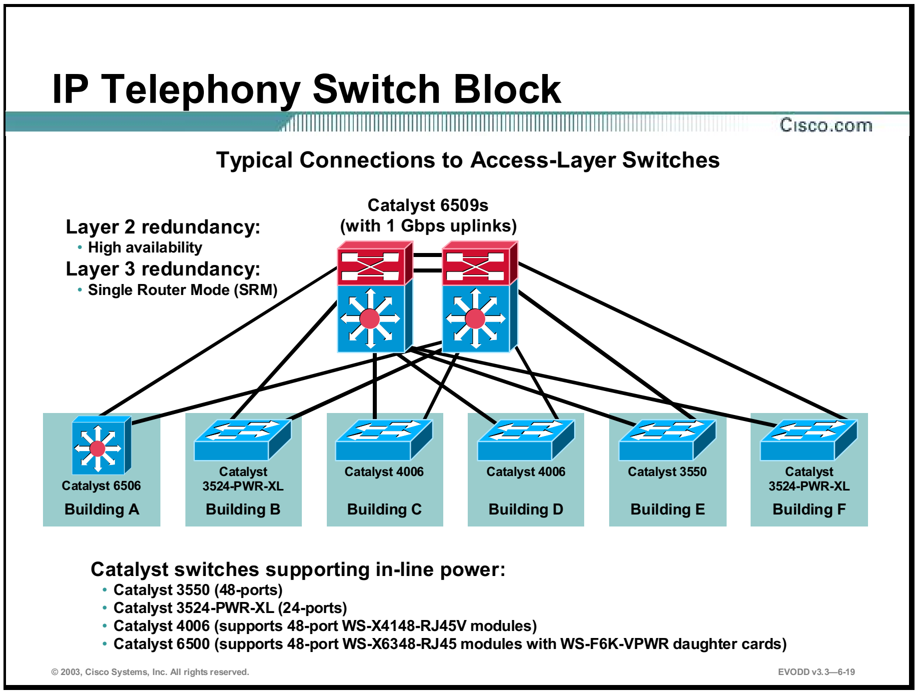
LAN Considerations

This topic reviews network resiliency technology that can be deployed in the LAN to maximize availability of the IP telephony network. Users are accustomed to high-availability of the PBX. Traditionally, PBX systems are said to have 99.999 percent availability, sometimes referred to as the five nines, which in reality is not an absolute value. An IP telephony replacement needs a similar level of availability.



The legacy data infrastructure used a collapsed core topology, where access-layer switches connected directly to the core block consisting of a single Catalyst 6509 switch. Within each of the campus buildings, access-layer switches were typically stacked, with a single link connecting back to the Catalyst 6509. Each building was in its own VLAN, with some buildings containing multiple VLANs.

Some of the access-layer switches were in-line power capable Catalyst 3524s. Other access-layer switches were a mixture of Catalyst 2900s, Catalyst 1900s, and a few shared hubs. Therefore, as part of the IP telephony migration, you must replace the hubs with switches; you will also need to upgrade any switch that does not support in-line power for IP Phones.



In the legacy LAN topology, the core Catalyst 6509 represented a single point of failure. Therefore, in the upgraded IP telephony switch block, you deployed dual 6509s at the core. Each building has dual connections back to the core. Dual connections provide you with alternate pathing in the event of a failure of one of the core Catalyst 6509s.

Redundant paths, however, do not guarantee minimum convergence time. The topology shown here requires redundancy at Layer 2 and Layer 3 with rapid convergence. A Catalyst 6509 supports dual supervisor engine modules and the High Availability feature. With High Availability enabled, one supervisor engine is active with the other in standby. The standby supervisor engine monitors Layer 2 states (for example, EtherChannel, Spanning Tree Protocol [STP] and Virtual Terminal Protocol [VTP]) of the active supervisor engine. If the active supervisor engine fails, the standby supervisor engine typically becomes active in less than three seconds.

For Layer 3 redundancy, the Catalyst 6500 supports the Single Router Mode (SRM) feature, which allows a standby Multilayer Switch Feature Card (MSFC) to mirror the configuration of the primary MSFC. SRM works with High Availability to provide you with rapid Layer 3 cutover within a chassis. For redundancy between the two chassis, you can configure Hot Standby Router Protocol (HSRP).

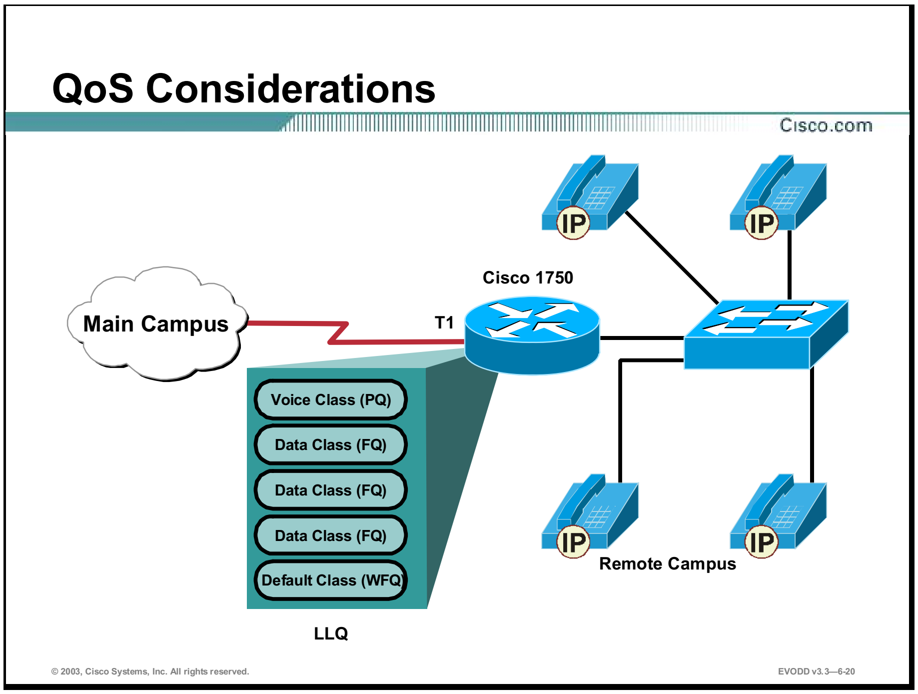
A Layer 2 network with redundant paths runs the STP to prevent issues associated with Layer 2 loops. With default timer settings, STP takes 50 seconds to converge should a link failure occur. Such a delay is intolerable for an IP telephony network. Therefore, you should deploy Cisco STP optimization features such as BackboneFast and UplinkFast. Specifically, you should configure wiring closet switches (for example, at the access layer) with the UplinkFast feature to decrease STP convergence time for directly attached failures. In addition, you should enable BackboneFast wherever possible in the network (for example, in Catalyst 4000 and 6500 series switches). BackboneFast decreases the STP convergence time for indirect failures.

