Cisco Unified Contact Center Express Solution Reference Network Design
Cisco Unified Contact Center Express, Release 4.1
August 2007

Americas Headquarters
Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134-1706
USA
http://www.cisco.com
Tel: 408 526-4000
800 553-NETS (6387)
Fax: 408 527-0883
CONTENTS

Preface vii
Purpose vii
Audience vii
Scope vii
Software Releases viii
Document Structure viii
Revision History ix
Obtaining Documentation ix
Cisco.com ix
Product Documentation DVD ix
Ordering Documentation x
Documentation Feedback x
Cisco Product Security Overview x
Reporting Security Problems in Cisco Products xi
Obtaining Technical Assistance xi
Cisco Technical Support & Documentation Website xii
Submitting a Service Request xii
Definitions of Service Request Severity xii
Obtaining Additional Publications and Information xiii

CHAPTER 1 Cisco Unified Contact Center Express Overview and Packaging 1-1
Unified CCX Overview 1-1
Unified CCX Packaging 1-2
Unified CCX Licensing 1-3
Basic IVR Functionality 1-4
Basic ACD Functionality 1-4
Basic CTI Functionality 1-7
Advanced IVR Functionality 1-7
Advanced ACD Functionality 1-9
Advanced CTI Functionality 1-11
Historical Reporting 1-11
## Chapter 2

**Unified CCX Terminology**  2-2  
**Unified CCX Call Processing**  2-3  
**Unified CCX System Management**  2-5  
**CRS Engine and Database Components**  2-5  
**Monitoring and Recording Components**  2-6  
**7920 Wireless IP Phone Support**  2-9  
**Citrix and Microsoft Terminal Services Support for CAD**  2-10  
**Unified CCX ASR and TTS**  2-11  
**Unified CCX Integration with Unified ICME Software**  2-11  
**Unified CCX Fault Tolerance**  2-14  
  - **CRS Engine Redundancy**  2-14  
  - **Database Redundancy**  2-15  
  - **Monitoring and Recording Redundancy**  2-16  
  - **Cold Standby Support**  2-17  
**Upgrading to Unified CCX 4.1**  2-18  
**CRS 4.1 Software Compatibility**  2-18

## Chapter 3

**Unified CCX General Rules for Design**  3-3  
**Cisco Unified CallManager Co-Resident Deployment Model (1)**  3-4  
**Single-Server Non-HA Deployment Model (2)**  3-5  
**Multi-Server Non-HA Deployment Model (3)**  3-5  
**Two-Server HA Deployment Model (4)**  3-6  
**Four-Server HA Deployment Model (5)**  3-6  
**Six-Server HA Deployment Model (6)**  3-7  
**Ten-Server HA Deployment Model (7)**  3-7  
**Other Design Considerations**  3-8

## Chapter 4

**Terminology**  4-1  
**Preliminary Information Requirements**  4-2  
**Principal Design Considerations for Call Center Sizing**  4-4  
**Planning Resource Requirements for Call Center Sizing**  4-5
Preface

Purpose

This document provides system-level best practices and design guidance for the Cisco Unified Contact Center Express (Unified CCX), Release 4.1. With proper planning, design, and implementation, Unified CCX provides a reliable and flexible voice processing and contact center solution for the enterprise.

Audience

This design guide is intended for the system architects, designers, engineers, and Cisco channel partners who want to apply best design practices for Unified CCX.

This design guide assumes that the reader is already familiar with the following concepts:

• Cisco Unified CallManager Administration
• Unified CCX and Cisco Unified IP IVR (Unified IP IVR) administration
• General system requirements and network design guidelines available from your local Cisco Systems Engineer (SE)

Scope

This document describes the various components used to build a Unified CCX system, and it gives recommendations on how to combine those components into an effective solution for your enterprise.

The following topics are not covered in this design guide:

• Installation and configuration of Unified CCX, Unified IP IVR, and Cisco Agent Desktop. For more information about these Cisco products, refer to the online product documentation available at Cisco.com.
• Unified IP IVR and Cisco Unified Queue Manager (Unified QM) programming guidelines. Unified CCX is a packaged solution built upon a Cisco software platform called Customer Response Solutions (CRS). The CRS platform supports other solution packages—Unified IP IVR and Unified QM. Unified IP IVR and Unified QM are primarily used with Cisco Unified Contact Center Enterprise (Unified CCE). Unlike Unified CCX, the Unified IP IVR and Unified QM solutions do not provide ACD and CTI functions. In Unified CCE deployments, the ACD and CTI functions are
Software Releases

Unless stated otherwise, the information in this document applies specifically to Unified CCX Release 4.1. Software releases are subject to change without notice, and those changes may or may not be indicated in this document. Refer to the Unified CCX release notes for the latest software releases and product compatibility information.

Document Structure

This guide contains the following chapters and appendices:

- Chapter 1, Unified CCX Overview and Packaging, provides an overview of the Unified CCX software and describes the Unified CCX packaging.
- Chapter 2, Unified CCX Architecture, describes the terminology, call processing, system management, CRS Engine and Database Service, Monitoring and Recording Services, ASR and TTS, integration with Unified ICME, fault tolerance, upgrades, and software compatibility for Unified CCX.
- Chapter 3, Unified CCX Deployment Models, describes the various ways Unified CCX can be deployed.
- Chapter 4, Basics of Call Center Sizing, introduces the basic concepts involved in call center sizing.
- Chapter 5, Sizing Unified CCX and Cisco Unified CallManager Servers, discusses the impact of performance criteria on the Unified CCX and Cisco Unified CallManager servers.
- Chapter 6, Bandwidth, Security, and QoS Considerations, discusses estimating bandwidth consumption, serviceability and security, and quality of service and call admission control.
- Appendix A, Server Capacities and Limits, provides a list of server capacities and limits.
- Appendix B, Voice Over IP Monitoring, provides design considerations for SPAN-based services.
• Appendix C, Unified CCX Integration with LDAP Server, provides information about directory services.
• The Index helps you find information in this guide.

Revision History

The following table lists the revision history for this document.

<table>
<thead>
<tr>
<th>Revision Date</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>July 18, 2006</td>
<td>First draft review</td>
</tr>
<tr>
<td>January 4, 2007</td>
<td>Final draft review</td>
</tr>
<tr>
<td>March 2, 2007</td>
<td>Draft for posting</td>
</tr>
</tbody>
</table>

Obtaining Documentation

Cisco documentation and additional literature are available on Cisco.com. Cisco also provides several ways to obtain technical assistance and other technical resources. These sections explain how to obtain technical information from Cisco Systems.

Cisco.com

You can access the most current Cisco documentation at this URL:
http://www.cisco.com/techsupport
You can access the Cisco website at this URL:
http://www.cisco.com
You can access international Cisco websites at this URL:
You can access more information about Unified CCX at this URL:
http://www.cisco.com/go/ipccexpress
You can access all of the Unified CCX 4.1 documentation at this URL:
www.cisco.com/univercd/cc/td/doc/product/voice/sw_ap_to/apps_4_0/english/index.htm

Product Documentation DVD

Cisco documentation and additional literature are available in the Product Documentation DVD package, which may have shipped with your product. The Product Documentation DVD is updated regularly and may be more current than printed documentation.
The Product Documentation DVD is a comprehensive library of technical product documentation on portable media. The DVD enables you to access multiple versions of hardware and software installation, configuration, and command guides for Cisco products and to view technical documentation in HTML. With the DVD, you have access to the same documentation that is found on the Cisco website without being connected to the Internet. Certain products also have .pdf versions of the documentation available.

The Product Documentation DVD is available as a single unit or as a subscription. Registered Cisco.com users (Cisco direct customers) can order a Product Documentation DVD (product number DOC-DOCDVD=) from Cisco Marketplace at this URL:
http://www.cisco.com/go/marketplace/

Ordering Documentation

Beginning June 30, 2005, registered Cisco.com users may order Cisco documentation at the Product Documentation Store in the Cisco Marketplace at this URL:
http://www.cisco.com/go/marketplace/

Nonregistered Cisco.com users can order technical documentation from 8:00 a.m. to 5:00 p.m. (0800 to 1700) PDT by calling 1 866 463-3487 in the United States and Canada, or elsewhere by calling 011 408 519-5055. You can also order documentation by e-mail at tech-doc-store-mkpl@external.cisco.com or by fax at 1 408 519-5001 in the United States and Canada, or elsewhere at 011 408 519-5001.

Documentation Feedback

You can rate and provide feedback about Cisco technical documents by completing the online feedback form that appears with the technical documents on Cisco.com.

You can send comments about Cisco documentation to bug-doc@cisco.com.

You can submit comments by using the response card (if present) behind the front cover of your document or by writing to the following address:

Cisco Systems
Attn: Customer Document Ordering
170 West Tasman Drive
San Jose, CA 95134-9883

We appreciate your comments.

Cisco Product Security Overview

Cisco provides a free online Security Vulnerability Policy portal at this URL:

From this site, you can perform these tasks:
• Report security vulnerabilities in Cisco products.
• Obtain assistance with security incidents that involve Cisco products.
• Register to receive security information from Cisco.

A current list of security advisories and notices for Cisco products is available at this URL:
http://www.cisco.com/go/psirt

If you prefer to see advisories and notices as they are updated in real time, you can access a Product Security Incident Response Team Really Simple Syndication (PSIRT RSS) feed from this URL:

**Reporting Security Problems in Cisco Products**

Cisco is committed to delivering secure products. We test our products internally before we release them, and we strive to correct all vulnerabilities quickly. If you think that you might have identified a vulnerability in a Cisco product, contact PSIRT:

- **Emergencies** — security-alert@cisco.com
  
  An emergency is either a condition in which a system is under active attack or a condition for which a severe and urgent security vulnerability should be reported. All other conditions are considered nonemergencies.

- **Nonemergencies** — psirt@cisco.com

In an emergency, you can also reach PSIRT by telephone:

- 1 877 228-7302
- 1 408 525-6532

---

**Tip**

We encourage you to use Pretty Good Privacy (PGP) or a compatible product to encrypt any sensitive information that you send to Cisco. PSIRT can work from encrypted information that is compatible with PGP versions 2.x through 8.x.

Never use a revoked or an expired encryption key. The correct public key to use in your correspondence with PSIRT is the one linked in the Contact Summary section of the Security Vulnerability Policy page at this URL:


The link on this page has the current PGP key ID in use.

---

**Obtaining Technical Assistance**

Cisco Technical Support provides 24-hour-a-day award-winning technical assistance. The Cisco Technical Support & Documentation website on Cisco.com features extensive online support resources. In addition, if you have a valid Cisco service contract, Cisco Technical Assistance Center (TAC) engineers provide telephone support. If you do not have a valid Cisco service contract, contact your reseller.

---

**Note**

For Unified CCX product management, pricing, packaging, or licensing questions, please send an e-mail to ask-icd-ivr-pm@external.cisco.com. For Unified CCX non-defect, usage, or technical questions, please send an e-mail to ask-icd-ivr-support@external.cisco.com. Information for Unified CCX can also be found at http://www.cisco.com/go/ipccexpress.
Obtaining Technical Assistance

Cisco Technical Support & Documentation Website

The Cisco Technical Support & Documentation website provides online documents and tools for troubleshooting and resolving technical issues with Cisco products and technologies. The website is available 24 hours a day, at this URL:

http://www.cisco.com/techsupport

Access to all tools on the Cisco Technical Support & Documentation website requires a Cisco.com user ID and password. If you have a valid service contract but do not have a user ID or password, you can register at this URL:


Note

Use the Cisco Product Identification (CPI) tool to locate your product serial number before submitting a web or phone request for service. You can access the CPI tool from the Cisco Technical Support & Documentation website by clicking the Tools & Resources link under Documentation & Tools. Choose Cisco Product Identification Tool from the Alphabetical Index drop-down list, or click the Cisco Product Identification Tool link under Alerts & RMAs. The CPI tool offers three search options: by product ID or model name; by tree view; or for certain products, by copying and pasting show command output. Search results show an illustration of your product with the serial number label location highlighted. Locate the serial number label on your product and record the information before placing a service call.

Submitting a Service Request

Using the online TAC Service Request Tool is the fastest way to open S3 and S4 service requests. (S3 and S4 service requests are those in which your network is minimally impaired or for which you require product information.) After you describe your situation, the TAC Service Request Tool provides recommended solutions. If your issue is not resolved using the recommended resources, your service request is assigned to a Cisco engineer. The TAC Service Request Tool is located at this URL:

http://www.cisco.com/techsupport/servicerequest

For S1 or S2 service requests, or if you do not have Internet access, contact the Cisco TAC by telephone. (S1 or S2 service requests are those in which your production network is down or severely degraded.) Cisco engineers are assigned immediately to S1 and S2 service requests to help keep your business operations running smoothly.

To open a service request by telephone, use one of the following numbers:

Asia-Pacific: +61 2 8446 7411 (Australia: 1 800 805 227)
EMEA: +32 2 704 55 55
USA: 1 800 553-2447

For a complete list of Cisco TAC contacts, go to this URL:

http://www.cisco.com/techsupport/contacts

Definitions of Service Request Severity

To ensure that all service requests are reported in a standard format, Cisco has established severity definitions.
Severity 1 (S1)—Your network is “down,” or there is a critical impact to your business operations. You and Cisco will commit all necessary resources around the clock to resolve the situation.

Severity 2 (S2)—Operation of an existing network is severely degraded, or significant aspects of your business operation are negatively affected by inadequate performance of Cisco products. You and Cisco will commit full-time resources during normal business hours to resolve the situation.

Severity 3 (S3)—Operational performance of your network is impaired, but most business operations remain functional. You and Cisco will commit resources during normal business hours to restore service to satisfactory levels.

Severity 4 (S4)—You require information or assistance with Cisco product capabilities, installation, or configuration. There is little or no effect on your business operations.

Obtaining Additional Publications and Information

Information about Cisco products, technologies, and network solutions is available from various online and printed sources.

- Cisco Marketplace provides a variety of Cisco books, reference guides, documentation, and logo merchandise. Visit Cisco Marketplace, the company store, at this URL:
  
  http://www.cisco.com/go/marketplace/

- **Cisco Press** publishes a wide range of general networking, training and certification titles. Both new and experienced users will benefit from these publications. For current Cisco Press titles and other information, go to Cisco Press at this URL:

  http://www.ciscopress.com

- **Packet magazine** is the Cisco Systems technical user magazine for maximizing Internet and networking investments. Each quarter, Packet delivers coverage of the latest industry trends, technology breakthroughs, and Cisco products and solutions, as well as network deployment and troubleshooting tips, configuration examples, customer case studies, certification and training information, and links to scores of in-depth online resources. You can access Packet magazine at this URL:

  http://www.cisco.com/packet

- **iQ Magazine** is the quarterly publication from Cisco Systems designed to help growing companies learn how they can use technology to increase revenue, streamline their business, and expand services. The publication identifies the challenges facing these companies and the technologies to help solve them, using real-world case studies and business strategies to help readers make sound technology investment decisions. You can access iQ Magazine at this URL:

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

  or view the digital edition at this URL:

  http://ciscoiq.texterity.com/ciscoiq/sample/

- **Internet Protocol Journal** is a quarterly journal published by Cisco Systems for engineering professionals involved in designing, developing, and operating public and private internets and intranets. You can access the Internet Protocol Journal at this URL:

  http://www.cisco.com/ipj

- Networking products offered by Cisco Systems, as well as customer support services, can be obtained at this URL:

• Networking Professionals Connection is an interactive website for networking professionals to share questions, suggestions, and information about networking products and technologies with Cisco experts and other networking professionals. Join a discussion at this URL:
  http://www.cisco.com/discuss/networking

• World-class networking training is available from Cisco. You can view current offerings at this URL:
Cisco Unified Contact Center Express Overview and Packaging

This chapter describes the basic architecture and capabilities of Cisco Unified Contact Center Express (Unified CCX) and explains how to match those capabilities to your system requirements. This chapter contains the following sections:

- Unified CCX Overview, page 1-1
- Unified CCX Packaging, page 1-2

Unified CCX Overview

Unified CCX is a tightly integrated contact center solution providing three primary functions—IVR, ACD, and CTI. The IVR function provides up to 300 IVR ports to interact with callers by way of either DTMF or speech input. The ACD function provides the ability to intelligently route and queue calls to up to 300 agents. The CTI function provides “screen pop” and interaction with other Windows-based desktop applications.

The Unified CCX software runs on approved Cisco MCS, HP, or IBM servers and uses Cisco Unified CallManager for call processing. The Unified CCX software can run on the same server with Cisco Unified CallManager (co-resident) or on a separate server. For more information on co-resident deployments, see the following web page:


For larger deployments requiring large amounts of historical reporting, silent monitoring, or recording, multiple servers might be required for the Unified CCX software. Unified CCX also offers a High Availability option which requires additional servers. A major purpose of this design guide is to help system designers determine the number and type of servers required for an Unified CCX deployment.

Cisco Unified CallManager provides the functionality typically associated with a PBX—call setup, teardown, and transition (transfer, conference, hold, retrieve). For calls requiring intelligent routing and queueing, Cisco Unified CallManager interacts with Unified CCX. Within Cisco Unified CallManager, a phone device called a CTI port is defined. A CTI port is an IP endpoint where a Voice over IP (VoIP) call can be terminated. A CTI port is defined in Cisco Unified CallManager for each Unified CCX IVR port. For the remainder of this document, we will refer to the IVR ports and CTI ports interchangeably.

When a new call arrives at Cisco Unified CallManager, if the dialed number is associated with the Unified CCX server, Cisco Unified CallManager asks the Unified CCX server which CTI port to route the call to. After the Unified CCX server selects an available CTI port and returns the directory number of that CTI Port to Cisco Unified CallManager, which sets up a VoIP data stream between the CTI port
Unified CCX Packaging

Unified CCX provides three primary functions—IVR, ACD, and CTI. Within the Unified CCX packaging, you have a choice of either basic or advanced feature sets for each of these functions. These feature sets are packaged into three different Unified CCX licensed packages—Standard, Enhanced, and Premium.

Unified CCX software requires both licensed server software and licensed seat software. Both the server and all seats must use the same packaging—Standard, Enhanced, or Premium. You cannot mix licenses. For example, you cannot have a Standard Unified CCX server with Enhanced Unified CCX seats.

The following table shows at a high level what functionality is included within each Unified CCX package. Details about each function are included in the sections that follow. The High Availability option is only available with Unified CCX Enhanced or Premium. The Unified ICME integration capability is available with all Unified CCX packages.

<table>
<thead>
<tr>
<th>Functionality</th>
<th>Standard Package</th>
<th>Enhanced Package</th>
<th>Premium Package</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basic IVR (prompt &amp; collect and queuing)</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Advanced IVR</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Basic ACD</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Advanced ACD</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Basic CTI</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Advanced CTI</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Additional packaging information can be found in *Getting Started with Cisco Unified Contact Center Express* located here:

Unified CCX Licensing

Throughout this document references are made to Unified CCX licenses. Please refer to this section for Unified CCX licensing definitions.