Your final consideration for Catalyst switch selection is in-line power support. Currently, Cisco supports in-line power for the following Catalyst switches:

- Catalyst 3550 (48 ports)
- Catalyst 3524-PWR-XL (24 ports)
- Catalyst 4006 (Supports 48-port WS-X4148-RJ45V modules)
- Catalyst 6500 (Supports 48-port WS-X6348-RJ45 modules with WS-F6K-VPWR daughter cards)

Quality of Service Considerations

This topic identifies appropriate quality of service (QoS) tools for the case study being considered. During the migration process, voice and data begin to share the same medium. Preservation of voice quality typically requires the network to treat voice traffic better than data traffic. Cisco supports a wide range of QoS tools that identify and give preferential treatment to voice traffic.



The WAN edge is the primary network location for deployment of QoS features and is applied in both remote and core WAN routers. In the case study being considered, your WAN connections exist between each remote campus and the main campus. Since each connection is a T-1 (for example, 1.544 Mbps), and since the estimated peak traffic demand is 791 kbps, you do not have to purchase additional bandwidth. However, if periods of congestion do occur, Low Latency Queuing (LLQ) can guarantee a minimum amount of bandwidth to voice traffic and give priority treatment to voice traffic.

With LLQ, traffic classes are defined using class maps. For example, voice traffic could be classified by its IP precedence value of 5, which is automatically assigned by Cisco IP Phones. The voice class could be assigned 96 kbps of bandwidth (for example, the bandwidth required to support four simultaneous calls), and voice traffic could be directed to the priority queue. This means that when traffic exists in multiple queues, the priority queue traffic is serviced first, up to the bandwidth limit specified. If required, other data types can be assigned specific bandwidth reservations, although those queues do not receive priority treatment. Flows within a class are treated with fair queuing. The non-classified traffic is assigned to class default where flows are treated with weighted fair queuing.

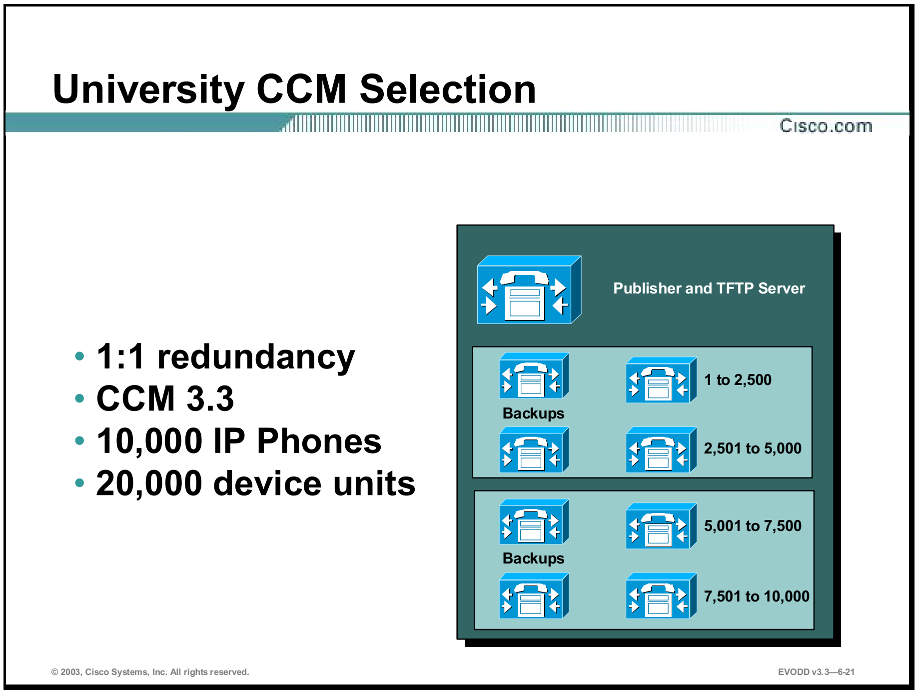
Other QoS tools for you to consider are Compressed Real-Time Transport Protocol (cRTP) and a Link Fragmentation and Interleaving (LFI) tool such as Multilink PPP (MLP). However, you should be aware that cRTP creates processing overhead on the router, and cRTP is not recommended for link speeds greater than 768 kbps.

Similarly, LFI tools are not required for link speeds greater than 768 kbps, because serialization delay is not an issue. For example, at 1.544 Mbps a 1500 byte frame exits a serial interface in approximately 7.5 ms, which is below the design serialization delay target of 10 ms. LFI tools also consume additional bandwidth on your network, because each fragment has header information added. Therefore, in this case study, LLQ is applied to the remote campus 3700 routers, main campus 3700 router (connecting to the WAN), and small office 1750 routers.

Cisco also supports QoS tools for the LAN. In this example, the main campus has two Catalyst 6500 series switches at the core. These switches can make queuing and forwarding decisions based on Class of Service (CoS) values, IP precedence values, or differentiated services code point (DSCP) values. To influence this behavior, port trust states are configured. Because Cisco IP Phones mark voice traffic with CoS and IP precedence values, and because IP precedence values can cross a router boundary, in this example, the Catalyst 6500s Ethernet ports are configured with a port trust state of *trust-ipprec*.

Cisco CallManager and Unity Scaling Considerations

This topic selects CCM and Unity platforms to support this requirement. The case study being considered requires call processing and messaging support for 6000 users.



Cisco frequently introduces new media convergence server (MCS) platforms, supporting various numbers of IP Phones. However, for this case study, the Cisco MCS-7835-1266 was chosen, which supports 2500 telephones per MCS. With CCM 3.3 installed in a 1:1 redundancy design, nine MCSs are used, as shown in the figure. The design supports over 10,000 IP Phones. Each primary CCM server has a dedicated backup server. One of the nine servers is dedicated as a TFTP server and a Structured Query Language (SQL) database publisher.

Unity Platforms

Cisco.com

Platform	Supported Unified Messaging Users
Cisco ICS7750 - SPE310 module (1 GB RAM)	500
Cisco MCS 7847	2200
Compaq ProLiant - DL380 G1/G2 - single processor	1599
Compaq ProLiant - DL380 G2 - dual processor	2200
Compaq ProLiant - ML370 G2 - dual processors	2200
Compaq ProLiant - DL580 G1 - quad processors	7500
Compaq ProLiant - DL570 G1 - dual processors	7500
Compaq ProLiant - ML570 G1 - quad processors	7500
Dell OptiPlex GX-150	499
Dell PowerEdge 1400 SC	1174
Dell PowerEdge 2500 - dual processor	2200
Dell PowerEdge 4400	6000
Dell PowerEdge 4600 - dual processor	6000
Dell PowerEdge 6400 - quad processor	7500
Dell PowerEdge 6450 - quad processor	7500
Dell PowerEdge 6600 - quad processor	7500
Dell PowerEdge 6650 - quad processor	7500
IBM x232 - single processor	1599
IBM x232 - dual processor	2200
IBM x342 - single processor	1599
IBM x342 - dual processor	2200
IBM x250 - quad processor	7500

 EVODD v3.3—6-22

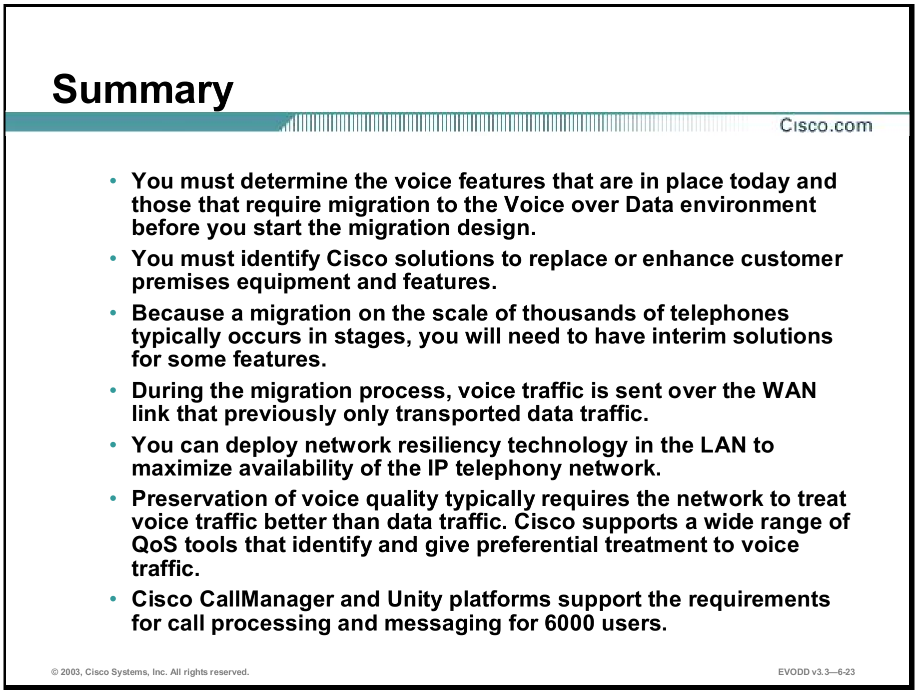
Cisco Unity provides a centralized repository for diverse message types, including voice mail, e-mail and faxes. Unity integrates with Microsoft Exchange 2000. However, Unity is supported on a specified set of platforms only, from vendors including Cisco, Compaq, Dell, and IBM. The figure shown here illustrates a subset of the supported models along with the number of unified messaging users supported by each platform. The numbers indicate the maximum number of users supported; the message store may be external to the Unity server.

Reference A more comprehensive listing of supported Unity platforms is located at: http://www.cisco.com/warp/public/cc/pd/unco/un/prodlit/ucutp_st.htm

For the case study being considered, any platform supporting 6,000 users could be chosen. However, for future expandability, you may choose a platform supporting 7,500 users instead.

Summary

This topic summarizes the key points discussed in this lesson.



Next Steps

After completing this lesson, go to:

- Your individual course curriculum for reference, and begin your next learning activity.

References

For additional information, refer to these resources:

- Cisco Emergency Responder: http://www.cisco.com/warp/public/cc/pd/unco/cer/
- Cisco IP IVR: http://www.cisco.com/warp/public/cc/pd/unco/ipivr/
- Cisco Unity: http://www.cisco.com/warp/public/cc/pd/unco/un/
- Unity Supported Platforms List: http://www.cisco.com/warp/public/cc/pd/unco/un/prodlit/ucutp_st.htm
- Survivable Remote Site Telephony: http://www.cisco.com/warp/public/cc/pd/unco/srstl/index.shtml

- Cisco Media Convergence Servers: http://www.cisco.com/warp/public/cc/pd/mxsv/mxcvsr/
- Cisco Unity Customer Profiles: http://www.cisco.com/warp/public/cc/pd/unco/un/profiles/index.shtml

Lesson Review

This practice exercise reviews what you have learned in this lesson.

Q1) What type of PBX model uses multiple PBX cabinets in separate physical locations with a centralized control cabinet?

A) distributed control model

B) dispersed control model

C) centralized control model

D) adjunct control model

Q2) What replacement does Cisco offer for a PBX operator console?

A) CCM 3.3 Attendant Console

B) CCM 3.3 WebAttendant

C) CCM 3.3 IP IVR

D) CCM 3.3 Unity

Q3) Which of the following migration states occur first?

A) CCM installation

B) dial plan migration

C) tie-line replacement

D) PBX removal

Q4) What formula determines the amount of WAN bandwidth required for voice calls between two locations?