Unified CCX provides licensing for two product components:

- Unified CCX Active and (optionally) Standby server software
- Unified CCX Seats.

Unified CCX server software licenses enable IVR port licenses as defined in Table 1-1.

Table 1-1 IVR Port Licenses by Unified CCX Package

<table>
<thead>
<tr>
<th>Unified CCX Package</th>
<th>IVR Port Licenses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard</td>
<td>As many basic IVR ports (prompt &amp; collect) as can be supported given the Unified CCX primary server on which the package is deployed and the mix of other features also deployed on that server. The Unified CCX Configuration &amp; Ordering Tool must be used to determine the number of basic IVR ports able to be supported.</td>
</tr>
<tr>
<td>Enhanced</td>
<td>As many basic IVR ports (prompt &amp; collect) as can be supported given the Unified CCX Active and (optionally) Standby servers on which the package is deployed and the mix of other features also deployed on those servers. The Unified CCX Configuration &amp; Ordering Tool must be used to determine the number of basic IVR ports able to be supported.</td>
</tr>
<tr>
<td>Premium</td>
<td>Advanced IVR ports are available only with Unified CCX Premium. No more than 2 advance IVR ports are licensed at no charge for each Premium seat licensed. Additional IVR ports can only be provided by purchasing additional Premium seats. Each additional Premium seat will provide 2 additional advanced IVR Ports.</td>
</tr>
</tbody>
</table>

Unified CCX seats provide a quantity one license for each of the Unified CCX product components as shown in Table 1-2. Each seat license entitles deployment by the customer as required, constrained only by technical constraints defined in the Unified CCX Configuration & Ordering Tool and best practices as defined in this document.
Table 1-2  Licensed Seat Product Components by Unified CCX Package

<table>
<thead>
<tr>
<th>Unified CCX Package</th>
<th>Licensed Seat Product Components</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard</td>
<td>Cisco Agent Desktop (CAD) Standard</td>
</tr>
<tr>
<td></td>
<td>Cisco IP Phone Agent (IPPA) Standard</td>
</tr>
<tr>
<td></td>
<td>Cisco Supervisor Desktop (CSD) Standard</td>
</tr>
<tr>
<td></td>
<td>Cisco Historical Reporting Standard</td>
</tr>
<tr>
<td>Enhanced</td>
<td>Cisco Agent Desktop (CAD) Enhanced</td>
</tr>
<tr>
<td></td>
<td>Cisco IP Phone Agent (IPPA) Enhanced</td>
</tr>
<tr>
<td></td>
<td>Cisco Supervisor Desktop (CSD) Enhanced</td>
</tr>
<tr>
<td></td>
<td>Cisco Historical Reporting Enhanced</td>
</tr>
<tr>
<td></td>
<td>Cisco On Demand Recording</td>
</tr>
<tr>
<td>Premium</td>
<td>Cisco Agent Desktop (CAD) Enhanced</td>
</tr>
<tr>
<td></td>
<td>Cisco IP Phone Agent (IPPA) Enhanced</td>
</tr>
<tr>
<td></td>
<td>Cisco Supervisor Desktop (CSD) Enhanced</td>
</tr>
<tr>
<td></td>
<td>Cisco Historical Reporting Enhanced</td>
</tr>
<tr>
<td></td>
<td>Cisco On Demand Recording</td>
</tr>
</tbody>
</table>

Basic IVR Functionality

All Unified CCX packages include basic IVR functionality. Basic IVR (prompt and collect) provides the ability to prompt callers for information and to collect information by way of DTMF. This feature is used for menus (such as press 1 for sales, press 2 for service...) and basic information collection (please enter your account number, order number...). The number of CTI ports allowed varies by server type and what other functions are running on that server. The maximum number of CTI ports possible for an Unified CCX deployment is 300. The ordering and configuration tool assists you in the sizing and selection of an appropriate server for any given deployment scenario. The basic CTI ports are not licensed separately. The cost for the basic IVR functionality is included in the server and seat licensing costs.

Basic call controls like terminate, transfer, and place call are also supported as part of the basic IVR functionality.

Basic XML document processing is also supported as part of the basic IVR functionality. This function could be used to access system-wide static data like a list of holidays, hours of operations, or a short list of hot customer accounts.

Basic ACD Functionality

All Unified CCX packages include basic ACD functionality. Here we define ACD functionality in the following five areas:

- Call routing and queuing
- Cisco Agent Desktop (CAD)
- IP Phone Agent (IPPA)
- Cisco Supervisor Desktop (CSD)
Call Routing and Queuing

The Unified CCX Basic ACD functionality provides the following call routing and queuing capabilities:

- **Historical Reporting**

**Conditional Routing.** Unified CCX supports routing based upon caller input to menus, real-time queue statistics, time of day, day of week, ANI, dialed number, and processing of data from XML text files.

- **Agent Selection.** Unified CCX supports longest available, linear, most handled contacts, shortest average handle time, and circular agent selection algorithms. With Basic ACD functionality, agents are associated with one resource group only.

- **Customizable Queuing Announcements.** Unified CCX supports the playing of customizable queuing announcements based upon any of the conditions specified above or based upon the skill group the call is being queued to. This includes announcements related to position in queue and expected delay.

- **Re-route on Ring No Answer.** If the selected agent does not answer within the allowed time limit, then the caller retains his position in queue. Any screen pop data is also preserved.

- **Cisco Unified Intelligent Contact Management Enterprise (Unified ICME) Integration.** Unified CCX has the ability to integrate with Unified ICME 7.x. Unified ICME integration provides the following:
  - The ability for Unified CCX to send agent, queue, and call state changes to Unified ICME software.
  - The ability of Unified ICME software to intelligently route and load balance (using pre-routing or post-routing) calls across multiple ACD sites, which can include one or more Unified CCX systems, Unified CCE systems, or traditional ACDs (that are supported by Unified ICME software). Calls routed to an Unified CCX application can also be sent call data so that it can be popped onto an agent’s screen.
  - The ability for Unified CCX to send post-route requests with call data to the Unified ICME software in order to request routing instructions. This could be in response to a new call that just arrived at Unified CCX or a call that is being transferred from an IVR port or agent. Call data included in the post-route request can be used by the Unified ICME software to profile route the call, and call data is also passed to the terminating ACD site (Unified CCX, Unified CCE, or traditional ACD) for an agent screen pop.
  - The ability for Unified ICME software to provide multi-site ACD reporting for a mixed network of ACD sites, which can include one or more Unified CCX systems, Unified CCE systems, or traditional ACDs.

**Cisco Agent Desktop (CAD)**

The Unified CCX Basic ACD functionality includes an agent desktop with the following features and options:

- **Agent State Control.** From the agent desktop, agents log in, log out, make themselves ready and not ready.

- **Call Control.** From the agent desktop, agents answer, release, hold, retrieve, conference, and transfer calls. Note that call control for agents using an IP Phone can also be done from the IP Phone. For example, to answer a call, the agent can simply pickup the IP Phone handset. The Unified CCX software ensures that the current call state for the IP Phone and CAD application are kept in synch.
- **Dynamic Regrouping.** Change of agent association with a resource group is applied immediately.
- **Real-Time Statistics.** Agents have access to real-time statistics for themselves and the queues to which they are associated. For example, from the agent desktop application, the agent can see how many calls they have handled today and how many calls are currently in queue for their team.
- **Integrated Text Messaging.** Agents can interact with their supervisor and other agents by way of text chat.
- **Reason Codes.** Agents can be configured to enter reason codes for not ready and logout.
- **Basic CTI.** Agent desktops provide an enterprise data window that is “popped” upon call ringing. See the section Basic CTI for more information on the enterprise data window.
- **Telephony Support.** CAD can be deployed with any Cisco 7900 Series IP Phone, including the Wireless IP Phone 7920. However, there are different features available on different phones. For considerations regarding the 7920, see Chapter 2, “Monitoring With the 7920 Wireless Phone. The 7902 and 7905 phones do not have a headset jack and therefore might not be appropriate for usage in a call center environment. In agent environments without a Cisco IP Phone, CAD also supports the agent using the Cisco IP Communicator softphone application running on the same workstation with CAD. An agent’s ACD (Unified CCX) extension is only valid with a single line. An agent’s ACD extension must not be configured to forward to no answer to voice mail or any other termination point. Doing so might impact re-routing on ring no answer of an Unified CCX routed call to another agent or back to queue. Agents who need to be contacted directly or who need voicemail should have their phones configured with a second extension (and multiple lines if necessary). Unified CCX does not monitor or report on activity on the non-ACD extensions on a phone. During the CAD login process, agents supply the designated ACD extension on their phone. Agents are associated with a specific Cisco Unified CallManager extension (directory number).
- **Hot Desking.** Hot desking allows agents to log in using CAD and any IP Phone registered with the same Cisco Unified CallManager cluster. Agents using CAD and IP Communicator can also use Extension Mobility. This also allows multiple agents to use the same phone—but only one at a time. In order to hot desk, agents must first log into Cisco Unified CallManager using the CallManager Extension Mobility feature. Extension Mobility brings a user specific phone profile (including configured extensions for that user) to the phone being logged in from. After logging into Cisco Unified CallManager with Extension Mobility, agents can log into Unified CCX using CAD.
- **Auto Update.** At CAD startup, it checks to see if a new version of the CAD program is available and automatically performs an update on the agent workstation. The workstation needs administrator rights for this auto update to occur.

### IP Phone Agent (IPPA)

For environments where agents do not have a workstation running the Windows operating system (Citrix terminals, Windows terminal, UNIX workstations, Macintosh workstations), Unified CCX offers IP Phone Agent (IPPA). IPPA is an XML application executing on a Cisco IP Phone 7940, 7960, or 7970, that provides an agent interface using the display and softkeys on the IP Phone. An agent cannot be using both CAD and IPPA simultaneously. However, an agent who typically uses CAD can also use IPPA in situations when the agent’s CAD workstation is inoperable. IPPA is not licensed separately. Unified CCX software is licensed by the seat. Seats are licensed based on maximum simultaneous logins. The Unified CCX basic ACD functionality includes IP Phone Agent with the following features:

- **Agent State Control.** From the IPPA XML application, agents log in, log out, and make themselves ready or not ready.
- **Call Control.** IPPA does not really provide call control using the IPPA XML application. The IP Phone itself provides the ability to perform call control.
• **Real-Time Statistics.** Agents have access to basic real-time statistics for the queues to which they are associated using the IP Phone Agent XML application.

• **Reason Codes.** Agents can be configured to enter reason codes for not ready and logout using the IPPA XML application.

• **Basic CTI.** IPPA allows for call data to be popped onto the IP Phone display upon call ringing.

• **Telephony Support.** IP Phone Agent is supported on the 791x, 794x, 796x, and 797x modes of IP Phones. Also, the wireless 7920 phone is supported.

• **Hot Desk.** Hot desk allows agents to log in using IPPA from any IP Phone registered with the same Cisco Unified CallManager cluster. This also allows multiple agents to use the same phone—but only one at a time. In order to hot desk, agents must first log into Cisco Unified CallManager using the Extension Mobility feature. Extension Mobility brings a user specific phone profile (including configured extensions for that user) to the phone being logged in from. After logging into Cisco Unified CallManager with Extension Mobility, agents can log into Unified CCX using IPPA.

**Cisco Supervisor Desktop (CSD)**

The Unified CCX Basic ACD functionality provides a separate supervisor desktop application (CSD). If a supervisor wishes to handle calls, then the supervisor uses both CAD and CSD. CSD and supervisors are not licensed separately. Supervisors are licensed the same as agents. If you need a call center with 10 agents and 1 supervisor, then you should order 11 seats. Seats are licensed based on maximum simultaneous logins.

The supervisor desktop provides the following features and options:

• **View / Change Agent State.** Supervisor desktops allow supervisors to view the current state of all agents that are part of that supervisor’s team. The supervisor desktop also allows supervisors to change an agent’s state (ready, not ready, logout).

• **Real-Time Agent and Skill Statistics.** Supervisors can view statistics for all agents and queues that are associated with their team. See the *Cisco Supervisor Desktop User’s Guide* for more details on statistics available through the supervisor desktop application.

• **Integrated Text Messaging.** Supervisors can send text messages to one or more agents.

• **Marquee Messages.** Supervisors can send a scrolling marquee (broadcast) message to all agents on their team.

**Basic CTI Functionality**

All Unified CCX packages include basic CTI functionality. The basic CTI functionality provides a customizable enterprise data window that is “popped” on the agent desktop upon call ringing. Data within the enterprise data window includes ANI, dialed number, and any caller input (account number, order number, case number, reason for calling...), plus details on how long the caller interacted with the IVR, how long the caller waited in queue, and how long the caller spent with all other agents if this was a transferred call.

**Advanced IVR Functionality**

The Unified CCX Premium Package includes both basic and advanced IVR functionality. Cisco provides no charge licenses for two advanced IVR ports for every licensed Unified CCX Premium seat.
The Unified CCX server has a single licensing flag which designates whether IVR ports have basic or advanced functionality. Therefore, all ports must be the same—all basic or all advanced. If you need any of the advanced IVR features, you must order the Unified CCX Premium packaging.

In addition to the functionality discussed above in the section Basic IVR Functionality, page 1-4, the advanced IVR functionality includes the following:

- **Database Integration.** The Unified CCX server can interoperate with any ODBC-compliant database. Databases tested and supported by Cisco are listed in the Cisco CRS Software and Hardware Compatibility Guide located here:


  Data retrieved from databases can be used with the conditional routing capabilities to provide customer profile-based routing and queueing. For example, a premier customer could be routed to a different group of agents or prioritized higher in queue than non-premier customers. Customers who have purchased products A & B, could be told about complementary product C while they are waiting for an agent. Database integration also provides the ability to offer complete self-service applications to callers. Database views are not supported using the CRS Editor database steps, but database views could be accessed using VoiceXML or Java logic modules.

- **HTTP Triggers.** The Unified CCX server can receive a customer contact request by way of an HTTP trigger. This allows web users to be offered service by way of a “click to talk to an agent” button. Information collected using the web (a customer call back number, account number, shopping cart content, and so forth) can be passed to the Unified CCX script to allow customer profile-based routing and a data-rich screen pop. These contacts can be prioritized and routed using the same methods available to normal inbound voice callers. HTTP Triggers could also be used to provide some simple browser-based e-mail routing, text chat request routing, call back request routing, and preview outbound dialing. The result is an integrated enterprise-wide multichannel, inbound and outbound blended queue on customer contacts. Note though that support for these application usage examples are not provided out of the box and require development, testing, and support by an experienced application developer from a Cisco IPC Specialized partner.

- **E-mail Generation.** The Unified CCX server can generate and send e-mails for things such as order confirmation. E-mail attachments are also supported. An e-mail could potentially be sent to a fax server solution which would accept the e-mail and fax it to the appropriate number. Please note this capability is not provided out-of-the-box by Unified CCX, but instead must be developed, tested, and supported by a Cisco IPC Specialized partner. Unified CCX Premium can be integrated with third-party FAX and paging services to provide on demand FAX and paging services under workflow control. Please refer to the following white paper with details and examples.


- **Voice XML 2.0 Support.** Unified CCX supports executing application logic developed with the Voice XML standard. VXML is required for certain complex grammar ASR and TTS interactions and is optional for an DTMF or simple ASR or TTS voice interaction service creation. This allows organizations to reuse application logic from other applications—like a transaction server to a mainframe database.

- **Java Support.** The Unified CCX server can support logic defined using Java. Java support allows for logic from existing web and Java applications to be reused.

- **IVR Port Call Recording.** The Unified CCX server can record input from callers. This could be used to allow call center staff to remotely record new announcements or prompts. This could also be used to prompt callers to leave a message and then by way of Unified CCX application development, the voicemail could be routed to an appropriately qualified agent using the same
competency-based routing and prioritized queueing techniques available for normal inbound voice calls. Support for this example is not provided out-of-the-box and requires development, testing, and support by an IPC Specialized partner.

- **MRCP Integration to Automatic Speech Recognition (ASR) and Text-to-Speech (TTS) Services.** Unified CCX integration to MRCP compliant ASR and TTS servers is provided as part of Unified CCX Premium. Tested ASR and TTS vendors are Nuance and Scansoft. ASR and TTS software must be purchased from one of these vendors. Cisco no longer sells ASR and TTS software as an option for Unified CCX.

- **Remote Silent Monitoring** provides a mechanism for silent monitoring of calls using an IP Phone or a PSTN phone. This form of silent monitoring does not require a CSD application to be running but does require a seat license for any supervisor engaged in remote silent monitoring. Remote silent monitoring also does not require any data network connectivity and is ideally suited for management of outsourcer customers of a call center service provider. The agent is unaware when being silent monitored using remote silent monitoring.

### Advanced ACD Functionality

The Unified CCX Enhanced and Premium packages include both basic and advanced ACD functionality. In addition to the basic ACD functionality discussed in the section, Basic ACD Functionality, advanced ACD functionality is provided in the following five areas:

- Call routing and queueing
- Cisco Agent Desktop (CAD)
- IP Phone Agent (IPPA)
- Cisco Supervisor Desktop (CSD)
- Historical Reporting

### Call Routing and Queuing

The Advanced ACD functionality provides the following call routing and queueing features:

- **Agent Skill and Competency-Based Routing.** Agents can be configured with multiple skills (up to 50), each with a different competency level (up to 10). Contact Service Queues (also known as skill groups) can be configured as requiring multiple skills (up to 50), each with a different minimum skill competency level (up to 10). The Unified CCX routing logic then matches the caller and contact requirements with agent skills to find the optimal match using one of the following agent selection criteria:
  - Longest Available, Most Handled Contacts, or Shortest Average Handle time
  - Most skilled, most skilled by weight, or most skilled by order
  - Least skilled, least skilled by weight, or least skilled by order

- **Dynamic Reskilling.** Changes to CSQ skills and competencies and agent skills and competencies are applied immediately.

- **Prioritized Queuing.** Customer contacts can be prioritized (up to 10 levels) based upon call or customer data, and calls may be moved within or among queues under workflow control using priority information.

- **Agent Routing.** Unified CCX routing applications can select a specific agent if that agent is in a “ready” state.
• **Wrap-up and Work Mode.** After call completion, an agent can be configured to be automatically placed into a work state, on a per CSQ basis. The agent can also optionally choose to enter work state if that option is provided by the agent’s desktop administrator. A wrap-up timer is also configurable on a per CSQ basis. Custom agent desktop and reporting development can be done to allow the entry of and reporting on wrap-up codes.

**Cisco Agent Desktop (CAD)**

The Advanced ACD functionality provides an agent desktop that includes the following additional features:

• **Application Integration.** CAD can be configured to allow call data to be passed to other desktop applications (like CRM applications) for an application screen pop. Passing data to other applications is performed by way of keystroke macros that are then associated with specific call events such as call ringing. With keystroke macros, no programming is required to develop a screen pop. Application integration can also be done upon call release to pop open a wrap-up application on the agent workstation.

• **Workflow Buttons.** CAD can be configured to have pre-defined workflow buttons that execute specified programs and keystrokes. Workflow buttons aid agents in completing repetitive tasks more quickly.