A) Bandwidth Required = Number of Simultaneous Calls X CIR

B) Bandwidth Required = Number of Simultaneous Calls X Codec Bandwidth

C) Bandwidth Required = Number of Simultaneous Calls X Codec Bandwidth/cRTP Header Size

D) Bandwidth Required = Number of Simultaneous Calls X BHCA Erlangs

Q5) Which two of the following are Catalyst switch models that support in-line power for Cisco IP Phones? (Choose two.)

A) Catalyst 3524-PWR-XL

B) Catalyst 2924-PWR-XL

C) Catalyst 4006

D) Catalyst 1900

Q6) Which two of the following statements are reasons to decide against using LFI techniques on a T-1 circuit? (Choose two.)

A) increased router processor overhead

B) increased bandwidth demands due to the addition of header information for each fragment

C) serialization delay for T-1 circuits

D) DTMF signaling cannot be reliably transported when using LFI techniques

Q7) Using CCM 3.3 and a 1:1 redundancy model, how many Cisco MCS-7835-1266 servers are recommended to support 10,000 users?

A) 4

B) 5

C) 8

D) 9

A

Answers to Review Questions

The lesson review items and solutions are contained here.

Module 1: Voice Over Data Overview

Lesson One: Enterprise and Public Switched Telephone Networks

Lesson Review

Q1) What is the role of toll bypass?

A) to allow telephone calls over the IP WAN if the PSTN becomes congested

B) to allow telephone calls over the IP WAN to avoid PSTN charges

C) to allow telephone calls over the PSTN if the IP WAN becomes congested

D) to allow telephone calls over the PSTN to avoid IP WAN charges

Answer: B

Q2) Which of the following is NOT considered Voice over Data building block?

A) PBX

B) PSTN

C) Unity

D) QoS

Answer: C

Q3) Which of the following best describes an OPX?

A) An OPX is an on-premises extension located at the main office.

B) An OPX is an off-premises extension located at a branch or home office.

C) An OPX is an outgoing PBX extension located at the main office.

D) An OPX is an outgoing PBX extension located at the branch or home office.

Answer: B

Q4) After the route replacement phase, which of the following would be a logical next step in a migration from a legacy PBX system to an IP telephony solution?

A) removing the legacy PBXs

B) connecting the digital telephones of the PBX to a Catalyst switch

C) installing CCM for toll bypass

D) connecting IP Phones to the PBX

Answer: C

Q5) Which of the following Voice over Data technologies associates a destination telephone number with a local DLCI?

A) VoIP

B) VoATM

C) VoHDLC

D) VoFR

Answer: D

Lesson Two: Voice Over Data Network

Lesson Review

Q1) What is the correct order for the Voice over Data deployment stages?

A) Implement, Operate, Design, and Plan

B) Plan, Design, Operate, and Implement

C) Design, Plan, Implement, and Operate

D) Plan, Design, Implement, and Operate

Answer: D

Q2) Which two of the following factors might be motivators for a customer who decides on a Voice over Data solution? (Choose two.)

A) cost savings

B) an opportunity to use the existing digital PBX telephones for the customer

C) IP telephony applications

D) an opportunity to increase departmental staffing

Answers: A and C

Q3) What is the Cisco recommendation regarding connecting IP Phones to a LAN?

A) use no more than three IP Phones per hub

B) connect IP Phones to switches

C) connect IP Phones to routers

D) connect IP Phones to gateways

Answer: B

Q4) What is an Erlang?

A) a compression mechanism for T.120 white board collaboration

B) a coder-decoder (codec) for very low-speed WAN links

C) an H.323 measure of quality

D) 3600 seconds of calls on the same circuit

Answer: D

Q5) Which three of the following costs are considered current costs for the customer? (Choose three.)

A) LEC

B) backup power

C) staffing

D) average busy hour call volume

Answers: A, B, and C

Q6) Which two of the following switch models would typically be upgraded during a Voice over Data migration? (Choose two.)

A) Cisco Catalyst 1900

B) Cisco Catalyst 3500

C) Cisco Catalyst 5500

D) Cisco Catalyst 6500

Answers: A and C

Q7) Which four of the following costs should be considered when calculating staff costs? (Choose four.)

A) internal voice staff costs

B) internal data staff costs

C) contracted staff

D) outsourced staff

E) telephone provider staff costs

Answer: A, B, C, and D

Q8) What is the default time that the Cisco Networking Feasibility Tool uses when estimating a project life?

A) 2 years

B) 3 years

C) 4 years

D) 5 years

Answer: B

Q9) What does the voice sensitivity analysis show?

A) The average daily call minutes required to achieve savings

B) The ROI for the project

C) The IRR for the project

D) The NPV for the project

Answer: A

Module 2: Standards for Voice Over Data Signaling

Lesson One: VoIP Gateway Protocols

Lesson Review

Q1) Which two of the following signaling protocols are considered peer-to-peer protocols? (Choose two.)

A) H.323/RAS

B) Simple Client Control Protocol

C) SIP

D) MGCP

Answers: A and C

Q2) What is the current version of H.323?

A) 2

B) 3

C) 4

D) 5

Answer: C

Q3) In the context of the H.323 standard, which of the following components is considered mandatory?

A) terminals

B) gateways

C) gatekeepers

D) MCUs

Answer: A

Q4) Which codec must be supported by H.323 terminals?

A) G.729

B) G.711

C) G.729 Annex A

D) G.723

Answer: B

Q5) Which of the following statements best describes registration, admission, and status (RAS)?

A) RAS is a protocol that runs between endpoints (terminals and gateways) and gatekeepers.

B) RAS is a protocol that runs between Cisco IP Phones and other IP Phones.

C) RAS is a protocol that runs between terminals and gateways. RAS is a protocol that runs between Cisco IP Phones and CCMs.

Answer: A

Q6) Which of the following protocols is used to set up connections between H.323 endpoints, over which real-time data can be transported?

A) RTCP

B) RTP

C) H.245

D) H.225

Answer: D

Q7) Which protocol does Cisco AVVID use between Cisco IP Phones and CCMs?

A) H.245

B) SSP

C) RTCP

D) H.225

Answer: B

Q8) Which of the following statements are true concerning MGCP?

A) MGCP uses a large set of complex transactions for maximum vendor interoperability.

B) MGCP uses TCP, instead of UDP, because TCP is a reliable transport protocol.

C) Endpoints can be mass-produced cheaply.

Answer: C

Q9) Which three of the following components belong to SIP? (Choose three.)

A) User agent client

B) User server client

C) SIP client

D) SIP server

Answers: A, C, and D

Lesson Two: Voice over Frame Relay Design Consideration

Lesson Review

Q1) Which two of the following benefits apply to VoFR? (Choose two.)

A) integrates voice and data in a homogeneous topology

B) enables managed network services

C) adds the performance benefits of tandem switching

D) utilizes Frame Relays built-in QoS service classes

Answers: A and B

Q2) What type of VoFR trunk is used for tie-line replacement between PBXs?

A) FRF.8

B) FRF 3.1

C) FRF.11

D) FRF.12

Answer: C

Q3) What encapsulation method is used to transport data across a Frame Relay network?

A) FRF.11

B) FRF.12

C) FRF.5

D) FRF 3.1

Answer: D

Q4) How many subchannels does FRF.11 support on a single DLCI?

A) 1

B) 23

C) 127

D) 255

Answer: D

Q5) Which two of the following applications would be used for a Cisco trunk FRF.11 connection with tandem avoidance? (Choose two.)

A) OPX extension

B) interoperability with third-party FRADs

C) PBX-to-PBX connection

D) End-to-end VoFR (such as, when users specify dialed digits)

Answers: A and C

Q6) In a hub and spoke Frame Relay network, where some routers carry VoIPoFR and data traffic, and some routers carry only data, where should FRF.12 be implemented?

A) only on the hub router

B) only on routers carrying voice traffic

C) on all spoke routers

D) on all routers

Answer: D

Q7) Approximately how much bandwidth is required for a VoFR call using the G.729 codec (assuming 50 voice samples per second, a 2 byte FRF.11 header, and a 1 byte sequence number)?

A) 26.4 kbps

B) 8 kbps

C) 32.6 kbps

D) 10.8 kbps

Answer: D

Module 3: Cisco CallManager Overview and Gateway Selection

Lesson One: Cisco CallManager Architecture

Lesson Review

Q1) Which of the following components are "Enterprise Cisco IP telephony Building Blocks?" (Choose three.)

A) PBX migration strategies

B) Properly provisioned QoS-enabled infrastructure

C) Unified messaging

D) Fax relay

Answers: A, B, and C

Q2) What database stores a Cisco IP telephony network configuration database?

A) IMAP

B) SQL

C) LDAP

D) DC Directory

Answer: B

Q3) What number of IP Phones on a network indicates that you should add a separate CCM to act as a SQL publisher and TFTP server?

A) 250

B) 500

C) 1000

D) 2500

Answer: C

Q4) What is the weight (per session per voice channel) for an H.323 client?

A) 1

B) 3

C) 5

D) 20

Answer: B

Q5) What is the protocol for signaling between CCM clusters?

A) H.323

B) T.120

C) RTCP

D) QSIG

Answer: A

Q6) Which of the following statements applies to transcoding via DSPs?

A) Transcoding is only supported between a high bit rate codec and a high bit rate codec.

B) Transcoding is only supported between a low bit rate codec and a low bit rate codec.

C) Transcoding is only supported between a high bit rate codec and a low bit rate codec.

D) All of the above are valid transcoding actions.

Answer: C

Q7) Which two of the following statements apply to location-based CAC? (Choose two.)

A) IP bandwidth limits are assigned per location (in kbps).

B) If insufficient bandwidth is available, the call attempt returns to the PSTN.

C) Cisco recommends using a single codec for all calls crossing the WAN.

D) Probes are sent to determine conditions within the network.

Answers: A and C

Q8) Which of the following statements best describes an isolated CCM deployment?

A) CCM are located at a single site.

B) CCM are located at multiple sites and are interconnected via the IP WAN.

C) CCM are located at multiple sites but are not interconnected.

D) None of the above.

Answer: C

Lesson Two: Gateway Types

Lesson Review

Q1) What is the function of DTMF relay?

A) DTMF relay converts pulse-dialed digits into tone-dialed digits.

B) DTMF relay preserves dialed digit information with low-compression codecs.

C) DTMF relay preserves dialed digit information with high-compression codecs.

D) DTMF relay uses TCP, instead of UDP, for tone signaling.

Answer: C

Q2) Which three features belong to Cisco H.323v2 gateways? (Choose three.)

A) DTMF relay

B) Supplementary services

C) CCM failover

A) Analog and digital DID

Answers: A, B, and C

Q3) Which statement best describes a Cisco VG248?

A) A digital gateway with two T1 interfaces

B) A 48-port analog gateway

C) An ISDN gateway with two PRIs

D) A SIP gateway module for the Cisco Catalyst 6500

Answer: B

Q4) Which gateway allows you to connect analog telephones to a Catalyst 6500 series switch?

A) WS-X6624-FXO

B) WS-X6608-T1

C) WS-X6624-FXS

D) WS-X6607-E1

Answer: C

Q5) What technology is designed for installing IP Phones in a small office without a CCM?

A) Cisco IOS Telephony Service

B) SRST

C) Cisco IVR

D) Cisco Unity

Answer: A

Q6) Which gateway does not support FXS interfaces?

A) Cisco Catalyst 6500

B) Cisco 1750

C) Cisco AS5300

D) Cisco 2600

Answer: C

Q7) You need to attach a remote location to the IP telephony network in the main office. You will NOT convert the telephones in the remote office from analog to IP Phones. This site must have the largest feature set possible. Which device would be best suited for this situation?

A) VG200

B) VG248

C) Catalyst 5500

D) Catalyst 6500

Answer: B

Module 4: Dial Plans and Voice-Mail Considerations

Lesson One: Signaling, Networking, and Transporting

Lesson Review

Q1) Which of the following best describes the Cisco Transport model, for forwarding PBX signaling?

A) The Transport model is used when the Cisco router can interpret the PBX signaling, such as QSIG.

B) The Transport model is used when the Cisco router cannot interpret the PBX signaling, such as DCS+.

C) The Transport model is only used when the PBX uses a single CCS channel.

D) The Transport model is only used when the PBX uses multiple CCS channels.

Answer: B

Q2) Which of the following is required to transport proprietary PBX signaling using multiple signaling channels?