• **On-Demand Call Recording.** CAD can be configured to allow clicking a single button to start and stop call recording on demand. The call recording only contains the portion of the call that occurs after the start record button is clicked. There are limits to how many simultaneous call recording sessions can be performed. Later chapters in this document discuss these limits.

• **Complete Call Recording.** CAD can be configured to automatically start and stop recording upon call answer and release. Conditions upon which calls are to be recorded are defined in the application script. There are limits to how many simultaneous call recording sessions can be performed. Later chapters in this document discuss these limits.

• **Automatic Failover.** Upon failure of the active Unified CCX server, CAD will automatically re-login agents on the standby server, and the agent will be placed into a not ready state. Upon failure of the active Unified CCX server, active calls on agents phones will survive. However, the call duration and other information associated with the call in the historical reporting database may be impacted. Historical reports generated for time frames in which a failover occurred will have missing or incorrect data. It will be called out in the report that a failover occurred.

**IP Phone Agent (IPPA)**

The advanced ACD functionality provides IPPA the following additional feature:

• **On-Demand Call Recording.** IPPA can be configured to allow clicking a single button to start and stop call recording on demand. The call recording only contains the portion of the call that occurs after the start record button is clicked. There are limits to how many simultaneous call recording sessions can be performed. Later chapters in this document discuss these limits.

**Cisco Supervisor Desktop (CSD)**

The Advanced ACD functionality provides a supervisor desktop that includes the following additional features:

• **Silent Monitoring.** CSD allows a supervisor to silently monitor agent calls. Agents can be configured to be aware or unaware that they are being monitored.
**Barge-in.** CSD allows a supervisor to barge in on an agent call. The barge-in feature enters the supervisor, the agent, and the caller into a three-way conference. This feature requires the supervisor to have the CAD application open and to be logged in as an agent. The agent is aware when the supervisor barges in. Barge-in is supported for agents using CAD with IP Communicator, CAD with IP Phone, or IPPA.

**Intercept.** CSD allows a supervisor to intercept an agent call. The Intercept feature transfers the call to the supervisor. This feature requires the supervisor to have the CAD application open and to be logged in as an agent. As the call releases from the agent desktop and phone, the agent is aware when Intercept occurs. The agent is then available to take another call. Intercept is supported for agents using CAD with IP Communicator, CAD with IP Phone, or IPPA.

**On-Demand Agent Call Recording.** CSD allows a supervisor to dynamically start and stop recording agent calls on demand. Agents are not aware that they are being recorded. The call recording only contains the portion of the call that occurs after the start record button is clicked. There are limits to how many simultaneous call recording sessions can be performed. The deployment models chapter discusses these limits. Call Recording is supported for agents using CAD with IP Communicator, CAD with IP Phone, or IPPA.

**Call Recording Playback and Exports.** The CSD Record Viewer application allows a supervisor to playback calls which were recorded with the last 7 days. Supervisors can sort the recorded call list by agent, DN, or date/time. Within Record Viewer, supervisors can tag selected recordings for a 30-day extended archiving, and supervisors can also save selected recordings in a .wav format into a specified folder for permanent archiving.

**Automatic Failover and Re-login.** Upon CRS Engine failover, the CSD automatically fails over to the standby CRS Engine so the supervisor does not have to re-login.

### Advanced CTI Functionality

The Unified CCX Enhanced and Premium packages include both basic and advanced CTI functionality. In addition to the basic CTI functionality discussed in the section, **Basic CTI Functionality, page 1-7**, the advanced CTI functionality allows call data to be passed to other Windows-based desktop applications (like CRM applications) for an application screen pop on ringing. Passing data to other applications is performed by way of keystroke macros that are then associated with specific call events such as call ringing or call release. With keystroke macros, no programming is required to develop a screen pop application. With the Enhanced package, internal CRS engine-generated data or data obtained from XML data sources may be used. The Unified CCX Premium package adds support for using data from supported databases using workflow-based SQL queries.

### Historical Reporting

Supervisors can view historical reporting statistics for the entire contact center using the Historical Reports client. See the *Cisco CRS Historical Reports User Guide* for more reporting details available through the Historical Reports Application.

Custom reporting templates can be generated using a combination of the Crystal Reports Developer’s Toolkit and SQL stored procedures using the *Cisco CRS Database Schema*. For more information on custom reporting, see the *Cisco CRS Historical Reporting Administrator and Developer Guide*. 
Cisco Unified Contact Center Express Solution Architecture

Cisco Unified Contact Center Express (Unified CCX) is a solution composed of many components. These components include not just the Unified CCX software and the servers upon which that software runs, but also include Cisco Unified CallManager, Cisco routers, Cisco data switches, Cisco Voice Gateways, and Cisco IP Phones. Unified CCX software is part of the Customer Response Solutions (CRS) software platform. CRS provides the software capabilities for not just Unified CCX, but also Unified IP IVR and Unified Queue Manager (Unified QM). (Note that Extended Services is no longer offered as part of Cisco Unified CallManager for either Cisco CRS 3.5, 4.0, or 4.1 releases.) Unified IP IVR and Unified QM are primarily used for Cisco Unified Contact Center Enterprise (Unified CCE) deployments. A single physical server can run only one of the CRS packages, either Unified CCX, Unified IP IVR, or Unified QM.

This chapter includes the following sections:

- Unified CCX Terminology, page 2-2
- Unified CCX Call Processing, page 2-3
- Unified CCX System Management, page 2-5
- CRS Engine and Database Components, page 2-5
- Monitoring and Recording Components, page 2-6
- 7920 Wireless IP Phone Support, page 2-9
- Citrix and Microsoft Terminal Services Support for CAD, page 2-10
- Unified CCX ASR and TTS, page 2-11
- Unified CCX Integration with Unified ICME Software, page 2-11
- Unified CCX Fault Tolerance, page 2-14
- Upgrading to Unified CCX 4.1, page 2-18
- CRS 4.1 Software Compatibility, page 2-18
Unified CCX Terminology

JTAPI is the mechanism that Unified CCX uses to communicate with Cisco Unified CallManager for call processing. Within Cisco Unified CallManager, a JTAPI user is defined and that user ID is utilized by the Unified CCX JTAPI Subsystem to log into Cisco Unified CallManager. This user id will be referred to as the CRSJTAPI user ID. This login process is what allows Unified CCX to begin communications with Cisco Unified CallManager and offer services like routing control.

When a caller dials the number of an extension configured on an IP Phone, Cisco Unified CallManager is capable of setting up that call without the aid of Unified CCX. However, sometimes callers dial generic numbers that are not associated with any particular phone. In those situations, Cisco Unified CallManager needs a mechanism to request routing instruction from some other application. One such mechanism is a JTAPI route request message and one such application that can provide routing control is Unified CCX. In order for Cisco Unified CallManager to request routing from another application for a particular dialed number, there must be a CTI Route Point defined within Cisco Unified CallManager for that dialed number. Within Cisco Unified CallManager, the CTI Route Point is also associated with the user (application) that can provide routing control. This Cisco Unified CallManager configuration is what enables Cisco Unified CallManager to ask Unified CCX how to route a call. The creation of a CTI Route Point, the association of that CTI Route Point to the dialed number, and the user association of that CTI Route Point to the JTAPI user responsible for routing control is done automatically by the Unified CCX Server as part of the creation of a JTAPI Trigger.

The JTAPI trigger also specifies what CTI port group and CRS application to use for a specified dialed number. As discussed in Chapter 1, Unified CCX provides IVR functionality. A Unified CCX system can provide up to 300 logical IVR ports (also called CTI Ports). The CTI ports within Unified CCX are logical VoIP endpoints where calls can be terminated—very similar to a softphone. The difference is that these softphones are controlled by an application that has the ability to encode .wav files from disk into one of the supported VoIP formats (G.711 or G.729) and then stream those VoIP packets out the Ethernet interface on the Unified CCX Server to the calling VoIP endpoint (IP Phone or Voice Gateway port).

Each CTI Port must be defined within CallManager as a device with a type of ‘CTI Port.’ Each CTI Port device is assigned a unique directory number (extension), just like a phone. This allows Cisco Unified CallManager to setup calls to these devices and endpoints. The creation of the CTI Ports on CallManager is done automatically by the Cisco Unified CallManager server when a group of CTI Ports (Call Control Group) is defined.

When a caller dials a dialed number that is associated with a CTI Route Point, Cisco Unified CallManager sends a route request to Unified CCX which has the dialed number associated with a group of CTI Ports. The Unified CCX software selects an available CTI Port from that CTI Port Group and returns the extension of that CTI Port to Cisco Unified CallManager. Cisco Unified CallManager then attempts to setup a call to that extension (CTI Port) by sending a ring message to the Unified CCX server. When the Unified CCX server gets the ring message for a particular CTI Port for a particular dialed number, the Unified CCX server begins executing the script associated with that trigger’s application. The first step in a script is typically an Accept step. The Accept step in the application will answer the call and trigger Cisco Unified CallManager to establish an RTP stream between the selected CTI Port and the Voice Gateway (VG) port (or calling IP Phone). The application can then prompt callers for input and provide the caller self service. When either the caller hangs up or the application executes a Terminate step, Cisco Unified CallManager tears down the call.

Within the application, it is also possible to route or transfer the call to an available agent. If no agents are available, queueing treatment is provided to the caller. Agents in Unified CCX are called resources. There is a subsystem within Unified CCX called the Resource Manager which is responsible for monitoring the state of agents and selecting agents based upon the agent skills and queue skills required. Queues in Unified CCX are called Contact Service Queues (CSQs). Agents use CAD or IPPA state controls to log in and make themselves ready. The Resource Manager is updated upon every agent state change.
Administrators use the CRS Administration web interface to configure agent skills and competencies. CRS Administration is also used to define CSQ skill and competency requirements and the agent selection criteria to be used for that CSQ. Applications use the Select Resource step to specify the CSQ into which the caller shall be placed. The Resource Manager subsystem is queried by the application to select the appropriate agent based upon the agent selection criteria. If no agent is available, the Select Resource step has a queued branch where queueing treatment is defined. When the Resource Manager finds an available and appropriately skilled agent, it will reserve that agent and then request for that call to be transferred to the agents IP Phone (using JTAPI messaging to Cisco Unified CallManager). After the call has been transferred to and answered by the agent, the CTI Port being used for that call is released.

An agent must be configured in Cisco Unified CallManager as a user. This adds a record to the Cisco Unified CallManager LDAP directory. Cisco Unified CallManager supports usage of one of the following LDAP directory servers—DC Directory (default), Netscape IPlanet, and Microsoft Active Directory. The LDAP directory is installed as part of the CallManager installation process and both Cisco Unified CallManager and the LDAP directory must be operational prior to beginning an Unified CCX installation. In Cisco Unified CallManager, an agent’s phone and directory number are associated with the agent’s Cisco Unified CallManager user name and the directory number is also marked as an Cisco Unified CallManager extension. This allows Cisco Unified CallManager to know that this Cisco Unified CallManager user is an agent, and the user then shows up in the resource list in CRS Administration. In Cisco Unified CallManager, agent phones are also associated with another JTAPI user called the Resource Manager JTAPI user. This user is referred to as the RMJTAPI user. The RMJTAPI user allows Unified CCX to monitor the state of the phone. For example, when an agent goes off hook to make an outbound call using the Unified CCX extension, the Unified CCX application needs to be notified so that the Resource Manager can update its agent state machine to show that agent being on an outbound call. The RMJTAPI user also allows Unified CCX to control the state of the phone. For example, when an agent clicks Answer on Cisco Agent Desktop, this triggers Unified CCX to have the RMJTAPI user signal to Cisco Unified CallManager to have that agent’s phone go off hook.

**Unified CCX Call Processing**

Figure 2-1 and the description that follows explain a typical Unified CCX call flow:
1. Call arrives at Voice Gateway (VG)

2. Voice Gateway asks Cisco Unified CallManager (Unified CM) how to route the call (using H.323 or MGCP).

3. Cisco Unified CallManager has the dialed number (DN) associated with a CTI Route Point that is associated with a CRS JTAPI user for Unified CCX. This triggers a JTAPI route request to be sent to Unified CCX.

4. Based upon the DN, which is mapped to a JTAPI trigger, the Unified CCX server selects an available CTI port and replies back to Cisco Unified CallManager with the extension of the CTI Port to send this call to. CallManager then sends a call setup (ring) message to Unified CCX, which then maps the DN to the appropriate Unified CCX script. The Accept step (typically the first step) in the script will answer the call and trigger Cisco Unified CallManager to establish an RTP stream between the Voice Gateway port and the selected CTI Port. Then the script prompts the caller for an account number and does a database lookup. Then the caller is prompted to select from a menu of choices and is provided self-service treatment. If the user presses 0, we go to the transfer to agent section of the script. In this scenario, we are assuming no appropriately skilled agents are available, so the script executes the queued loop logic until an appropriately skilled agent becomes available.

5. An appropriately skilled agent becomes available as a result of logging in and going ready or completing a previous call.
Chapter 2      Cisco Unified Contact Center Express Solution Architecture

Unified CCX System Management

Several applications are available for administering and monitoring an Unified CCX deployment. The primary tool an administrator uses to manage an Unified CCX deployment is the CRS Administration web interface. CRS Administration is a web-based application accessed using a Windows Internet Explorer 6.0 or above browser. Using CRS Administration, administrators perform tasks such as uploading applications, uploading prompts, mapping applications to dialed numbers, configuring agent skills and CSQs, starting and stopping CRS subsystems, and monitoring overall cluster status.

In addition to CRS Administration, an administrator uses the Cisco CRS Editor. The Cisco CRS Editor is a client-based utility that produces .aef files which the administrator uploads using CRS Administration. The Cisco CRS Editor is automatically installed on servers with the CRS Engine by default. The Cisco CRS Editor can also be downloaded and installed from CRS Administration onto other workstations.

The Cisco Desktop Administrator (CDA) is another client-based utility that can be downloaded and installed from CRS Administration. It is installed on the servers with the CRS Engine by default. CDA allows an administrator to perform tasks such as configuring the agent interface, setting up reason codes, and defining agent workflows and keystroke macros.

Another client utility to monitor an Unified CCX deployment is the Historical Reports client application. You download and install the Historical Reports client from CRS Administration. There are 27 historical reporting templates available. Taken in combination with filtering parameters and chart or no chart options, there are 255 possible reports available. Custom reporting templates can be created with Crystal Reports development toolkit. Within CRS Administration, there are also 11 browser-based real-time reports. The Cisco Supervisor Desktop (CSD) and Cisco Agent Desktop (CAD) both also provide reports to allow real-time monitoring of an Unified CCX deployment. Both CSD and CAD are downloaded and installed from CRS Administration.

For additional information about CRS Administration, see the Cisco CRS Administration Guide.

CRS Engine and Database Components

Unified CCX has four core software components:

- CRS Engine
- Database
- Monitoring
- Recording
Every Unified CCX deployment must have a CRS Engine component and a Database component. These software processes can run on the same physical server or on separate servers. When separated, they must be located in the same Campus LAN with the maximum round trip delay of 2 ms between them. A standby CRS Engine component and Database component are options that will be discussed later in this chapter. The Monitoring component and Recording component are optional and are discussed in the next section of this chapter. The combined set of physical servers upon which these software modules run is called an Unified CCX cluster. The maximum cluster size for Unified CCX is 10 physical servers. The cluster concept for Unified CCX is a little different than Cisco Unified CallManager. In a Cisco Unified CallManager cluster, every node has a copy of the database. In an Unified CCX cluster, at most two nodes will have a database. The other nodes within an Unified CCX cluster will be running one or more of the other software server processes. Design rules for Unified CCX deployment models will be discussed in the next chapter.

The CRS Engine (and closely related subsystems) is the component that provides functions like the following:

- JTAPI communications with Cisco Unified CallManager
- Execution of scripts
- Encoding and streaming of .wav files for all CTI Ports defined
- Communications with CAD for agent state control, call control, and screen pop
- Agent monitoring and selection
- CRS Administration web interface.

Simply put, one can think of the CRS Engine component as providing the core ACD, IVR, and CTI services. The other components—Database, Monitoring, and Recording—are auxiliary software components that can be run on separate physical servers from the CRS Engine if the Unified CCX deployment is large enough to warrant additional hardware computing resources.

The Database component is a required component for any Unified CCX deployment and is the component that manages access to the database. The CRS Database contains four data stores. They are as follows:

- Configuration data store
- Repository data store
- Agent data store
- Historical data store

The configuration data store contains Unified CCX configuration information like Resources (agents), skills, resource groups, teams, and CSQ information. The repository data store contains user prompts, grammars, and documents. The agent datastore contains agent logs, statistics, and pointers to the recording files. The historical data store contains Contact Call Detail Records (CCDRs).

**Monitoring and Recording Components**

The previous section introduced the CRS Engine and Database components. This section introduces the Monitoring and the Recording components. These components are software processes that can run on the same physical server or separate servers from the CRS Engine and Database components. Details on supported deployment scenarios are covered in Chapter 3.

Unified CCX Enhanced and Premium provide the ability for a supervisor to silently monitor agents. Unified CCX Enhanced and Premium also provide the ability for agent calls to be recorded. Agent call recording can be triggered in the following ways:
Monitoring and Recording Components

- Supervisor clicks record button on Cisco Supervisor Desktop (CSD) for a specified agent call
- Agent clicks record button on Cisco Agent Desktop (CAD) or IP Phone Agent (IPPA)
- Workflow configuration automatically triggers complete call recording on certain types of calls for agents using CAD.

In order to use the silent monitoring or recording features, access to the RTP (Real-Time Protocol) packet streams is required. Silent monitoring and recording will work with either G.711 or G.729 RTP streams and a mixture of agents using G.711 and G.729 phones is supported. However, silent monitoring and recording will not work with encrypted media streams. Unified CCX provides two mechanisms for access to the RTP packet stream—SPAN port monitoring and desktop monitoring.

SPAN port monitoring requires a physical server running a Monitoring software process to be connected to the SPAN port of a VLAN on a Catalyst switch where the agent’s phone is installed. The SPAN port is like a broadcast port for all data traffic (including voice RTP streams) traversing a VLAN segment. When a supervisor clicks the silent monitor button on the CSD, it signals to the appropriate Monitoring component to forward a copy of the captured RTP streams for the selected agent to the requesting CSD. The CSD then plays the packets through the sound card on the CSD workstation. No IP Phone (or any type of phone) is involved when the silent monitoring stream is being played using CSD. The CSD can reside anywhere on the Unified Communications network, but the agent’s phone must be on the same VLAN where the SPAN port Monitoring component is installed. The Catalyst switch RSPAN feature allows a VLAN to extend across multiple Catalyst switches. Please refer to Appendix B for more detail on SPAN port monitoring design guidance.

Note

For any deployment in which an agent desktop is the IP Phone Agent or any deployment in which the desktop is a Cisco Agent Desktop or Cisco Supervisor Desktop and the associated phone is not a Cisco IP Phone 7940, 7941, 7960, 7961, or 7970, SPAN port monitoring and recording is the required deployment model.