A) Cross-connect

B) Frame forwarding

C) CES

D) STUN

Answer: A

Q3) Which of the following is a PBX interconnection requiring a call to pass through one or more intermediate PBXs on the way from the source PBX to the destination PBX?

A) Tunnel PBX

B) CES PBX

C) Tandem PBX

D) CCS PBX

Answer: C

Q4) Which of the following signaling protocols can use the Cisco Translate model to forward PBX signaling?

A) DCS

B) DCS+

C) MCDN

D) QSIG

Answer: D

Q5) Which proprietary protocol CANNOT use frame forwarding?

A) DCS

B) DCS+

C) MCDN

D) ISL

Answer: A

Lesson Two: Dial Plans

Lesson Review

Q1) A dial plan is analogous to which of these items?

A) router routing table

B) access-list

C) subnet

D) CoS

Answer: A

Q2) Which of these options best describes a partition?

A) partitions define which users can access which calling search spaces

B) partitions are CoS groups

C) partitions are groups of devices with similar accessibility

D) partitions are subnets

Answer: C

Q3) Which of the following do CCMs use to address problems with overlapping dial plans?

A) partitions

B) calling search spaces

C) CoS restrictions

D) translation patterns

Answer: D

Q4) Which three of the following are considered PBX to CCM migration issues? (Choose three.)

A) cost

B) ROI

C) scalability

D) security and regulatory impact

Answers: A, C, and D

Q5) Which of these E911 features provide location information to a 911 operator?

A) ANI

B) ALI

C) PSAP

D) MLTS

Answer: B

Q6) Which Cisco product addresses E911 concerns?

A) NetRanger

B) Emergency Responder

C) PSAP-IP

D) SRST

Answer: B

Q7) What is the only standards-based signaling protocol used between a PBX and a voice-mail system?

A) H.323

B) QSIG

C) Q.931

D) SMDI

Answer: D

Q8) What are the three components of the Cisco unified messaging solution? (Choose three.)

A) Cisco Unity server

B) directory server

C) CCM server

D) message store

Answers: A, B, and D

Module 5: Voice Over Data Characteristics

Lesson One: Delay Characteristics

Lesson Review

Q1) Echo becomes a significant problem when the round-trip delay becomes greater than what amount of time?

A) 10 ms

B) 50 ms

C) 125 ms

D) 150 ms

Answer: B

Q2) Which three of the following components are fixed delay components? (Choose three.)

A) propagation delay

B) queuing delay

C) serialization delay

D) processing delay

E) packetization delay

Answers: A, C, and D

Q3) Which two of the following QoS mechanisms can be used to influence queuing delay? (Choose two.)

A) CBWFQ

B) WRED

C) IP RTP Priority

D) CRTP

Answers: A and C

Q4) What is the ITU-T's G.114 recommendation for maximum one-way delay?

A) 10 ms

B) 20 ms

C) 125 ms

D) 150 ms

Answer: D

Q5) What effect does tandem encodings have on a MOS value?

A) It increases the MOS value.

B) It decreases the MOS value.

C) It has no effect on the MOS value.

D) Tandem switching uses a different MOS scale (such as, 1-10).

Answer: B

Lesson Two: Compression Technologies and Packet Compensation

Lesson Review

Q1) Which three of the following types are voice coders? (Choose three.)

A) waveform coders

B) vocoders

C) hybrid coders

D) linear coders

Answers: A, B, and C

Q2) G.711 is an example of which voice coder type?

A) waveform

B) vocoder

C) hybrid

D) linear

Answer: A

Q3) G.729 is an example of which voice coder type?

A) waveform

B) vocoder

C) hybrid

D) linear

Answer: C

Q4) Which of the following codecs require 20 MIPS of CPU processing power?

A) G.729

B) G.729A

C) G.729B

D) G.723.1

Answer: A

Q5) Background noise varies based on location. Which of the following is synthetically generated noise approximating the appropriate background noise for a conversation?

A) white noise

B) purple noise

C) red noise

D) pink noise

Answer: D

Q6) Which three of the following methods accommodate a lost packet? (Choose three.)

A) expand the playout time of the previous packet

B) interpolate the packet

C) replay the previous packet

D) play noise

Answers: B, C, and D

Q7) What is the primary cause of echo?

A) a dB level that is too high on the transmit side

B) a dB level that is too high on the receive side

C) an impedance mismatch in 2-wire to 4-wire hybrid circuits

A) EMI

Answer: C

Q8) What is the standard analog fax protocol?

A) T.120

B) T.30

C) FRF.11

D) T.38

Answer: B

Module 6: Voice Over Data Migration

Lesson One: Voice Quality Overview

Lesson Review

Q1) What is the queuing strategy that Cisco recommends for voice traffic?

A) WFQ

B) CBWFQ

C) IP RTP priority

D) LLQ

Answer: D

Q2) Which two of the following are marking types? (Choose two)

A) IP precedence

B) WFQ

C) LLQ

D) DSCP

Answers: A and D

Q3) How many ToS bits are used by DSCP?

A) 3

B) 6

C) 8

D) 16

Answer: B

Q4) What is the serialization delay for a 1,500-byte frame exiting a serial interface running at 56 kbps?

A) 10 ms

B) 20 ms

C) 150 ms

D) 214 ms

Answer: D

Q5) When configuring traffic shaping, what is the timing interval if the CIR = 64 kbps and the Bc = 8,000 bits?

A) 20 ms

B) 125 ms

C) 150 ms

D) 214 ms

Answer: B

Q6) What are the two steps for campus QoS implementation? (Choose two.)

A) Classification

B) Marking

C) Queuing

D) Policing

Answers: A and C

Q7) A Cisco IP Phone marks voice frames with what CoS value?

A) 2

B) 5

C) 7

D) 64

Answer: B

Q8) What type of admission control tools uses SAA probes?

A) local

B) measurement-based

C) end-to-end

D) remote

Answer: B

Q9) What LFI tool would be appropriate for a PPP network?

A) FRF.11 Annex C

B) FRF.12

C) MLP

D) CRTP

Answer: C

Lesson Two: Comprehensive Design Strategies

Lesson Review

Q1) What type of PBX model uses multiple PBX cabinets in separate physical locations with a centralized control cabinet?

A) distributed control model

B) dispersed control model

C) centralized control model

D) adjunct control model

Answer: B

Q2) What replacement does Cisco offer for a PBX operator console?

A) CCM 3.3 Attendant Console

B) CCM 3.3 WebAttendant

C) CCM 3.3 IP IVR

D) CCM 3.3 Unity

Answer: A

Q3) Which of the following migration states occur first?

A) CCM installation

B) dial plan migration

C) tie-line replacement

D) PBX removal

Answer: C

Q4) What formula determines the amount of WAN bandwidth required for voice calls between two locations?

A) Bandwidth Required = Number of Simultaneous Calls X CIR

B) Bandwidth Required = Number of Simultaneous Calls X Codec Bandwidth

C) Bandwidth Required = Number of Simultaneous Calls X Codec Bandwidth/cRTP Header Size

D) Bandwidth Required = Number of Simultaneous Calls X BHCA Erlangs

Answer: B

Q5) Which two of the following are Catalyst switch models that support in-line power for Cisco IP Phones? (Choose two.)

A) Catalyst 3524-PWR-XL

B) Catalyst 2924-PWR-XL

C) Catalyst 4006

D) Catalyst 1900

Answers: A, and C,

Q6) Which two of the following statements are reasons to decide against using LFI techniques on a T-1 circuit? (Choose two.)

A) increased router processor overhead

B) increased bandwidth demands due to the addition of header information for each fragment

C) serialization delay for T-1 circuits

D) DTMF signaling cannot be reliably transported when using LFI techniques

Answers: B and C

Q7) Using CCM 3.3 and a 1:1 redundancy model, how many Cisco MCS-7835-1266 servers are recommended to support 10,000 users?

A) 4

B) 5

C) 8

D) 9

Answer: D

B

Course Glossary

Acronym or Term	Expansion of Acronym
AbS	Analysis by Synthesis
AbS	Analysis-by-Synthesis coding
ACELP	Algebraic Code Excited Linear Prediction
ACF	Admission Confirmation
ACL	Access Control List
ADPCM	adaptive differential pulse code modulation
AHT	average handle time
AIS	alarm indication signal
AMIS	Audio Messaging Interchange Specification
ANI	automatic number identification
API	application programming interface
ARJ	Admission Reject
ARQ	Admission Request
ASP	application service provider
AVBO	advanced voice busyout
AVVID	Cisco Architecture for Voice, Video and Integrated Data
Bc	committed burst
BCF	Bandwidth Confirm
BE	border elements
BECN	backward explicit congestion notification
Bps	bits per second
BRJ	Bandwidth Reject
BRQ	Bandwidth Request
CAC	Call Admission Control
CAMA	Centralized Automated Messaging Accounting
CAS	channel associated signaling
CBWFQ	class-based weighted fair queuing
CCM	Cisco CallManager
CCS	common channel signaling
CDP	Coordinated Dialing Plan
CEF	Cisco Express Forwarding
CELP	code excited linear prediction compression
CES	Circuit emulation service
CFB	Call Forwarding-Busy
CID	channel ID
CIPT	Cisco IP telephony
CIR	committed information rate

Acronym or Term	Expansion of Acronym
CLEC	competitive local exchange carrier
CLI	command-line interface
CLID	calling line identification
CO	central office
codec	coder/decoder
CoS	class of service
CPE	customer premises equipment
CRC	cyclic redundancy check
cRTP	Compressed Real-Time Transport Protocol
CS-ACELP	Conjugate Structure Algebraic Code Excited Linear Prediction
DCS	Lucent Distributed Communications System
DCS+	Lucent Distributed Communications System
DID	Direct Inward Dialing
DLCI	data-link connection identifier
DN	Directory Number
DOD	direct outward dialing
DSO	digital service zero
DS-0	digital signal level 0
DS-1	digital signal level 1
DS-3	digital signal level 3
DSCP	differentiated services code point
DSI	digital speech interpolation
DSL	digital subscriber line
DSP	digital signal processor
DTMF	dual tone multifrequency
DVM	digital voice module
E & M	recEive and transMit
EF	Expedited Forwarding
EIGRP	Enhanced Interior Gateway Routing Protocol
ETSI	European Telecommunication Standards Institute
EVODD	Cisco Enterprise Voice over Data Design
FB	forward-busy
FNA	Forward-No Answer
FRAD	Frame Relay Access Device
FRF.11	Frame Relay Forum implementation agreement for VoFR
FRF.12	Frame Relay Forum implementation agreement for VoIPoFR
FRF.8	Frame Relay-to-ATM Service Internetworking

Acronym or Term	Expansion of Acronym
FTE	FAX Terminal Equipment
FXO	Foreign Exchange Office
FXS	Foreign Exchange Station
GCF	Gatekeeper Confirmation
GRQ	Gatekeeper Request
GUI	graphical user interface
HDLC	High-Level Data Link Control
HSRP	Hot Standby Router Protocol
HSSI	High-Speed Serial Interface
IETF	Internet Engineering Task Force
IPDC	Internet Protocol Device Control
IRQs	interrupt requests
ISL	Inter-Switch Link
ISO	International Organization for Standardization
ISP	Internet service provider
ITSP	Internet Telephony Service Provider
ITU	International Telecommunication Union
ITU-T	International Telecommunication Union-Telecommunication Standardization Sector
IVR	interactive voice response
IXC	inter-exchange carrier
JTAPI	Java Telephony Application Programming Interface
LCR	Least Cost Routing
LDCELP	low-delay CELP
LEC	local exchange carrier
DFI	Link Fragmentation and Interleaving
LLQ	low latency queuing
LPC	linear predictive coding
MAN	metropolitan-area network
MCDN	Meridian Customer Defined Network
MCS	Media Convergence Server
MCU	multipoint control unit
MFT	multiflex trunk module
MGC	Media Gateway Controller
MGCP	Media Gateway Control Protocol
MHSRP	Multigroup Hot Standby Router Protocol
MIME	Multipurpose Internet Mail Extensions
MIPS	millions of instructions per second