Desktop monitoring provides a mechanism for the CAD application to obtain a copy of the RTP packet streams directly from the phone and therefore removing the need for a Monitoring component connected to the SPAN port on the Catalyst switch. A 7940 or above phone is required for desktop monitoring and the agent workstation running CAD must be connected to the data port on the back of the agent phone. The IP Communicator (softphone) also supports using desktop monitoring for silent monitoring and recording.

Note

For all deployments in which agents use Cisco Agent Desktop and supervisors use Cisco Supervisor Desktop and the phone deployed is an IP Phone 7940, 7941, 7960, 7961, or 7970, use desktop monitoring.

When a supervisor clicks the silent monitor button on the Cisco Supervisor Desktop for an agent using desktop monitoring, the RTP streams are sent directly from Cisco Agent Desktop to Cisco Supervisor Desktop, and no SPAN port monitoring component is required. However, in order for silent monitoring to occur with desktop monitoring, there must still be at least one Monitoring process running somewhere. The Monitoring process is responsible for setting up the media streams from CAD to CSD for silent monitoring. For desktop monitoring, the agent workstation must have a NIC that supports 802.1Q. This allows the NIC to process packets from both the data and voice VLANs. Appendix C of the Cisco CAD Installation Guide provides a quick and simple test to determine if a workstation NIC will operate properly with the desktop monitoring feature of CAD.

An Unified CCX deployment and individual locations and sites can have a mixture of some agents using desktop monitoring and some agents using SPAN port monitoring.
If an agent call requires recording, then a copy of the RTP packet streams is sent to the Recording Server. If desktop monitoring is being used by the agent being recorded, then CAD sends the RTP streams to the Recording component. If SPAN port monitoring is being used by the agent being recorded, then the Monitoring component (on the VLAN where the agent phone is connected) sends the RTP streams to the Recording component. Agents can be silently monitored and recorded at the same time. When that occurs, CAD or the Monitoring component are sending two copies of the RTP packet streams.

A normal G.7xx VoIP RTP call has two RTP streams (one representing what the agent is hearing and one representing what the agent is saying). These two streams flow in opposite directions across the network. When an agent call is being silent monitored or recorded, both of those RTP streams must be sent. For example, if a supervisor is silent monitoring an agent, two G.7xx RTP streams will be sent from either CAD (desktop monitoring) or the Monitoring component to the CSD. If an agent call is being recorded, two G.7xx RTP streams are sent to the Recording component. If the agent is being silent monitored and recorded, four RTP streams are being sent. This is in addition to the two bi-directional RTP streams of the actual call.

The monitoring and recording packet streams are true G.7xx RTP streams and should be tagged like any other RTP stream to ensure these packets are delivered with appropriate priority and minimal latency. Chapter 6 further discusses bandwidth requirements.

An Unified CCX cluster can have up to five Monitoring components with one of them running on the logical Recording component. The Recording component requires a co-resident Monitoring component (but only one—regardless of whether the Recording component is simplex or redundant). Four additional Monitoring components can be deployed if SPAN port monitoring at remote agent sites is needed.

The agent call recordings are stored on the hard drive of the Recording component server with agent data store locator records pointing to the actual recording files. If a redundant Recording component server is deployed, they operate in a load balancing fashion, and recordings are only stored on the hard drive of the Recording component server that actually received the RTP stream. The call recordings in Unified CCX 4.1 are stored in a raw format that is only playable using the Cisco Supervisor Desktop (CSD) Record Viewer. The CSD Record Viewer shows 7 days worth of call recording as well as those tagged for 30-day extended archiving. The CSD Record Viewer also provides the supervisor the option to save selected individual recordings into a .wav format in a specified folder.

The recording capability of Unified CCX is not intended for usage as a permanent recording archival solution. However, an export utility is also available to bulk export all recordings into a .wav format. The export utility has no ability to specify selected recordings and will export all recordings on a Recording component. System administrators could build their own customized command macros or process that would perform regular (at least weekly) exporting of the recordings for permanent archival of agent call recordings.

With Unified CCX Enhanced, up to 32 simultaneous agent calls can be recorded. With Unified CCX Premium, up to 80 simultaneous agent calls can be recorded. When a supervisor is playing back or saving a recording using the CSD Record Viewer application, a recording resource is used and therefore counts against the maximum simultaneous call recording capacity for the duration of that recording playback. Maximum simultaneous call recording and playback capacity is dependent upon the deployment model and server sizing. If more than 32 simultaneous recordings and playbacks are required, the Recording component must be separated from the CRS Engine and Database components. A dedicated 7845-class server for the Recording process is required for 80 simultaneous recordings and playbacks. The Configuration & Ordering tool can assist you in determining an appropriately sized server for the amount of recording required. See Appendix A, “Server Capacities and Limits” for the full capacity matrix.
Because IP Phone Agent (IPPA) does not include an agent using CAD, IPPA requires a SPAN port Monitor component on the local VLAN segment for silent monitoring or recording. Also the 7902, 7905, and 7912 phones require a SPAN port Monitor component as there are either no data ports on these phones or these data ports are not compatible with desktop monitoring. IPPA also cannot be configured to have calls automatically recorded.

If no agent call recording is required but silent monitoring is required, then a deployment may contain 5 Monitoring components that are for SPAN port monitoring only. Even if all agents will use desktop monitoring, at least one Monitoring component must still be installed. The Monitoring component is responsible for establishing the media stream from CAD to CSD.

Unified CCX Premium is required for remote supervisory monitoring. Remote supervisory monitoring provides a mechanism to silent monitor calls using an IP Phone or PSTN phone. This form of silent monitoring does not require a CSD or any data network connectivity and is ideally suited for management from outsourcer customers of a call center service provider. Agents are unaware when they are being silent monitored using remote supervisory monitoring.

A remote supervisor is configured with a numeric user ID and password and also with the CSQs and agents that the remote supervisor is allowed to silent monitor in this fashion. The remote supervisor then dials a specific number that invokes a Unified CCX application. The application begins by prompting the supervisor for the user ID and password. After the remote supervisor is authenticated, the remote supervisor is prompted on whether they wish to silent monitor calls for a specific agent or for a specific CSQ. Then the Unified CCX application requests a copy of the RTP streams for the selected types of calls, and the Unified CCX application and CTI Port relays those packets to the remote supervisor's phone. Remote supervisory monitoring works with both SPAN port monitoring and desktop monitoring. However, remote supervisory monitoring only works with a CRS Engine and CTI Ports and agent phones using G.711 encoding. Remote supervisory monitoring also places an additional impact on the CRS Engine. This activity is reflected in the Unified CCX 4.1 Configuration & Ordering Tool.

**7920 Wireless IP Phone Support**

Unified CCX supports usage of the 7920 Wireless IP Phone by agents. Agents can be using Cisco Agent Desktop (CAD) with the 7920 phone or agents can use the IP Phone Agent (IPPA) interface with the 7920. When planning to use the 7920 for agents with Unified CCX, the following considerations need to be taken into account:

- If a logged in agent using a 7920 roams outside Wireless Access Point (WAP) range for greater than 60 seconds (possibly slightly longer depending upon Cisco Unified CallManager time out), Cisco Unified CallManager unregisters the 7920 device (and ends any call in progress if the 7920 was off hook). This generates a device unregistered JTAPI event to be sent to Unified CCX which causes the Unified CCX agent state to change to 'not ready.' When agents roam between WAPs, the hand off occurs within a second or two (depending upon wireless LAN design, encryption, and authentication techniques used). Therefore, roaming between WAPs is supported. If an agent is using the 7920 with CAD, but is away from the CAD workstation, there is no way for the agent to know that the agent state is 'not ready' and there is no way for the agent to change the agent state to 'ready.' If the agent is using the 7920 with IPPA, then the agent can check the agent state via IPPA and can change the agent state to 'ready' via IPPA. Therefore, if agents anticipate roaming outside WAP range for greater than 60 seconds, then it is recommended that they log in to Unified CCX via IPPA for that login session. If agents anticipate working at their desk or not roaming outside WAP range, then it is okay for them to login to Unified CCX via CAD for that login session.
- The 7920 is not supported with a shared line for the agent Unified CCX extension as shared lines are not supported with Unified CCX agent extensions.
Citrix and Microsoft Terminal Services Support for CAD

Citrix and Microsoft Terminal Services Support for CAD

Unified CCX supports the running of Cisco Agent Desktop (CAD) within a Citrix or Microsoft terminal services environment. When planning to use terminal services for CAD, the following considerations need to be taken into account:

- The agent must be using a Cisco IP Phone (that is, no softphone)
- Cisco Supervisor Desktop (CSD) and Cisco Desktop Administrator (CDA) are not supported in a terminal services environment.
- Desktop monitoring (for silent monitoring and recording) is not supported with terminal services. SPAN port monitoring must be used.
- Macros work only if they involve applications running on the terminal server, and not those running on the client PC.
- Only one user name is supported per CAD application login.
- The login ID and extension that appear by default in the login dialog box when CAD is started are those associated with the last login by any user.
- The Citrix web client is not supported.
- Supported server platforms for terminal services deployments are Citrix 4.0 running on Windows 2000 SP4, Citrix 4.0 running on Windows 2003 SP1, or Microsoft Terminal Services running on Windows 2003 SP1.
- Supported client platforms for terminal services deployments are Windows 2000 Professional SP4, Windows XP SP1, and Windows XP SP2.

Please reference Integrating CAD into a Citrix MetaFrame Presentation Server Environment for implementation details. This document can be found at:

Unified CCX ASR and TTS

Unified CCX allows integration with Media Resource Control Protocol (MRCP) compliant Automatic Speech Recognition (ASR) and Text-To-Speech (TTS) servers. Nuance and Scansoft are the only ASR and TTS providers whom have been tested and will be supported. ASR and TTS software must be purchased from one of these vendors. These vendors can provide design and server sizing requirements for their software. Cisco no longer re-sells Nuance ASR and TTS as an Unified CCX option.

From CRS Administration, you must configure the address of an MRCP server and the number and type of resources provided by that MRCP server. The MRCP servers (ASR and TTS servers) are not considered to be part of the Unified CCX cluster. Multiple Unified CCX clusters can interact with the same MRCP servers. An Unified CCX cluster can also define multiple MRCP servers, and resources from those servers are selected based upon the system and application configuration.

Calls requiring ASR require the CRS Engine to pass the media stream from the CTI port to the ASR Server. This activity impacts system performance and system sizing. The impact is reflected in the Unified CCX 4.1 Configuration & Ordering Tool.

When using ASR, the ASR resource is allocated at the time of the first step that uses ASR. The ASR resource is then allocated for the duration of the call. When using ASR, you must calculate the required number of ASR resources (ports) similar to the way you calculate any IVR port requirement. You will need the average time the ASR port is used (similar to average call treatment time) and the number of calls using ASR in the busy hour. You can then apply this data to any Erlang-B traffic calculator or other tool to compute the number of ASR resources required. In environments where you have long queue times, it might be economical to transfer the call to another CTI Route Point and pass call data to the second application in order to allow the ASR resource to be released.

For TTS, each 'Generate TTS Prompt' allocates and releases a TTS resource, and the TTS resource is typically only allocated for a couple of seconds and then released (this might vary depending on the application). To determine the number of TTS resources, use the same methodology described above for ASR resources.

A product bulletin is available to help with the configuration of the speech software. It is posted here: http://www.cisco.com/en/US/products/sw/custcosw/ps1846/prod_bulletin0900aecd8033f6c0.html

Unified CCX Integration with Unified ICME Software

Unified CCX can also be implemented as a child ACD of Cisco Intelligent Contact Management Enterprise (Unified ICME) 7.x software. The Unified CCX integration with Unified ICME software requires an IPCC Express Gateway PG process to be co-resident on the Unified CCX server with the CRS Engine. This integration provides the following capabilities:

• The ability for Unified CCX to send agent, queue, and call state changes to the Unified ICME.
• The ability for Unified ICME software to intelligently route and load balance calls across multiple ACD systems which can include one or more Unified CCX systems, Unified CCE systems, or traditional ACDs (that are supported by Unified ICME). Calls routed to a Unified CCX application can also be sent call data so that it can be popped onto an agent’s screen.
• The ability for Unified CCX to send post-route requests with call data to Unified ICME in order to request intelligent routing instruction. This could be in response to a transfer request from an agent or from a step within and Unified CCX application running on a CTI port.
• The ability for Unified ICME to provide multi-site ACD reporting for a mixed network of ACD sites which can include one or more Unified CCX systems, Unified CCE systems, or traditional ACD's.
Unified CCX Integration with Unified ICME Software

- The ability for Unified CCX to send post-route requests with call data to the Unified ICME software in order to request routing instructions. This could be in response to a new call that just arrived at Unified CCX or a call that is being transferred from an IVR port or agent. Call data included in the post-route request can be used by Unified ICME to profile route the call, and call data is also passed to the terminating ACD site (Unified CCX, Unified CCE, or traditional ACD) for an agent screen pop.

**Note**
In a Cisco IPCC Gateway deployment, Unified CCX cannot be co-resident with Cisco Unified CallManager. Cisco Unified CallManager must be installed on a different machine.

**Figure 2-2** Cisco IPCC Gateway Solution with Two Unified CCX Sites

**Figure 2-3** shows another Unified ICME integration deployment scenario. In this scenario, the Unified ICME routes and load balances calls between an Unified CCX 4.1 site, an Unified CCE 7.x site, and a traditional ACD site. Call data for agent screen pop can be passed between these sites via Unified ICME.
In order for Unified CCX to integrate with Unified ICME software, there must be a IPCC Express Gateway PG installed on each server with a CRS Engine. It is not supported for the IPCC Express Gateway PG to run on a separate server when integrating Unified CCX with Unified ICME.

The IPCC Express Gateway PG must be ordered as a part of Unified ICME 7.x software suite. The IPCC Express Gateway PG software is installed from Unified ICME software installation CD—not from the CRS software CD.

**Note**

Partners must have Unified ICME/ Unified CCE ATP status to order and deploy the IPCC Gateway PG with Unified ICME software.

When running in a High Availability Unified CCX deployment, only one IPCC Express Gateway PG process is active. The active PG is the one on the server with the active CRS Engine. Even though there are two IPCC Express Gateway PGs installed, the Unified ICME software views these PGs as a simplex PG because only one is operational at a time.

Running the IPCC Express Gateway PG might reduce the maximum supported number of agents and call processing capabilities of a Unified CCX deployment. The Unified CCX 4.1 Configuration & Ordering Tool can assist solution planners and designers in sizing the hardware required for an Unified CCX deployment.

When Unified ICME routes calls to Unified CCX, it is really routing them to a Cisco Unified CallManager dialed number. Then Cisco Unified CallManager goes through the process of resolving the dialed number association to the CTI Route Point and CRS JTAPI User and offering the call to Unified CCX. Unified CCX then invokes the appropriate script.
Unified CCX Fault Tolerance

The Unified CCX solution offers a number of different capabilities to provide fault tolerance. To begin with, an Unified CCX deployment utilizes a Cisco Unified Communications network composed of Cisco data switches and routers which provide for a highly available data network with many options for redundancy. Cisco campus and network design guides discuss best practices for designing highly available networks with Cisco switches and routers.

A Cisco Unified CallManager deployment utilizes a cluster approach with up to 8 call processing servers per Cisco Unified CallManager cluster. Unified CallManager groups devices (voice gateways, IP Phones, and CTI Ports) into device pools and allows for device pools to have a primary, secondary, and tertiary Cisco Unified CallManager server. When a device pool’s primary Cisco Unified CallManager server fails, the devices within that device pool automatically fail over to the secondary or tertiary Cisco Unified CallManager server. Unified CCX CTI Ports are grouped together into CTI Call Control Groups (often called a CTI Port Group). Each CTI Port Group is configured as part of a device pool. Cisco Unified CallManager also supports Voice Gateways deployed at many locations with trunks from different service providers.

Cisco Unified CallManager has a subsystem called the CTI Manager that abstracts the device management from the JTAPI communications to an application server (like Unified CCX). This allows an application to not be concerned with what specific server a device (voice gateway, agent phone, or CTI port) is currently registered. Unified CCX has the ability to communicate with up to two CTI Managers within a Cisco Unified CallManager cluster, but only actively communicates with one at a time. If the active CTI Manager subsystem or the Cisco Unified CallManager node running the active CTI Manager fails, then Unified CCX closes the sockets for all CTI ports and immediately begins JTAPI communications with the backup CTI Manager. Calls being handled by agents survive, but if their phones are registered with the failed Cisco Unified CallManager, then they will not be able to perform any subsequent call control. Upon completion of existing calls, agent phones will automatically re-register to the secondary Cisco Unified CallManager server. For agents who were not off hook, their phones will re-register to the secondary Cisco Unified CallManager immediately.

In addition to being able to fail over to another Cisco Unified CallManager node within the cluster, Unified CCX itself provides a clustering mechanism. However, Unified CCX clustering is a little bit different than CallManager clustering. In Cisco Unified CallManager clustering, each of the 8 servers in the cluster has an identical database. In Unified CCX clustering, you can have up to 10 servers in an Unified CCX cluster, but only 2 will be servers with Database components. With Unified CCX clustering, configuration changes are made to both databases (assuming both are operational).

The four different components (CRS Engine, Database, Monitoring, and Recording) all provide some level of redundancy and fault tolerance, but each functions a little bit differently.

CRS Engine Redundancy

When deploying with High Availability, two CRS Engine components must be deployed on separate servers. If one server initiates the engine mastership election first, it becomes master. The other server becomes standby. If both servers are started approximately at the same time, it is not specified which server becomes master. When the CRS Engine component server fails over, the standby server becomes the master server and remains as the master server until another failure occurs. Any active calls being processed by applications on CTI Ports will be released upon failure of the master CRS Engine server.
Unified CCX Fault Tolerance

All ACD, IVR and desktop services will failover within 5 seconds. Any incoming call arriving at Cisco Unified CallManager destined for CRS route points can be accepted by the CRS engine and all CRS call treatment and ACD routing services are operational. Automatically logging on large numbers of agents may take up to 1 minute. For a given agent, the ACD is not able to route calls to agents until the automatic login process completes and the agent manually sets the state to 'ready.' Agents on Unified CCX routed calls will see those calls survive and CAD will automatically relog agents back in within one minute, and they will see a visual indicator that a fail over has occurred. After being logged back in, agents will have to set the state to 'ready' when they are ready to begin receiving calls. Agents using IPPA will need to manually relogin to the new master CRS Engine server.

Note
Historical Report generation is done by giving preference to non-Engine master node so that generation of Historical Reports does not impact the CRS Engine performance. In a two-node scenario where the CRS Engine and the Database are collocated and the Engine master is on the Database Publisher node, the Historical Reports are generated on the Database Subscriber node. If the Engine master fails over to the Database Subscriber node, then the Historical Reports are generated on the Database Publisher node.