Acronym or Term	Expansion of Acronym
MLP	Multilink PPP
MOH	music on hold
MOS	mean opinion score
MPLS	Multiprotocol Label Switching
MPLS VPN	Multiprotocol Label Switching Virtual Private Network
MSDL	Multirate Subscriber Digital Line
MS-GSM	Microsoft-Global System Mobile
MTP	media termination point
MTU	maximum transmission unit
MWI	Message Waiting Indicator
NPA	Numbering Plan Area
NPV	net present value
NTP	Network Time Protocol
OAM	Operation, Administration, and Maintenance
OC-3	Optical Carrier 3
OOS	Out-of-Service
OPX	Off-Premises eXtension
OSI	Open System Interconnection protocol stack
OSI model	Open System Interconnection reference model
OSPF	Open Shortest Path First
PCM	pulse code modulation
PFC	Policy Feature Card
PHB	per-hop behavior
PIR	peak information rate
PLAR	private line, automatic ringdown
POTS	plain old telephone service
pps	packets per second
PQ	priority queue
PQ/CBWFQ	priority queuing/class-based weighted fair queuing
PSAP	Public Safety Answering Point
PSTN	Public Switched Telephone Network
PTNX	Private Telco Network Exchange
PVC	permanent virtual circuit
QoS	quality of service
QPM	QoS Policy Manager
QSIG	Q Signaling
RAS protocol	registration, admission, and status protocol

Acronym or Term	Expansion of Acronym
RBOC	regional Bell operating company
RIP	Routing Information Protocol
ROI	return on investment
RSP	Route/Switch Processor
RSVP	Resource Reservation Protocol
RTCP	Real-Time Transport Control Protocol
RTP	Real-Time Conference Protocol
RTP	Real-Time Transport Protocol
SAA	Service Assurance Agent
SAR	segmentation and reassembly
SCCP	Signaling Connection Control Port
SCN	Switched Circuit Network
SCR	sustainable cell rate
SGCP	Simple Gateway Control Protocol
SID	Silence Insertion Descriptors
SIP	session initiation protocol
SLA	service level agreement
SMDI	Simplified Message Desk Interface
SMDS	Switched Multimegabit Data Service
SMTP	Simple Mail Transfer Protocol
SNA	Systems Network Architecture
SQL	Microsoft Structured Query Language
SRST	Survivable Remote Site Telephony
SSP	Skinny Station Protocol
STUN	Serial Tunneling
SVC	switched virtual circuit
SVCC	Switched virtual channel connection
TAPI	Telephony Application Programming Interface
TCP	Transmission Control Protocol
TDM	time-division multiplexing
ToS	type of service
TTL	Time to Live
UA	User Agent
UAC	user agent client
UDP	User Datagram Protocol
UNI	User-Network Interface
UPS	uninterruptible power supply

Acronym or Term	Expansion of Acronym
VAD	voice activity detection
VCI	virtual channel identifier
VCW	voice card ware
VIC	voice interface card
VoATM	Voice over ATM
VoFR	Voice over Frame Relay
VoHDLC	Voice over HDLC
VoIP	Voice over IP
VoIPoATM	Voice over IP over ATM
VoIPoFR	Voice over IP over Frame Relay
VPN	Virtual Private Network
VWIC	voice/WAN interface card
WFQ	weighted fair queuing
WIC	WAN interface card
WRED	weighted random early detection
WRR	Weighted Round Robin

C

Answers to Laboratory Exercises

The Laboratory Exercise questions and solutions are contained here.

Module 1: Voice Over Data Overview

Laboratory Exercise: Calculating Trunk Capacity

Questions and Answers:

Assuming with 18 trunks carrying 9 Erlangs of traffic with an average of 3 minutes determine the following:

Q1) What is the average number of busy trunks?

Answer: With 9 Erlangs of traffic, 9 trunks will be busy since an Erlang is the amount of traffic one trunk can handle in one hour.

Q2) What is the number of calls that can originate in an hour?

Answer: 9 in one hour *60 minutes/hour divided by 3 minutes/call= 180

Q3) What is the time it takes to complete all the calls?

Answer: 180 calls lasting 3 minutes per call, the total is 540 minutes or 9 hours

Module 2: Standards for Voice Over Data Signaling

Laboratory Exercise: Voice over Frame Relay

Questions and Answers:

Q1) Both voice and data will use the Frame Relay ports in the diagram. What protocol should be implemented to keep data from placing too much delay on the voice?

Answer: FRF.11 or FRF.12

Q2) What link speed would provide enough bandwidth that data would not have to be fragmented?

Answer: 1.544 Mbps

Q3) Which ports should FRF.12 be implemented on?

Answer: Ports A, B, C, and D

Q4) If third-party vendor interoperability is required for voice, then what fragmentation method should be placed on the ports?

Answer: FRF.11 – on Port A and Port B

Q5) If the call endpoints are permanently fixed and no third-party interoperability is required, then what fragmentation method should be placed on the ports that voice will be crossing?

Answer: Cisco Trunk FRF.11 – on Port A and Port B

Q6) If dynamic switched voice is required, then what fragmentation method should be placed on the ports?

Answer: Switched FRF.11 – Port A and Port B

Q7) Will traffic shaping be required for this network?

Answer: Yes

Module 3: Cisco CallManager Overview and Gateway Selection

Laboratory Exercise: CCM

Answers:

Table: CallManager Chart

Questions	Main	Branch	Remote
How many servers?	3 – 2 primaries 1 publisher 1 in each building	2 – 1 primary 1 publisher	None – Use OPX SRST
What type of gateways?	Intercluster Trunk	Intercluster Trunk	4224
Bandwidth Requirement?	480 for voice	480 for voice	120 for voice
What type of call admission control?	Gatekeeper	Gatekeeper for calls to main. Locations for call to remote	Locations
What code is used?	7.11 in campus 7.29 between sites	7.11 in office 7.29 between sites	7.11 in office 7.29 between sites

Module 6: Voice Over Data Migration

Laboratory Exercise: QoS

Practice

The company wants to merge its voice traffic onto an existing data network. Determine the interfaces that require QoS tools and identify the QoS tools that need to be implemented. Use the QoS Chart to record your solutions.

Answers:

Table: QoS Chart

Tools	Leased Lines	VoIPoFR	VoFR
Classification	IP Prec	IP Prec	IP Prec
Prioritization	LLQ	IPRTP	LLQ
Slow-Speed Link Efficiency	MLPPP	FRF.12	FRF.12
Traffic Shaping Required	No	Yes	Yes
Admission Control Required	Yes	Yes	Yes