Database Redundancy

When deploying with High Availability, for Historical Data store, Agent Data store, and Repository Data store, the two servers running the Database components are set up with one being the Publisher and one being the Subscriber. These roles do not change in the event of a failure. If both the Publisher and the Subscriber are up and running, then the server running the Publisher Database component is given DB mastership, where data is written to and read from. If the server running the Publisher Database component is down (or any of the SQL Services such as MSSQLSCRSSQL, Distributed Transaction Coordinator or SQL Agent$CRSSQL on the server with the Publisher Database is down) then the Subscriber is given the DB mastership, where data is written to and read from. SQL Merge Replication replicates the data between the Publisher and Subscriber. If the Subscriber or Publisher is down for less than the retention period (default is 4 days for Hardware with more than 18GB Hard Drive, and it is 2 days for Hardware with smaller 18GB Hard Drive), replication will automatically kick-in to sync data from the Publisher when the Subscriber comes back in service and vice versa. If the Subscriber is down for more than the retention period, the Subscriber has to be reinitialized at off-peak hours from the CRS Administration Dat astore Control Center page.

Under normal call load volume, a latency of 1 to 3 minutes for SQL Merge Replication is expected; this could be higher for higher call load. The impact could be more when Historical Reports are running as it impacts the SQL processing. Due to replication latency, the Historical Reports which get generated from a Subscriber, might not have the latest call records; the Historical Reports will be generated up to the last replicated time.

SQL Server “linked server” technique is used to replicate configuration data store data in High Availability deployments. The way it works is that when both servers with Databases components are operational, configuration data store changes, such as skills and resource groups, are written to both the servers with Database components. If one server with a Database component is down, then configuration data store changes are not possible. However, configurations can be read in CRS Administration; that is, no configuration data store data writes are possible, but data reads are possible when one server with a Database component is down. However, call processing, historical data writing, and call activity reporting can continue even when one Database component is down.

In the case when one of the Database servers is not operational and configuration data store changes are required, you can temporarily "deactivate" the configuration data store component on the off-line Database component server using CRS Administration. After that, you can make configuration data store
changes on the active Database server. Once the off-line Database server is back in service, you can "activate" the configuration data store component on that Database server during off-peak hours as the whole active database configuration data store data will get synchronized.

Network Partitioning

When the network is partitioned (split into two or more islands), every island elects its own set of masters. When the partition is restored, all masters are dropped, therefore all calls receiving call treatment or in queue are dropped, and a new election process is initiated. As a result of this new election, the Database Publisher is elected as a master for the ADS, RDS, and HDS. It is not specified which Engine or CDS becomes the master.

Monitoring and Recording Redundancy

Unified CCX supports up to six servers running the Monitoring component. Five of these servers can provide SPAN port monitoring service. The first Monitoring component must be installed on a server that is running the Recording component. If there is a second server running a Recording Component, a Monitoring component has to be installed on that second server as well, but SPAN port Monitoring is supported only on one of these two servers. These two servers running the Monitoring service are sometimes considered as one “monitoring domain.” The other four monitoring components are typically for SPAN port monitoring for agents at remote sites. When configuring a phone with SPAN port monitoring, only one SPAN port monitoring server can be assigned to this phone. There is no redundancy when desktop Monitoring is used.

When desktop monitoring, CAD forwards the RTP stream to CSD. But a server running the Monitoring component is still required in order to set up the media stream. Any one of the six monitoring servers could be chosen for this purpose. If one or multiple Monitoring components fail, desktop monitoring will still work, as long as one server running the Monitoring component is still available in the CRS cluster.

It is possible to configure and enable both SPAN port monitoring and desktop monitoring for a phone. However, only one method is used at any time for that phone. If both SPAN port monitoring and desktop monitoring are configured correctly, desktop monitoring is chosen. If desktop monitoring fails, SPAN port monitoring is used as a backup. Please refer to the Cisco Desktop Administrator User's Guide for more information.

When deploying with High Availability and agent call recording, two physical servers running Recording components must be deployed. The two physical Recording servers work as a single logical Recording server (a "recording domain") and recording tasks are load balanced in a round robin fashion across the two physical Recording Servers. An Unified CCX deployment only supports one "recording domain." The actual call recordings are stored only on the disk of the physical Recording component server where the recording task took place. Therefore, if a Recording server fails, the Supervisor will be unable to playback those recordings on the failed Recording server until that Recording server is operational again.

The two servers where the Recording components are running also serve as a backup for each other. In order to function properly during a period when one of the servers fails, the two Recording servers must be sized to be capable of supporting all recording for the Unified CCX cluster. For example, under normal operations, a large call center may be setup to handle 40 recording sessions on each Recording component (for a total of 80 simultaneous call recordings). If either server with a Recording component were to fail, the other server would be capable of providing 80 simultaneous call recordings—which no loss in recording capacity. The Unified CCX Configuration & Ordering tool takes into account this type of failover in the sizing of hardware resources for agent call recording.
The two servers with Recording components must be deployed together on the same VLAN segment. Typically these two servers are installed on the same VLAN segment as the CRS and Cisco Unified CallManager servers. In certain scenarios, in an effort to reduce WAN bandwidth, it might make sense for the two servers with the Recording components to be deployed at a remote site with a large number of agents. However, if this is done, then both servers must be deployed at this remote site.

Recording requires a Monitoring component. When SPAN port monitoring is configured for silent monitoring, the SPAN port monitoring server forwards the RTP stream to the Recording component. If that SPAN port monitoring server fails, recording is not possible. When desktop monitoring is configured, the Monitoring component is still required in order to set up the media stream. Any one of the six monitoring servers could be used for this purpose. If one or multiple Monitoring components fail, recording still works, as long as one server running the Monitoring component is still available in the CRS cluster.

**Cold Standby Support**

CRS high availability requires that the CRS Engine and Database components and the CTI Managers with which the CRS servers communicate should be located in the same Campus LAN and the maximum round trip delay between these servers be less than 2 ms.

For disaster recovery deployments where the backup CRS servers need to be in a different geographic location, CRS high availability is therefore not supported. However, for this requirement, it is possible to deploy identically configured cold standby servers in the disaster recovery site. These cold standby servers at the second site should be shut down while the primary servers are in service at the first site. If a disaster occurs and the primary site is down, the standby servers are turned on, restored, and become the active servers.

Cold standby is supported for all Unified CCX deployment models described in Chapter 3. For example, for the ten-server HA deployment model, ten coldstandby servers can be added on the disaster recovery site.

When deploying cold standby, the following rules apply:

- Use the Backup and Recovery System (BARS) tool to backup the primary servers and restore the cold standby server. Follow the backup and restore procedure in the *Cisco CRS Installation Guide* when doing this.
- The deployment in the disaster recovery site needs to be identical to the deployment in the primary site (same number of servers, same server type).
- The standby servers should be shut down if the primary servers are running, and vice versa.
- The servers at the disaster recovery site need to have the same IP addresses and hostnames as their corresponding servers at the primary site during the restore procedure.
- During the restore procedure at the disaster recovery site, the server attempts to write to the LDAP directory. If it cannot write to the LDAP directory, the restore fails.
- The deployment at the disaster recovery site needs to follow the Unified CCX rules for design that are described in Chapter 3, “Cisco Unified Contact Center Express Deployment Models”. For example, when deploying with high availability at each site, the CRS Engine and Database components and the CTI Managers with which the CRS servers communicate must be located in the same Campus LAN and the maximum round trip delay between these servers should be less than 2 ms.
- It might be necessary to modify the Unified CCX configuration. For example, the appropriate region and location for the CTI ports might need to be reconfigured for proper CAC and codec negotiation.
Once the restore procedure is completed, changing IP addresses on the CRS servers is possible. In this case, follow the procedure described in the Cisco CRS Administration Guide to change the IP addresses of the CRS servers. Also update every CAD and CSD desktop by running the postinstall.exe program at C:\Program Files\Cisco\Desktop\bin on the desktop PC.

Upgrading to Unified CCX 4.1

Unified CCX 3.x and 4.0 users can upgrade to Unified CCX 4.1. Unified CCX 2.x users must first upgrade to Unified CCX 3.5 before upgrading to Unified CCX 4.1.

Some 3.x deployments might have the Database component running on an expansion server (separate from the CRS Engine). If you wish to upgrade those deployments to 4.1 and add High Availability, then two additional physical servers must be added. The two additional physical servers need to be equal or better in performance than the existing servers.

There is a limitation on the database size based on system hardware size. For more details, refer to the section "Cisco CRS Disk Space Usage" in the Cisco CRS Installation Guide:


Further, the Pre-upgrade check tool, which can be downloaded from CCO, can be run to find out how much data can be migrated before upgrade. This tool can be found on the same web site as the Configuration & Ordering tool.

Some 3.x deployments might have the Database component and CRS Engine component running on the same physical server. If those deployments wish to upgrade to 4.1 and split the Database component from the CRS Engine component, move the 3.x Database Server to an expansion server first. Then upgrade both physical servers to 4.1, and then add two additional servers for High Availability. Upgrading in this order allows a greater amount of historical call data to be retained.

CRS 4.1 Software Compatibility

CRS software is dependent upon integration with many other software components, especially Cisco Unified CallManager. Please be sure to check to ensure the CRS release you are planning is supported with the Cisco Unified CallManager release for which this deployment is planned. The Cisco CRS Software and Hardware Compatibility Guide can be found at:

Cisco Unified Contact Center Express Deployment Models

As discussed in the previous chapter, when designing a Unified CCX deployment, there are four core software components—CRS Engine, Database, Recording, and Monitoring—for you to consider placing. This chapter discusses the placement of these core software components for the most common deployment models along with rules and considerations for those deployment models. A later chapter discusses usage of the Unified CCX 4.1 Configuration & Ordering Tool to help you determine the number and types of servers required for any supported deployment model and call processing requirements. Prior to using that tool though, it is a good idea to have a pretty close idea of what deployment model you desire.

This chapter contains the following sections:

- Unified CCX General Rules for Design, page 3-3
- Cisco Unified CallManager Co-Resident Deployment Model (1), page 3-4
- Single-Server Non-HA Deployment Model (2), page 3-5
- Multi-Server Non-HA Deployment Model (3), page 3-5
- Two-Server HA Deployment Model (4), page 3-6
- Four-Server HA Deployment Model (5), page 3-6
- Six-Server HA Deployment Model (6), page 3-7
- Ten-Server HA Deployment Model (7), page 3-8
- Other Design Considerations, page 3-8
Figure 3-1 depicts the maximum sized deployment model for Unified CCX 4.1 and will be referred to as the Unified CCX 4.1 Reference Architecture and also the 10-Server HA deployment model. The reference architecture provides high availability and contains 10 physical servers running Unified CCX software components. The reference architecture has 2 CRS Engines, 2 Database Servers, 2 Recording Servers (which requires one of the Monitoring Servers), and 4 additional Monitoring Servers.

There are many supported CRS deployment models. In this document we will discuss the most common deployment models. They are as follows:

1. Cisco Unified CallManager (Unified CM) Co-Resident
2. Single-Server non-HA
3. Multi-Server non-HA
4. Two-Server HA
5. Four-Server HA
6. Six-Server HA
7. Ten-Server HA (reference architecture)

Table 3-1 depicts the placement of the core software components for each of these seven deployment models. Note that many other similar deployment models are possible and supported. These models have no bearing on which specific server model is used. The minimum server model required is identified by the Cisco Unified Contact Center Express Configuration and Ordering Tool. The following sections in this chapter discuss the general rules for design and considerations and limitations for each of the six deployment models listed here. This information allows an Unified CCX system planner or designer to understand what other similar deployment models are supported and to understand how to determine the best solution for a given set of requirements.
Table 3-1  Software Components of Deployment Models

<table>
<thead>
<tr>
<th>Model</th>
<th>Server 1 Unified CCX</th>
<th>Server 2 Unified CCX</th>
<th>Server 3 Unified CCX</th>
<th>Server 4 Unified CCX</th>
<th>Server 5 Unified CCX</th>
<th>Server 6 Unified CCX</th>
<th>Servers 7-10 Unified CCX</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Unified CM Co-resident</td>
<td>Unified CM+Eng+DB+Rec+Mon</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 Single Server non-HA</td>
<td>Eng+DB+Rec+Mon</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Multi-Server non-HA</td>
<td>Eng</td>
<td>DB + Rec + Mon</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Mon x 4</td>
</tr>
<tr>
<td>4 Two-Server HA</td>
<td>Eng+DB+Rec+Mon*</td>
<td>Eng + DB + Rec + Mon*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 Four-Server HA</td>
<td>Eng</td>
<td>Eng</td>
<td>DB + Rec + Mon*</td>
<td>DB + Rec + Mon*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 Six-Server HA</td>
<td>Eng</td>
<td>Eng</td>
<td>DB</td>
<td>DB</td>
<td>Rec+Mon*</td>
<td>Rec+Mon*</td>
<td></td>
</tr>
<tr>
<td>7 Ten-Server HA</td>
<td>Eng</td>
<td>Eng</td>
<td>DB</td>
<td>DB</td>
<td>Rec+Mon*</td>
<td>Rec+Mon*</td>
<td>Mon x 4</td>
</tr>
</tbody>
</table>

Note: The Recording component will always require a co-resident Monitoring component. For redundant servers running Recording components, only one of the 5 supported Monitoring components is required—even though the Monitoring component is installed on both servers running the Recording components. Only one of the Monitoring components on the redundant servers running the Recording components is ever utilized. The asterisk (*) in deployment models 4, 5, 6, and 7 of Table 3-1 relates to this rule that even though there are two Monitoring components installed, only one is counted against the Unified CCX cluster limit of 5 Monitoring components.

Unified CCX General Rules for Design

When designing a Unified CCX deployment, the following rules apply:

- When deploying for high availability, the CRS Engines and Databases servers must be located in the same Campus LAN and the round trip delay between these servers should be less than 2 ms. The links between these servers must be highly available and the available bandwidth should always be considerably higher than the load, and there should be no steady-state congestion. For additional information about LAN and campus infrastructure design, refer to the Cisco Unified Communications solution Reference Network Design (SRND) and to the Designing a Campus Network for High Availability SRND, available at www.cisco.com/go/srnd.
Note

If the network connectivity is lost between the two CRS engines, both CRS engines become active (island mode). When the network connectivity is restored, both CRS servers become inactive for a few seconds in order to negotiate the CRS engine mastership, and all the calls in queues and receiving call treatment would be lost.

- The Unified CM servers running CTI Managers with which Unified CCX will communicate must be located in the same campus LAN as the CRS Engine servers. The available bandwidth for these LAN links should always be considerably higher than the load, and there should be no steady-state congestion, and the maximum round-trip delay between the servers should be less than 2 ms. Other Unified CM nodes within the same Unified CM cluster can be located across the WAN in a Unified CM cluster split over the WAN. In such deployments, the WAN requirements that are described in Cisco Unified Communications SRND must be satisfied.

- When deploying HA, the CRS Engine and Database components must both be redundant.

- If recording is going to be used for a HA deployment, then the Recording component must be redundant.

- All agents for a Unified CCX deployment must be using phones which register to the same Cisco Unified CallManager cluster. Calls can be received from devices and callers on another Cisco Unified CallManager cluster (using inter-cluster trunks).

All Unified CCX deployments must be configured using the Unified CCX 4.1 Configuration & Ordering Tool in order to be supported.

With regard to a couple of key server characteristics, only the 7835 and 7845 models provide redundant power supplies, redundant fans, and redundant hot-swappable disk drives. For any HA deployment, use either the 7835 or the 7845 server model.

Cisco Unified CallManager Co-Resident Deployment Model (1)

Unified CCX software can be installed co-resident with Cisco Unified CallManager software. For co-resident installations, there is a limit of 10 simultaneous agents logged in, 1 supervisor logged in, 1 silent monitoring or recording session, 1 after hours historical reporting client session, and either 10 basic IVR ports or 5 advanced IVR ports. The Unified CCX call volume must not exceed 300 BHCC.

In addition to the 10 IP Phones for agents, the number of non-agent IP Phones supported on the Cisco Unified CallManager server varies by server model. The supported BHCC call volume of non call center traffic on the Cisco Unified CallManager server also varies by server model. The following table lists the supported limits for non call center phone and call volume per server model.

<table>
<thead>
<tr>
<th>Server Model</th>
<th>Max Unified CM IP Phones</th>
<th>Max Unified CM BHCC Call Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>7815</td>
<td>150</td>
<td>900</td>
</tr>
<tr>
<td>7825</td>
<td>500</td>
<td>3,000</td>
</tr>
<tr>
<td>7835</td>
<td>1,000</td>
<td>6,000</td>
</tr>
<tr>
<td>7845</td>
<td>3,000</td>
<td>18,000</td>
</tr>
</tbody>
</table>
Co-resident installations do not support the following:

- Expansion servers
- Unified ICME integration using Unified Contact Center Gateway PG
- High Availability
- Failover to a secondary CTI Manager

If the Cisco Unified CallManager server is part of a multi-node Cisco Unified CallManager cluster, the CTI Manager which Unified CCX communicates with must be running on the same Cisco Unified CallManager node where Unified CCX is installed. All CTI Ports and Route Points must register with this Cisco Unified CallManager node as primary and no secondary is supported.

Install the maximum allowed memory on the Cisco Unified CallManager server. The Cisco Unified CallManager server platforms have particular memory requirements, which are documented in Product Bulletin 2864, *Physical Memory Recommendations for Cisco Unified CallManager, Version 4.0 and Later*, available at:


### Single-Server Non-HA Deployment Model (2)

This deployment model is for small deployments that do not meet the requirements of the Cisco Unified CallManager Co-Resident deployment model and will only require one single processor server or two dual processor servers simultaneous reporting sessions. The 7845 will allow two reporting client sessions during operating hours. All others servers only support one reporting client during operating hours. This deployment model places a single instance of all 4 software components on the same server. This deployment model may use either MSDE (default) or SQL Server as the database. The MSDE database limits the number of simultaneously connected historical reporting clients to five and the max database size is 2 GB.

This deployment model can support silent monitoring and recording for agents at any WAN connected site using desktop monitoring (CAD with a 7940 or higher phone). This deployment model can also support SPAN port monitoring for agents on the same VLAN segment as the Unified CCX server.

This deployment model will allow the CRS Engine to fail over to a backup CTI Manager in the event that the primary CTI Manager fails. CTI Ports and CTI Route Points should be grouped into device pools that have the same primary and secondary server list as that used for JTAPI communications with the CTI Managers.

Up to four additional Monitoring Servers for SPAN port monitoring of agents at remote sites can be added to this model if needed.

### Multi-Server Non-HA Deployment Model (3)

This deployment model is for scenarios requiring more than two simultaneous historical reporting clients during operating hours or more IVR applications, agents, supervisors, monitoring, and recording than could be supported on a single-server deployment or for any deployment requiring ASR or TTS. When any software component is split off onto a separate server from the server with the CRS Engine, we call that server an expansion server.

This deployment model may use either MSDE (default) or SQL Server as the database. The MSDE database limits the number of simultaneously connected historical reporting clients to five and the maximum database size is 2 GB.
This deployment model can support silent monitoring and recording for agents at any WAN connected site using desktop monitoring (CAD with a 7940 or higher phone). This deployment model can also support SPAN port monitoring for agents on the same VLAN segment as the Unified CCX server and four remote sites.

This deployment model allows the CRS Engine to fail over to a backup CTI Manager in the event that the primary CTI Manager fails. CTI Ports and CTI Route Points should be grouped into device pools that have the same primary and secondary server list as that used for JTAPI communications to the CTI Managers.

Two-Server HA Deployment Model (4)

This deployment model is for small to medium-sized contact centers requiring High Availability (HA) and average contact center requirements for supported agents, supervisors, IVR applications, historical reporting, recording, and silent monitoring.

This deployment model incorporates redundant CRS Engine, Database, Recording, and Monitoring components. Both servers and the Unified CM servers that are running the CTI Managers with which Unified CCX will be communicating must be collocated in the same Campus LAN and the round-trip delay between these servers should be less than 2 ms.

SQL Server 2000 is required for HA as SQL Server 2000 replication services are utilized to keep the databases synchronized.

Historical reporting during operating hours is limited to 2 reporting clients on a 7845 and 1 on all other servers.

This deployment model can support silent monitoring and recording for agents at any WAN connected site using desktop monitoring (CAD with a 7940 or higher phone). This deployment model can also support SPAN port monitoring for agents on the same VLAN segment as the Unified CCX server. This deployment model provides redundancy for both recording and silent monitoring for all agents using desktop monitoring (regardless of location) or agents on the local VLAN using SPAN port monitoring.

This deployment model does not incorporate any additional remote Monitoring components, so no SPAN port monitoring for agents at remote sites using IPPA or low end phones can be performed. Up to four SPAN port Monitoring components can be easily added to this deployment model to add that capability.

This deployment model will allow either CRS Engine component to fail over to a backup CTI Manager in the event that the primary server fails. CTI Ports and CTI Route Points should be grouped into device pools that have the same primary and secondary server list as that used for JTAPI communications to the CTI Managers.

Four-Server HA Deployment Model (5)

This deployment model is for medium to large-sized contact centers requiring High Availability and average contact center requirements for historical reporting, recording, and silent monitoring. This deployment model is very similar to the two server HA deployment model, except that by offloading the database, recording, and monitoring components, the CRS Engine is capable of handling more Busy Hour Call Completion (BHCC), agents, supervisors, and IVR self-service applications, and the Database Server is capable of handling more historical reporting clients during operating hours.
This deployment model incorporates redundant CRS Engines running separately from redundant Database Servers, Recording Servers, and Monitoring Servers all running co-resident on a pair of expansion servers. All four servers must be co-located in the same Campus LAN and along with the two Unified CM servers running CTI Manager. The round-trip delay between these servers should be less than 2 ms.

SQL Server 2000 is required for HA as SQL Server 2000 replication services is utilized to keep the databases synchronized.

This deployment model can support silent monitoring and recording for agents at any WAN-connected site using desktop monitoring (CAD with a 7940 or higher phone). This deployment model can also support SPAN port monitoring for agents on the same VLAN segment as the Unified CCX server. This deployment model provides redundancy for both recording and silent monitoring for all agents using desktop monitoring (regardless of location). This deployment model does not incorporate any additional remote Monitoring Servers, so no SPAN port monitoring for agents using IPPA or low end phones at remote sites can be performed. Up to four SPAN port Monitoring Servers can be easily added to this deployment model to add that capability.

This deployment model will allow either CRS Engine to fail over to a backup CTI Manager in the event that the primary server fails. CTI Ports and CTI Route Points should be grouped into device pools that have the same primary and secondary server list as that used for JTAPI communications to the CTI Managers.

**Six-Server HA Deployment Model (6)**

This deployment model is for large contact centers requiring High Availability, a large number of agents, supervisors and high volumes of IVR applications, historical reporting, recording, and silent monitoring. This deployment model is very similar to the two-server and four-server HA deployment models, except that by offloading the database, recording, and monitoring components, the CRS Engine is capable of handling more BHCC, agents, supervisors, and IVR self-service applications. By splitting the database from the recording and monitoring components, increased levels of reporting and recording can be handled. SQL Server 2000 is required for HA, as SQL Server 2000 replication services are utilized to keep the databases synchronized.

This deployment model can support silent monitoring and recording for agents at any WAN-connected site using desktop monitoring (CAD with a 7940 or higher phone). This deployment model can also support SPAN port monitoring for agents on the same VLAN segment as the Recording Servers. This deployment model provides redundancy for both recording and silent monitoring for all agents using desktop monitoring (regardless of location). Because the redundant Recording Servers are separate from the CRS Engine and Database Servers, the redundant Recording Servers could be placed at a remote site separate from the central site where the CRS Engine, Database Servers, and Cisco Unified CallManager servers are running. This might be a good idea when most agents are at this remote site and a high volume of call recordings are being performed. Placing the Recording Servers at the remote site might reduce WAN bandwidth requirements as the recording media streams remain on the remote site LAN.

This deployment model allows either CRS Engine to fail over to a backup CTI Manager in the event that the primary server fails. Group CTI Ports and CTI Route Points into device pools that have the same primary and secondary server list as that used for JTAPI communications to the CTI Managers.
Ten-Server HA Deployment Model (7)

This deployment model is our reference architecture and represents the largest deployment model. This deployment model is for large contact centers requiring High Availability, a large number of agents, supervisors and high volumes of IVR applications, historical reporting, recording, and silent monitoring and agents on up to five different VLANs that will be using SPAN port monitoring. This deployment model is very similar to the two-server and four-server HA deployment models, except that by offloading the database, recording, and monitoring components, the CRS Engine is capable of handling more BHCC, agents, supervisors, and IVR self-service applications. By splitting the database from the recording and monitoring components, increased levels of reporting and recording can be handled. This model also includes four additional SPAN port monitoring servers for agents at remote sites.

SQL Server 2000 is required for HA as SQL Server 2000 replication services is utilized to keep the databases synchronized.

This deployment model can support silent monitoring and recording for agents at any WAN-connected site using desktop monitoring (CAD with a 7940 or higher phone). This deployment model can also support SPAN port monitoring for agents on the same VLAN segment as the Unified CCX servers. This deployment model provides redundancy for both recording and silent monitoring for all agents using desktop monitoring (regardless of location). This deployment model also incorporates four additional remote Monitoring Servers, so SPAN port monitoring for agents using IPPA or low end phones at remote sites can be performed.

Because the redundant Recording Servers are separate from the CRS Engine and Database Servers, the redundant Recording Servers could be placed at a remote site separate from the central site where the CRS Engine, Database Servers, and Cisco Unified CallManager servers are running. This might be a good idea when most agents will be at this remote site and a high volume of call recording will be performed. Placing the Recording Servers at the remote site will reduce WAN bandwidth requirements as the recording media streams will remain on the remote site LAN.

This deployment model will allow either CRS Engine to fail over to a backup CTI Manager in the event that the primary server fails. CTI Ports and CTI Route Points should be grouped into device pools that have the same primary and secondary server list as that used for JTAPI communications to the CTI Managers.

Other Design Considerations

Consider the following when designing your Unified CCX 4.1 system:

- High Availability requires additional disk space so historical call reporting capacity (maximum retention period) may be reduced. Historical call reporting capacity is also dependent upon BHCC, hours of operation per day, and days of operation per week. The Configuration & Recording tool will estimate the number of weeks of historical call reporting that may be stored for different deployment models. Moving the database to an expansion server will increase the call reporting capacity. In some very high volume large agent count deployments, it may no longer be possible to retain a full 13 months (which was assured in 3.x releases) of historical call reporting data.

- G.711 call recording requires about 1MB per minute. G.729 call recording requires about 256KB per minute. If significant amounts of call recording are planned, it might make sense to split the Recording component(s) from the CRS Engine component(s) and Database component(s) to ensure adequate disk space is available.

- Five categories of data utilize hard disk space. They are as follows:
  - Windows Server OS, CRS Software, and SQL Server Database Management Software
– CRS Logs
– The CRS Database (comprised of 4 data stores)
– BARS (Backup and Recovery System) data
– Recording Files.

Systems planners and designers should attempt to estimate the impact of each in order to determine hard disk requirements. The Cisco Customer Response Solutions Installation Guide provides more information about disk size requirements for very large installations.

• The Cisco Unified CallManager sizing tool assumes devices are evenly distributed across all servers within a Cisco Unified CallManager Cluster. Because it is recommended that ALL CTI Ports and CTI Route Points be configured as part of a device pool that homes primarily to the same Cisco Unified CallManager Server as the primary CTI Manager being used, it may be required to run the Cisco Unified CallManager sizing tool on a per location or per server basis.

• The Cisco Unified CallManager QSIG path replacement feature is not supported for Unified CCX calls. For additional information about Unsupported Features in Cisco Unified CallManager, see the 4.1x release notes for Cisco Customer Response Solutions available on line at http://www.cisco.com/en/US/products/sw/customosw/ps1846/prod_release_notes_list.html

• Cisco Unified CallManager Forced Authorization Codes and Client Matter Codes are not allowed to be used on the same Cisco Unified CallManager cluster where Unified CCX is installed. This is for any Cisco Unified CallManager phones and users—not just Unified CCX agents. The resolution to this is to deploy two separate Cisco Unified CallManager clusters. For details on why this is not supported, see the Cisco CRS Release Notes available on line at: http://www.cisco.com/en/US/products/sw/customosw/ps1846/prod_release_notes_list.html

• The Unified CCX Gateway PG can be added to any of these models except for Model 1, the CallManager co-resident model. The PG process must run on the server(s) with the CRS Engine component(s).

• Sometimes new releases of Cisco Unified CallManager will not support Unified CCX immediately at Cisco Unified CallManager first customer ship (FCS) time. Some organizations may be early adopters of new Cisco Unified CallManager releases and may be slowed from migrating to new Cisco Unified CallManager releases and using new Cisco Unified CallManager features if Unified CCX is installed with that same Cisco Unified CallManager cluster. Therefore, in some scenarios, it may make sense to have a separate Cisco Unified CallManager cluster for Unified CCX.

• Configurations with a remote IVR with Cisco Unified CallManager are generally not supported except on a case-by-case basis using the bid assurance process.
Basics of Call Center Sizing

This chapter introduces the basic concepts involved in call center sizing. This chapter contains the following sections:

- **Terminology**, page 4-1
- **Preliminary Information Requirements**, page 4-2
- **Principal Design Considerations for Call Center Sizing**, page 4-4
- **Planning Resource Requirements for Call Center Sizing**, page 4-5.

**Terminology**

Figure 4-1 illustrates the common port types and how they map to Unified CCX.

**Figure 4-1**  Call Center Port Types

Call center sizing differentiates the port types as follows:

- **Gateway or PSTN trunk ports** — handle calls originating from the PSTN. They are purchased separately from Unified CCX.
**Queue ports** — are IVR ports that queue calls (when no agents are available) prior to transferring the caller to an available agent. These ports are included at no additional cost with Unified CCX Standard or Enhanced, but they must be sized for proper capacity planning for the Unified CCX server. Refer to the Cisco Unified Contact Center Express Configuration and Ordering Tool for more details.

**IVR ports** — are full-featured IVR ports with all the capabilities found in the standalone Cisco Unified IP IVR product, except that the Unified CCX IVR ports require Unified CCX Premium and do not support Cisco Unified Intelligent Contact Management Enterprise (Unified ICME) integration.

If you want additional supporting features, such as automatic speech recognition (ASR), text-to-speech (TTS), e-mail notification, web server or client functionality, and database operations, you simply need to purchase the Premium package. Additional seats may also be purchased for IVR port licenses if the number of port licenses that come with the seat licenses is not sufficient.

The goal of the system architect is to determine the appropriate number and types of IVR ports to provision for the Unified CCX system. However, as shown in Figure 4-1, the Unified CCX architecture differs slightly from the example TDM call center configuration in that IVR ports and queue ports (and P&C ports as well) are combined into one logical CTI port. Therefore, the call sizing approach in this document calculates trunk, IVR, and queue ports. The remaining sections of this chapter use the term **IVR port** to denote the combined queue port and IVR port (both full-service and P&C ports).

**Preliminary Information Requirements**

System designers are advised to create a sizing document to do the following:

- Scope out the preliminary configuration information for the Unified CCX server.
- Size the gateways for the system.

To determine the size of the call center, obtain answers to the following questions:

- How many IVR ports do you need?
- How many PSTN gateway trunk ports do you need?
- How many agents will answer incoming calls?

To answer these questions properly, you will need the sizing metrics and information listed in Table 4-1.

**Table 4-1  Call Center Sizing Metrics**

<table>
<thead>
<tr>
<th>Metric</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average handle time (AHT)</td>
<td>Average duration (talk time) of a call plus after-call work time, which is</td>
</tr>
<tr>
<td></td>
<td>the wrap-up time after the caller hangs up.</td>
</tr>
<tr>
<td>Average IVR port usage time</td>
<td>The total time for prompt playout and/or menu navigation (if any) in the</td>
</tr>
<tr>
<td></td>
<td>Unified CCX script. This should not include the queue time the caller</td>
</tr>
<tr>
<td></td>
<td>spends waiting in queue before an agent becomes available. Queue time is</td>
</tr>
<tr>
<td></td>
<td>calculated using Erlang-C automatically as shown in Figure 4-2.</td>
</tr>
<tr>
<td>Service level goal for agents</td>
<td>Percentage of calls answered by agents within a specific number of seconds.</td>
</tr>
</tbody>
</table>
All of the metrics in Table 4-1 are basic call sizing metrics. Once this information is obtained, the number of gateway trunk ports, IVR ports, and agents can be calculated using the IPC Resource Calculator available at: http://tools.cisco.com/partner/ipccal/index.htm.

The IPC Resource Calculator uses Erlang C for sizing agents, and Erlang B for sizing IVR ports. The output of this sizing process will provide you with the total number of Gateway trunk ports, IVR ports and total number of agents to size the Unified CCX system properly.

See Figure 4-2 for an overview of the IP call center sizing process, and see the section on Planning Resource Requirements for Call Center Sizing, page 4-5, for detailed sizing information for both IVR ports and Unified CCX agents.

Note

If the system being designed is a replacement for an existing ACD or an expansion to an installed Unified CCX or Unified IP IVR system, you might be able to use the historical reporting information from the existing system to arrive at the above metrics.

In addition, call sizing design considerations may vary if the call center is more self-service oriented.
Principal Design Considerations for Call Center Sizing

Figure 4-2 illustrates the principal steps and design considerations for sizing a call center.

**Figure 4-2**  Unified CCX Design Process – Call Center Sizing

Figure 4-2 is a general overview of the design considerations for call sizing. For a detailed description of the call center sizing design process, refer to the section on sizing call center resources in the Unified Contact Center Enterprise Solution Reference Network Design (SRND) for Cisco Unified Contact Center Enterprise, available online at http://www.cisco.com/warp/public/779/largeent/it/ese/unifiedCommunications.html

There are similar basic call center sizing considerations and steps for Unified CCE, and they also can be used in sizing a smaller contact center for Unified CCX. This call sizing approach will provide you with the minimum number of IVR ports to support the total BHCA.
In addition, you should include the following design considerations, specific to Unified CCX, in your call center sizing calculations:

- At a minimum, plan on enough capacity to replace your existing system. The replacement system should perform at least as well as the one it is replacing.

- After all of the Erlang (C and B) calculations are complete for the call center sizing, any changes in queue times or agents will affect the total number of trunks and IVR ports required for an Unified CCX solution.

- As you increase the size of the agent pool, very small changes in the average queue time and percentage of queued calls will affect the required number of gateway trunks and IVR ports.

- Even if you perform all of the calculations for a call center, there are still some variables that you cannot plan for but that will affect the ports needed on an Unified CCX system. For example, one or more agents could call in sick, and that would affect the port count and queue time for each call. Just two agents calling in sick could increase the port count by over 12%. This would affect the price of the system and, if not planned for, would affect the ability of the call center to meet caller requirements. Properly sizing call center resources is integral to designing an effective Unified CCX system.

Note: Not all of the Unified CCX system limits are available at the same time.

If all of the call sizing information is available, the next step is to apply Unified CCX sizing limits to the call center requirements. For this step, use the Cisco Unified Contact Center Express Configuration and Ordering Tool, available online at


Planning Resource Requirements for Call Center Sizing

To assist you with planning resource requirements, this section illustrates how to size an Unified CCX Standard application with 25 agents.

Example of Sizing Unified CCX Standard Application with 25 Agents

This example is not intended to be a comprehensive contact center design example, but it illustrates how changing metrics such as BHCA, AHT, and Service Levels can affect provisioning of agents.

The following information applies to this example of Unified CCX Standard with 25 agents:

<table>
<thead>
<tr>
<th>Metric</th>
<th>Metric Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Busy Hour Call Attempts (BHCA)</td>
<td>800 calls in 60-minute interval</td>
</tr>
<tr>
<td>Service level goal</td>
<td>90% of all calls handled within 15 seconds</td>
</tr>
<tr>
<td>Average handle time (AHT)</td>
<td>90 seconds:</td>
</tr>
<tr>
<td></td>
<td>• Average talk time = 90 seconds</td>
</tr>
<tr>
<td></td>
<td>• Wrap-up time = 0 seconds</td>
</tr>
<tr>
<td>Wait before Abandon</td>
<td>120 seconds</td>
</tr>
<tr>
<td>Grade of service (% blockage) for gateway ports to the PSTN</td>
<td>1% (0.01)</td>
</tr>
</tbody>
</table>
Using the IPC Resource Calculator available (http://tools.cisco.com/partner/ipccal/index.htm), we can
determine that 25 agents are needed for this system. Checking the Cisco Unified Contact Center Express
Configuration and Ordering Tool indicates that all of these parameters fit within a single-server Unified
CCXsystem.

Figure 4-3 provides a basic example of the IPC Standard Resource Calculator.

The IPC Resource Calculator also uses Erlang B and C to calculate the number of IVR ports needed
for call treatment (prompt and collect) and queuing. An example of this is the default icd.aef script
logic that is available with all the Unified CCX packages. Note in Figure 4-4 how the script logic allows the
application developer to insert various delays in the script; these delays must be included in Average Call
Treatment Time, (IVR) input to the IPC Resource Calculator.
Figure 4-4  Application Processing Time for Unified CCX

The following steps detail the procedure for calculating IVR ports for our example Unified CCX application:

Step 1  Calculate the number of IVR ports required to handle IVR call treatment functionality:

   a. Estimate the average time the call is being processed by the Unified CCX script, from the time the initial call enters the application until the time the call is queued. This value is the call treatment time (CTT, also called Average IVR Delay). Using the default icd.aef script for our example, this value would be the time the welcome prompt is played. The welcome prompt used by this particular Unified CCX application was estimated at two seconds. (Note that a lengthy prompt/collect sequence for caller self-service will result in much longer CTT).

   b. Now enter the CTT (Average IVR Delay) of 20 seconds into the IPC Resource Calculator, and notice that in this example ten IVR ports are required for call treatment.

Step 2  Calculate the number of IVR Ports required to handle queuing functionality.

   In this case the IPC Resource Calculator has already performed the calculation from the previous inputs, yielding a value of five IVR ports required for queuing.

Step 3  Calculate the total number of IVR Ports required.

   The IPC Resource Calculator automatically adds up all IVR ports required (queuing, call treatment and self service (using the advanced IPC Resource Calculator). In this example a total of fifteen IVR ports are required.

   Note at this point that the IPC Resource Calculator has also determined the number of Gateway Voice Trunks needed to support the required number of Agents and IVR ports. In this example, 36 PSTN trunks (DS0’s) are required.
**Figure 4-5 Standard Resource Calculator Call Treatment Example**

<table>
<thead>
<tr>
<th>Project Identification: Call Treatment Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calls per interval (BHCA): 60 min ▼ 800 calls</td>
</tr>
<tr>
<td>Service Level Goal (SLG): 90% ▼ within 15 sec</td>
</tr>
<tr>
<td>Avg call talk time: 90 sec</td>
</tr>
<tr>
<td>Avg after call work time: 0 sec</td>
</tr>
<tr>
<td>Avg handle time (Agent calls): 90 sec</td>
</tr>
<tr>
<td>Avg Call treatment Time (IVR): 20 sec</td>
</tr>
<tr>
<td>Wait before abandon (Tolerance): 120 sec</td>
</tr>
<tr>
<td>Blockage % (PSTN Trunks): 1% of calls lost (Busy)</td>
</tr>
</tbody>
</table>

**Recommended Agents: 25**

- Calls completed (BHCC): 792 calls
- Calls answered within SLG: 92% within 15 sec
- Calls answered beyond SLG: 8% beyond 15 sec
- Queued calls: 19.3% 152 Q Calls 0.8 Erlangs
- Calls answered immediately: 80.7% 639 calls
- Avg Queue Time (AQT): 17 sec 0m 17s
- Avg Speed to Answer (ASA): 3 sec 0m 3s
- Avg call duration: 113 sec 1m 53s
- Agents utilization: 79%
- Calls exceeding Abandon Tolerance: 0% 0 Calls
- PSTN Trunk Utilization: 69%
- Voice trunks required: 36 Trunks T1/PRI ▼ 1.7 T1/PRI
- IVR ports required for queueing: 5 IVR Ports
- IVR ports required for call treatment: 10 IVR Ports
- Sum of Required IVR Ports: 15

Note that changes in BHCA, CCT, and service level will affect the overall number of ports and agents required in a call center. Each increase or decrease in call handling time will affect the number of ports much more dramatically than in a smaller call center.

Sizing Cisco Unified Contact Center Express and Cisco Unified Communications Manager Servers

This chapter helps you size the Cisco Unified Contact Center Express (Unified CCX) Server and the Cisco Unified Communications Manager Server. This chapter contains the following sections:

- Cisco A2Q Bid Assurance Requirements, page 5-1
- Sizing Tools, page 5-1
- Affect of Performance Criteria on the Unified CCX Server, page 5-2
- Impact of Performance Criteria on the Unified CM Server(s), page 5-2

Cisco A2Q Bid Assurance Requirements

The Assessment to Quality (A2Q) process is a Cisco design review and deployment assessment initiative that identifies and solves solution issues before a deal gets booked. Cisco Customer Contact Business Unit (CCBU) requires that all new Unified CCX deals be submitted to the A2Q Contact Center team. Please note the following requirements:

- The Cisco A2Q process must be followed for every Unified CCX deployment.
- Every Unified CCX deployment must use the Unified CCX Configuration & Ordering Tool. The tool will either automatically bid assure a configuration or will flag that a manual bid assurance review is required.
- Every Unified CCX configuration must be bid assured prior to making a final offer to a customer.

Sizing Tools

The Unified CCX Configuration and Ordering Tool for Unified CCX and Unified IP IVR 5.0 must be used to size Unified CCX and Unified IP IVR systems. Also, the Unified CM Capacity Tool (CMCT) is the only approved tool and must be used to properly size the Unified CM server(s). Before sizing the servers, first familiarize yourself with the on-line Help and frequently asked questions (FAQs) of the tools before using them to size your systems.
Affect of Performance Criteria on the Unified CCX Server

System performance criteria fall into two general categories:

- Unified CCX and Unified IP IVR components - Applications, SW versions, capabilities, server types, and options and quantities that your system requires.
- System usage - The average number of calls placed and received per hour, the average call length, the scripts being executed, grammar used for ASR, and so forth.

Effect of Performance Criteria

Each performance criterion can have an effect on the performance of the Unified CCX or Unified IP IVR system. In general, the more Unified CCX or Unified IP IVR components that you install and the heavier the system usage, the higher the demand on the server. However, the performance criteria can also interact in various non-linear ways to affect performance. The Configuration and Ordering Tool for Unified CCX and Unified IP IVR 5.0 can help you see and evaluate the effects of performance criteria on the Unified CCXs and Unified IP IVR server.

Impact of Performance Criteria on the Unified CM Server(s)

Similarly, Unified CM system performance is influenced by many criteria such as:

- Software release versions— Using the capacity tool, make sure to select the Cisco Unified CM software version with which Unified CCX will be working. Unified CCX 5.0 works only with Unified CM 5.1 and 6.0.
- The type and quantity of devices registered
  - CTI ports (IP IVR ports for queuing, call treatment and self service)
  - Gateway (GW) ports
  - Agent phones

Note

Cisco recommends that you count outbound agents as inbound agents when performing sizing in CMCT.

For deployments with more than 50 agent phones, Cisco strongly recommends that you deploy a minimum of two subscriber servers and a combined TFTP publisher.


Only the 5.0 version of the Unified CCX Configuration and Ordering Tool is valid for Unified CCX 5.0 configurations.

The Unified CM Capacity Tool is available online for partner access to account for the capacity required in Unified CM servers to handle CTI and call processing for Unified CCX.

The capacity tool is available online at:
http://www.cisco.com/partner/WWChannels/technologies/resources/CallManager/
- Route points, and so forth
- The load processed by these devices (calls per second)
- Application call flows
  - IVR self-service
  - Call treatment/Prompt and collect
  - Routing to agents, % transfers and conferences
- Special Unified CM configuration and services
  - Other non-Unified CCX devices—IP phones, GW ports, Unity ports, dial plan, and so forth.
  - Music on Hold (MOH)
  - Tracing levels—Unified CM CPU resource consumption varies depending on trace level enabled. Changing trace level from Default to Full on Unified CM can increase CPU consumption significantly under high loads. Changing tracing level from Default to No tracing can also decrease CPU consumption significantly at high loads (this is not a recommended configuration and would not be supported by Cisco TAC). CPU consumption due to default trace will vary based on load, Unified CM release, applications installed; call flow complexity, and so on.
- Server platform type
Bandwidth, Security, and QoS Considerations

This chapter presents some design considerations for provisioning network bandwidth, providing security and access to corporate data stores, and ensuring Quality of Service (QoS) for Unified CCX applications.

This chapter contains the following sections:

- Estimating Bandwidth Consumption, page 6-1
- Serviceability and Security, page 6-12
- QoS and Call Admission Control, page 6-13

Estimating Bandwidth Consumption

Bandwidth plays a large role in deployments involving:

- The centralized call processing model (Unified CCX at the central site)
- Any call deployment model that uses call admission control or a gatekeeper

Remote Agent Traffic Profile

Remote Agent Traffic Profile Unified CCX signaling represents only a very small portion of control traffic (Agent/Supervisor Desktop to and from the CRS Server) in the network. For information on TCP ports and Differentiated Services Code Point (DSCP) marking for Unified CCX ICD and CTI traffic, see the sections on Serviceability and Security, page 6-12, and QoS and Call Admission Control, page 6-13.

Bandwidth estimation becomes an issue when voice is included in the calculation. Because WAN links are usually the lowest-speed circuits in an IP Telephony network, particular attention must be given to reducing packet loss, delay, and jitter where voice traffic is sent across these links. G.729 is the preferred codec for use over the WAN because the G.729 method for sampling audio introduces the least latency (only 30 msecs) in addition to any other delays caused by the network.

Where voice is included in bandwidth, system architects should consider the following factors:

- Total delay budget for latency (taking into account WAN latency, serialization delays for any local area network traversed, and any forwarding latency present in the network devices). The generally agreed-upon limit for total (one-way) latency for applications in a network is 150 milliseconds.
Estimating Bandwidth Consumption

- Impact of delays inherent in the applications themselves. 8 seconds is the average Unified CCX agent login time with no WAN delay. This includes the exchange of approximately 1,000 messages between the agent application and various servers. The overall time to log in agents increases by approximately 30 seconds for each 30 milliseconds of WAN delay.
- Impact of routing protocols. For example, Enhanced Interior Gateway Routing Protocol (EIGRP) uses quick convergence times and conservative use of bandwidth. EIGRP convergence also has a negligible impact on call processing and Unified CCX agent logins.
- Method used for silently monitoring and recording agent calls. The method used dictates the bandwidth load on a given network link.

Silent Monitoring Bandwidth Usage

The silent monitoring feature of the CAD desktop software, which includes both listening to and recording agent calls, has the largest bandwidth requirements for the CAD product. Properly configuring this feature is especially important for remote agents who are connected to the main site by a WAN connection.

An agent's call can be listened to or recorded by the CAD software. To do this, a request is sent to a VoIP provider. The VoIP provider, captures the voice streams representing the call (two voice streams per call) and sends them back to the requestor. The bandwidth requirements detailed in this section are for the network links between the requestor and provider.

Silent Monitoring Requestors

There are two possible requestors in the CAD software:
- Cisco Supervisor Desktop
- Recording service

Cisco Supervisor Desktops will send requests when the supervisor wishes to listen to an agent's call in real-time. The VoIP provider will capture the voice streams and send them back to the supervisor's desktop where they can be listened to over the desktop's speakers.

A Recording service will send requests when either a supervisor or agent wishes to record the call. The VoIP provider will send the voice streams and the Recording service will save the streams to disk so they can be listened to at a later time.

An installation may have one or two Recording services. An off-board Recording service may be installed at a remote office location, if all Agents and all Supervisors are at that remote location.

Silent Monitoring Providers

There are also two possible VoIP providers in the CAD software:
- Cisco Agent Desktop
- VoIP Monitor service

The Cisco Agent Desktop application contains a service referred to as the Desktop Monitor service that runs on the agent's desktop. It is responsible for processing silent monitoring requests only for the agent logged into the CAD application on the desktop. It captures voice packets sent to the IP or soft phone associated with the logged in agent. The IP phone must be connected in series with the agent desktop on the network for this to work.
By default, this service is active on all agent desktops when the application is started. After initial installation of the CAD servers, all agents are already configured to use the Desktop Monitor service for the silent monitoring feature.

A VoIP Monitor service is able to handle multiple requests for silent monitoring simultaneously. It captures packets directly from the switch via the switch’s Switched Port Analyzer (SPAN) configuration. An installation may have up to four VoIP Monitor services on different machines. Off-board VoIP services may be installed at remote office locations. In some instances, this may be required due to network complexity and capacity planning.

Note

IP Phone agents, who don’t have a desktop, must be configured to use a VoIP Monitor service for the silent monitoring feature.

Figure 6-1 shows a representative Unified CCX installation supporting a remote office over a WAN. Both the main office and the remote office have a VoIP and Recording service on site.

Figure 6-1 Contact Center Representation

It is easy to see where the bandwidth will be required for the silent monitoring feature when you can locate the requestors and providers. There are some notes of interest regarding bandwidth that are shown in the figure:

- Although an administrator is able to assign a specific VoIP service to an agent device, the Recording service that is used is determined at the time the request is made. This applies if two Recording services were installed in order to load-balance the installation. This may result if the provider and requestor being separated by a WAN and requiring the bandwidth on the WAN.
If the VoIP provider is a VoIP Monitor service, the requestor is a Recording service, and these services reside on the same machine, there is no additional bandwidth used on the network to record the call.

Regardless of who the requestor and VoIP provider are, the bandwidth requirement between these two points is the bandwidth of the IP call being monitored and/or recorded. You can think of each monitoring and/or recording session as being a new phone call (2 voice streams) for calculating bandwidth. Therefore, to calculate bandwidth to support the Silent Monitoring feature, you can use the same calculations used to provisioning the network to handle call traffic.

**IP Call Bandwidth Usage**

An IP phone call consists of two streams of data. One stream is sent from phone A to phone B. The other stream is sent from phone B to phone A. The voice data is encapsulated into packets that are sent over the network. The amount of data required to store a voice stream is dependent upon the CODEC used to encode the data. The CAD software can support both the G.711 and G.729 CODEC.

The voice data itself is transmitted over the network using the Real-Time Transport Protocol (RTP). The RTP protocol supports the idea of silence suppression. When silence suppression is used, no voice packets are sent over the network if there is no sound. Otherwise, even packets that contain silence are sent. This lowers the average required bandwidth for a call. Although CAD supports silence suppression, the lower bandwidth requirements for silence suppression should not be used when provisioning the network because the worst case scenario would be where there is no silence in the call, requiring the full call bandwidth as if silence suppression was not enabled.

When calculating bandwidth for an IP call, you must use the size of the RTP packet plus the additional overhead of the networking protocols used to transport the RTP data through the network.

For example, G.711 packets carrying 20 ms of speech data require 64 kbps (kilobytes per second) of network bandwidth per stream. These packets are encapsulated by four layers of networking protocols (RTP, UDP, IP, and Ethernet). Each of these protocols adds its own header information to the G.711 data. As a result, the G.711 data, once packed into an Ethernet frame, requires 87.2 kbps of bandwidth per data stream as it travels over the network. Since an IP phone call consists of two voice streams, in this example, a call would require 174.4 kbps.

The amount of voice data in a single packet also influences the size of the packet and bandwidth. The example above used packets containing 20 milliseconds of speech for its calculations, but this value can be changed in the Cisco Unified CallManager configuration for each supported CODEC. Configuring packets to contain more speech information reduces the number of packets sent over the network and reduces the bandwidth since there are fewer packets containing the additional networking headers, but the packet sizes increase. Table 6-1 shows the bandwidth required for a phone call for the different combinations of CODEC and amount of speech per packet.

<table>
<thead>
<tr>
<th>CODEC</th>
<th>Milliseconds of speech per packet</th>
<th>Bandwidth required (Kbps) for a call</th>
</tr>
</thead>
<tbody>
<tr>
<td>G.711</td>
<td>10</td>
<td>220.8</td>
</tr>
<tr>
<td>G.711</td>
<td>20</td>
<td>174.4</td>
</tr>
<tr>
<td>G.711</td>
<td>30</td>
<td>159.0</td>
</tr>
<tr>
<td>G.729</td>
<td>10</td>
<td>108.8</td>
</tr>
<tr>
<td>G.729</td>
<td>20</td>
<td>62.4</td>
</tr>
<tr>
<td>G.729</td>
<td>30</td>
<td>47.0</td>
</tr>
</tbody>
</table>
Estimating Bandwidth Consumption

Table 6-1 Per-Call Packet Size Bandwidth Requirements (continued)

<table>
<thead>
<tr>
<th>CODEC</th>
<th>Milliseconds of speech per packet</th>
<th>Bandwidth required (Kbps) for a call</th>
</tr>
</thead>
<tbody>
<tr>
<td>G.729</td>
<td>40</td>
<td>39.2</td>
</tr>
<tr>
<td>G.729</td>
<td>50</td>
<td>34.6</td>
</tr>
<tr>
<td>G.729</td>
<td>60</td>
<td>31.4</td>
</tr>
</tbody>
</table>

Note: These calculations are based on G.711 using a sampling rate of 64 kbps speech encoding and the G.729 using 8 kbps. This means one second of speech encoded into the G.711 CODEC requires 65,536 bits (or 8,192 bytes) to represent one second of sound.

For full-duplex connections, the bandwidth speed applies to both incoming and outgoing traffic. (For instance, for a 100-Mbps connection, there is 100 Mbps of upload bandwidth and 100 Mbps of download bandwidth.) Therefore, an IP phone call consumes the bandwidth equivalent of a single stream of data. In this scenario, a G.711 IP phone call with no silence suppression and containing 20 milliseconds of speech per packet requires 87.2 kbps (174.4 / 2) of the available bandwidth.

Table 6-2 and Table 6-3 display the percentage of total bandwidth available, based on the network connection, that is required for simultaneous monitoring sessions handled by a VoIP provider.

The following notes apply to the bandwidth requirements shown in Table 6-2 and Table 6-3:

- The bandwidth values are calculated based on the best speed of the indicated connections. A connection’s true speed can differ from the maximum stated due to various factors.
- The bandwidth requirements are based on upload speed. Download speed affects only the incoming stream for the IP phone call.
- The values are based upon each voice packet containing 20 milliseconds of speech.
- The number of bytes in each packet include the entire Ethernet encapsulation.
- The data represents the CODECs without silence suppression. With silence suppression, the amount of bandwidth used may be lower.
- The data shown does not address the quality of the speech of the monitored call. If the bandwidth requirements approach the total bandwidth available and other applications must share access to the network, latency (packet delay) of the voice packets can affect the quality of the monitored speech. However, latency does not affect the quality of recorded speech.
- The data represents only the bandwidth required for monitoring and recording. It does not include the bandwidth requirements for other Cisco Agent Desktop modules as outlined in other sections of this document.

Table 6-2 Available Upload Bandwidth Percentage for Simultaneous Monitoring Sessions with G.711 CODEC

<table>
<thead>
<tr>
<th>Number of Simultaneous Monitoring Sessions</th>
<th>Percentage of Available Bandwidth Required (No Silence Suppression)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>100 Mbps</td>
</tr>
<tr>
<td>Call only</td>
<td>0.1</td>
</tr>
<tr>
<td>1</td>
<td>0.3</td>
</tr>
<tr>
<td>2</td>
<td>0.4</td>
</tr>
<tr>
<td>3</td>
<td>0.6</td>
</tr>
</tbody>
</table>

¹ Not supported (NS)
Estimating Bandwidth Consumption

Chapter 6  Bandwidth, Security, and QoS Considerations

Table 6-2  Available Upload Bandwidth Percentage for Simultaneous Monitoring Sessions with G.711 CODEC

<table>
<thead>
<tr>
<th>Number of Simultaneous Monitoring Sessions</th>
<th>Percentage of Available Bandwidth Required (No Silence Suppression)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>100 Mbps</td>
</tr>
<tr>
<td>4</td>
<td>0.8</td>
</tr>
<tr>
<td>5</td>
<td>1.0</td>
</tr>
<tr>
<td>6</td>
<td>1.1</td>
</tr>
<tr>
<td>7</td>
<td>1.3</td>
</tr>
<tr>
<td>8</td>
<td>1.5</td>
</tr>
<tr>
<td>9</td>
<td>1.7</td>
</tr>
<tr>
<td>10</td>
<td>1.8</td>
</tr>
</tbody>
</table>

1. The bandwidth of the connection is not large enough to support the number of simultaneous monitoring sessions.

Table 6-3  Available Upload Bandwidth Percentage for Simultaneous Monitoring Sessions with G.729 CODEC

<table>
<thead>
<tr>
<th>Number of Simultaneous Monitoring Sessions</th>
<th>Percentage of Available Bandwidth Required (No Silence Suppression)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>100 Mbps</td>
</tr>
<tr>
<td>Call only</td>
<td>0.0</td>
</tr>
<tr>
<td>1</td>
<td>0.1</td>
</tr>
<tr>
<td>2</td>
<td>0.2</td>
</tr>
<tr>
<td>3</td>
<td>0.2</td>
</tr>
<tr>
<td>4</td>
<td>0.3</td>
</tr>
<tr>
<td>5</td>
<td>0.3</td>
</tr>
<tr>
<td>6</td>
<td>0.4</td>
</tr>
<tr>
<td>7</td>
<td>0.5</td>
</tr>
<tr>
<td>8</td>
<td>0.5</td>
</tr>
<tr>
<td>9</td>
<td>0.6</td>
</tr>
<tr>
<td>10</td>
<td>0.7</td>
</tr>
</tbody>
</table>

1. The bandwidth of the connection is not large enough to support the number of simultaneous monitoring sessions.

Bandwidth Requirements for VoIP Monitor Service

Although the bandwidth requirements are the same between the VoIP Monitor service and the Desktop Monitor service, the VoIP Monitor service can handle more simultaneous sessions (since it runs on the server). Table 6-4 and Table 6-5 expand upon the Table 6-2 and Table 6-3 by increasing the number of simultaneous sessions.
Estimating Bandwidth Consumption

For silent monitoring or recording, use QoS mechanisms to optimize WAN bandwidth utilization. Use advanced queuing and scheduling techniques in distribution and core areas as well.

<table>
<thead>
<tr>
<th>Number of Simultaneous Monitoring Sessions</th>
<th>Percentage of Available Bandwidth Required (No Silence Suppression)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>100 Mbps</td>
</tr>
<tr>
<td>1</td>
<td>0.3</td>
</tr>
<tr>
<td>5</td>
<td>1.0</td>
</tr>
<tr>
<td>10</td>
<td>1.8</td>
</tr>
<tr>
<td>15</td>
<td>2.6</td>
</tr>
<tr>
<td>20</td>
<td>3.5</td>
</tr>
<tr>
<td>25</td>
<td>4.4</td>
</tr>
<tr>
<td>30</td>
<td>5.2</td>
</tr>
<tr>
<td>35</td>
<td>6.1</td>
</tr>
<tr>
<td>40</td>
<td>7.0</td>
</tr>
<tr>
<td>45</td>
<td>7.8</td>
</tr>
<tr>
<td>50</td>
<td>8.7</td>
</tr>
</tbody>
</table>

1. The bandwidth of the connection is not large enough to support the number of simultaneous monitoring sessions.

<table>
<thead>
<tr>
<th>Number of Simultaneous Monitoring Sessions</th>
<th>Percentage of Available Bandwidth Required (No Silence Suppression)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>100 Mbps</td>
</tr>
<tr>
<td>1</td>
<td>0.1</td>
</tr>
<tr>
<td>5</td>
<td>0.3</td>
</tr>
<tr>
<td>10</td>
<td>0.7</td>
</tr>
<tr>
<td>15</td>
<td>0.9</td>
</tr>
<tr>
<td>20</td>
<td>1.2</td>
</tr>
<tr>
<td>25</td>
<td>1.6</td>
</tr>
<tr>
<td>30</td>
<td>1.9</td>
</tr>
<tr>
<td>35</td>
<td>2.2</td>
</tr>
<tr>
<td>40</td>
<td>2.5</td>
</tr>
<tr>
<td>45</td>
<td>2.8</td>
</tr>
<tr>
<td>50</td>
<td>3.1</td>
</tr>
</tbody>
</table>

1. The bandwidth of the connection is not large enough to support the number of simultaneous monitoring sessions.

For silent monitoring or recording, use QoS mechanisms to optimize WAN bandwidth utilization. Use advanced queuing and scheduling techniques in distribution and core areas as well.
Both monitoring and recording packets are real RTP streams and are not tagged with QoS marking by default. Silent monitoring sessions are time sensitive because of the real-time nature. Tag the packets in the network infrastructure to the same QoS marking as other real-time voice traffic. Properly provision the priority queue to include silent monitoring traffic. Recording packets can be tagged with QoS marking that is lower priority than the real-time voice traffic, because the packets do not need to be delivered in real time. The ports used for monitoring and recording are listed in Port Utilization for Product Revisions, page 6-12, and can be used to classify monitoring or recording traffic.

For more information about QoS traffic classification, see QoS and Call Admission Control, page 6-13. For provisioning guidelines for centralized call processing deployments, see the Cisco IP Telephony Solution Reference Network Design (SRND) Cisco CallManager Releases 4.3 documentation, available online at:


**CAD Desktop Applications Bandwidth Usage**

The CAD desktop applications include the following:

- Cisco Agent Desktop
- Cisco Supervisor Desktop
- Cisco Desktop Administrator

These applications also require a certain amount of bandwidth, although far less than the Desktop Monitor service. In addition, the type of communication across the network is bursty. In general, bandwidth usage is low when the agents are not performing any actions. When features or actions are requested, the bandwidth increases for the time it takes to perform the action, which is usually less than one second, then drop down to the steady state level. From a provisioning standpoint, one must determine the probability of all the CAD agents performing a particular action at the same time. It might be more helpful to characterize the call center and determine the maximum number of simultaneous actions (in the worst case) to determine instantaneous bandwidth requirements, then determine what amount of delay is tolerable for a percentage of the requested actions.

For example, the raw bandwidth requirement for 300 CAD agents logging in simultaneously is about 4.5 Kilobytes/second and the login time is about 9 seconds (with no network delay) for each agent. If the WAN link did not have this much bandwidth, logins would take longer as packets were queued before being sent and received. If this caused the login attempts to take twice as long (18 seconds), would this delay be acceptable? If not, more bandwidth should be provisioned.

Each of these applications communicates with the base CAD services running on server machines. In addition, the agent desktop application communicates with the CTI server for call control actions and state changes. **Table 6-6** displays the types of messaging for each application.
Table 6-7 displays the average bandwidth requirements for different numbers of agents. This information is derived from bandwidth testing and extrapolation of bandwidth data. Since there are many variables that can affect bandwidth, a configuration that resulted in higher bandwidth usage was chosen to provide near worst-case scenarios. If the agent’s WAN link meets or exceeds the bandwidth requirements shown in this table, Cisco Agent Desktop will be able to run without delays in message passing.

The configuration parameters that affect bandwidth and apply to Tables 6-7 and 6-8 are shown below.

- Number of skills per agent: 10
- Number of agents per team: 20
- Number of teams: 50
- Number of agent state changes per agent per hour: 10 (This does not count state changes due to handling calls)
- Calls per agent per hour: 60
- Team Performance Messages per team per hour: 8
- Chat messages sent/received per hour: 20
- Average chat message size (in bytes): 40
- Number of calls recorded per hour: 10

Cisco Agent Desktop Bandwidth Usage

CAD agents are able to login and logoff their agents, change their agent state, handle calls, and send reporting information to the base servers. The bandwidth requirements for these activities are fairly small but can add up when many agents are considered.
Note

The bandwidth requirements shown do not include the bandwidth of the RTP streams for the call, recording, or monitoring session.

Table 6-7  Average Bandwidth Requirements For Cisco Agent Desktop

<table>
<thead>
<tr>
<th>Number of agents</th>
<th>Average Download Bandwidth (Kilobytes/second)</th>
<th>Average Upload Bandwidth (Kilobytes/second)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.02</td>
<td>0.003</td>
</tr>
<tr>
<td>10</td>
<td>0.2</td>
<td>0.02</td>
</tr>
<tr>
<td>50</td>
<td>0.8</td>
<td>0.07</td>
</tr>
<tr>
<td>100</td>
<td>1.7</td>
<td>0.1</td>
</tr>
<tr>
<td>150</td>
<td>2.5</td>
<td>0.2</td>
</tr>
<tr>
<td>200</td>
<td>3.4</td>
<td>0.3</td>
</tr>
<tr>
<td>250</td>
<td>4.2</td>
<td>0.4</td>
</tr>
<tr>
<td>300</td>
<td>5.0</td>
<td>0.4</td>
</tr>
</tbody>
</table>

Cisco Supervisor Desktop Bandwidth Usage

A Cisco Supervisor Desktop will receive events for all the agents of the team that the supervisor is logged into. This information includes state changes, call handling, login/logoff, etc. The more agents, skills, and calls there are, the more data will be sent to supervisors. In addition, particular reports are automatically refreshed periodically to provide real-time data while the supervisor is viewing the report. Refreshing reports requires additional bandwidth.

Table 6-8 uses the same basic configuration parameters used to determine the bandwidth numbers in Table 6-7. The CSD has the optional parameter to view and refresh some reports, which requires additional bandwidth.

Table 6-8  Bandwidth Requirements For Cisco Supervisor Desktop

<table>
<thead>
<tr>
<th>Number of agents</th>
<th>Average Download Bandwidth (Kilobytes/second)</th>
<th>Average Upload Bandwidth (Kilobytes/second)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.02</td>
<td>0.01</td>
</tr>
<tr>
<td>10</td>
<td>0.1</td>
<td>0.02</td>
</tr>
<tr>
<td>50</td>
<td>0.6</td>
<td>0.07</td>
</tr>
<tr>
<td>100</td>
<td>1.3</td>
<td>0.1</td>
</tr>
<tr>
<td>150</td>
<td>1.9</td>
<td>0.2</td>
</tr>
<tr>
<td>200</td>
<td>2.5</td>
<td>0.3</td>
</tr>
<tr>
<td>250</td>
<td>3.1</td>
<td>0.3</td>
</tr>
<tr>
<td>300</td>
<td>3.7</td>
<td>0.4</td>
</tr>
</tbody>
</table>
Cisco Desktop Administrator Bandwidth Usage

The bandwidth requirements for CDA are very small and are only seen when an administrator is actively changing configurations. In general, the bandwidth used by CDA is negligible from a provisioning standpoint.

Remote Agent Traffic Profile

Unified CCX signaling represents only a very small portion of control traffic (Cisco Unified CallManager CTI and ICD subsystems) in the network. For information on TCP ports and Differentiated Services Code Point (DSCP) marking for Unified CCX ICD and CTI traffic, see the sections on Serviceability and Security, page 6-12, and QoS and Call Admission Control, page 6-13.

Bandwidth estimation becomes an issue when voice is included in the calculation. Because WAN links are usually the lowest-speed circuits in an IP Telephony network, particular attention must be given to reducing packet loss, delay, and jitter where voice traffic is sent across these links. G.729 is the preferred codec for use over the WAN because the G.729 method for sampling audio introduces the least latency (only 30 milliseconds) in addition to any other delays caused by the network.

Where voice is included in bandwidth, system architects should consider the following factors:

- Total delay budget for latency (taking into account WAN latency, serialization delays for any local area network traversed, and any forwarding latency present in the network devices). The generally agreed-upon limit for total (one-way) latency for applications in a network is 150 milliseconds.
- Impact of delays inherent in the applications themselves. 25 seconds is the initial Unified CCX agent login setup time with no WAN delay. The overall time to log in agents and base delay adds approximately 30 seconds of delay per 70 milliseconds of WAN delay.
- Impact of routing protocols. For example, Enhanced Interior Gateway Routing Protocol (EIGRP) uses quick convergence times and conservative use of bandwidth. EIGRP convergence also has a negligible impact on call processing and Unified CCX agent logins.

Use Table 8-1 to estimate the number of Unified CCX agents that can be maintained across the WAN (with IP Telephony QoS enabled). These numbers are derived from testing where an entire call session to Unified CCX agents, including G.729 RTP streams, is sent across the WAN. Approximately 30% of bandwidth is provisioned for voice. Voice drops are more of an issue when you are running RTP in conjunction with Cisco Agent Desktop and other background traffic across the WAN. These voice drops might occur with a specific number of agents at a certain link speed, and those possible scenarios are denoted by the entry N/R (not recommended) in Table 6-9.

<table>
<thead>
<tr>
<th>Frame Relay</th>
<th>128 KB</th>
<th>256 KB</th>
<th>512 KB</th>
<th>768 KB</th>
<th>T1</th>
</tr>
</thead>
<tbody>
<tr>
<td>G.729</td>
<td>3</td>
<td>7</td>
<td>15</td>
<td>25</td>
<td>38</td>
</tr>
<tr>
<td>G.711</td>
<td>N/R</td>
<td>N/R</td>
<td>N/R</td>
<td>N/R</td>
<td>14</td>
</tr>
</tbody>
</table>

In remote agent deployments, QoS mechanisms should be used to optimize WAN bandwidth utilization. Advanced queuing and scheduling techniques should be used in distribution and core areas as well. For information on QoS traffic classification, see QoS and Call Admission Control, page 6-13. For provisioning guidelines for centralized call processing deployments, refer to the Cisco IP Telephony Solution Reference Network Design (SRND) Cisco Unified CallManager Releases 4.0 and 4.1 documentation, available online at http://www.cisco.com/warp/public/779/largeent/it/ese/unifiedCommunications.html
Serviceability and Security

Security can be implemented on many levels. Applications security is clearly dependent upon security implemented at the infrastructure level. For more details on security at the network infrastructure level, refer to security design considerations in the *Cisco IP Telephony Solution Reference Network Design (SRND) for Unified CallManager 4.3* documentation, available online at


Corporate Data Access

Aside from call routing, Unified CCX or Unified IP IVR scripts often process enterprise data from existing corporate data stores such as a database or a corporate directory server for functions such as account authorization and order status. Often, these data stores already exist and share data with other enterprise applications. Figure 6-2 shows an example of a network where voice and data components reside in separate VLANs and are separated by a firewall.

![Figure 6-2 Unified CCX Accessing Data Stores](image)

Unified CCX can communicate with these external sources through its subsystems, provided Network Address Translation (NAT) is not used.

Port Utilization for Product Revisions

For a list of the TCP and UCP ports used by Cisco CRS 4.1(1), including Unified IP IVR and Unified CCX, see:

Ping, NAT, PAT, and Reverse DNS Lookups

The following configurations and information are required for the CAD software to work properly. The Cisco Agent Desktop application uses the TCP Ping command to verify that it can communicate with the active VoIP servers. This is done even if no agents are configured to use a VoIP Monitor service for the silent monitoring feature. If Ping is disabled on the machine running a CAD VoIP Monitor Server, the silent monitoring feature will not work properly.

There are certain CAD modules that rely upon reverse DNS lookups. If this feature is turned off on the machines running CAD services, there will be a loss of some functionality and errors will be generated and logged.

Network Address Translation (NAT) is only supported between the Cisco Agent Desktop and the CRS servers. Port Address Translation (PAT) is not supported.

QoS and Call Admission Control

Quality of Service (QoS) becomes an issue when more voice and application-related traffic is added to an already growing amount of data traffic on your network. Accordingly, Unified CCX and time-sensitive traffic such as voice need higher QoS guarantees than less time-sensitive traffic such as file transfers or emails (particularly if you are using a converged network).

QoS should be used to assign different qualities to data streams to preserve Unified CCX mission-critical and voice traffic. The following are some examples of available QoS mechanisms:

- Packet classification and usage policies applied at the edge of the network, such as Policy Based Routing (PBR) and Committed Access Rate (CAR).
- End-to-end queuing mechanisms, such as Low Latency Queuing (LLQ). Because voice is susceptible to increased latency and jitter on low-speed links, Link Fragmentation and Interleaving (LFI) can also be used to reduce delay and jitter by subdividing large datagrams and interleaving low-delay traffic with the resulting smaller packets.
- Scheduling mechanisms such as Traffic Shaping to optimize bandwidth utilization on output links.

Classifying Unified CCX and Application-Related Traffic

Table 6-10 and the following section list TCP ports and DSCP markings for use in prioritizing Unified CCX and Cisco Unified CallManager mission-critical CTI traffic. The performance criteria used in classifying such traffic should include:

- No packet drops on the outbound or inbound interface of the WAN edge router
- Voice (G.729) loss under 1%
- One-way voice delay under 150 msecs

A detailed description of QoS is not within the scope of this design guide. For QoS design recommendations, refer to the Enterprise Quality of Service Solution Reference Network Design guide available online at:

QoS Considerations for CAD software

The most important network traffic for quality of service consideration in the CAD software is the voice streams sent between VoIP requestors and providers. The processes that send and receive these voice streams have been set to have higher priorities than other processing threads. This helps assure that there will be no delays in processing these voice streams. However, the voice streams themselves contain no QoS markings. These markings are stripped off when the voice streams are captured by the VoIP provider's software. The networking components used to send these data streams (switches, routers, gateways) should be configured with the appropriate QoS settings to ensure the delivery of these voice streams to meet the intended QoS requirements.
### Server Capacities and Limits

This appendix provides a list of server capacities and limits as shown in Table A-1.

**Table A-1 Server Capacities and Limits**

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Cisco MCS-7845H1-CC1 and MCS-7845I1-CC1 (dual CPU using Win2003 Advanced Server OS) with call duration &gt; 3 minutes</th>
<th>Same server described to the left with call duration ≤ 3 minutes</th>
<th>Dual CPU Server using Win2003 Advanced Server (except MCS-7845H1-CC1 and MCS-7845I1-CC1) with call duration &gt; 3 minutes</th>
<th>All Other Supported Servers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of agents</td>
<td>300</td>
<td>150</td>
<td>200</td>
<td>75</td>
</tr>
<tr>
<td>Number of supervisors</td>
<td>32</td>
<td>30</td>
<td>32</td>
<td>10</td>
</tr>
<tr>
<td>(If a supervisor takes a call, the supervisor counts as an agent.)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of IVR ports</td>
<td>300</td>
<td>300</td>
<td>300</td>
<td>150</td>
</tr>
<tr>
<td>Number of automatic speech recognition (ASR) ports</td>
<td>150</td>
<td>150</td>
<td>150</td>
<td>75</td>
</tr>
<tr>
<td>Number of text-to-speech (TTS) ports</td>
<td>200</td>
<td>200</td>
<td>200</td>
<td>80</td>
</tr>
<tr>
<td>Number of Contact Service Queues (CSQs)</td>
<td>150</td>
<td>75</td>
<td>100</td>
<td>25</td>
</tr>
<tr>
<td>Number of skills</td>
<td>150</td>
<td>100</td>
<td>100</td>
<td>50</td>
</tr>
<tr>
<td>Number of skills with which an agent can associate</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>Number of CSQs with which an agent can associate</td>
<td>25</td>
<td>25</td>
<td>25</td>
<td>25</td>
</tr>
<tr>
<td>Number of skills with which a CSQ can associate</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>Number of CSQs for which a call can queue</td>
<td>25</td>
<td>25</td>
<td>25</td>
<td>25</td>
</tr>
<tr>
<td>Number of simultaneous Historical Reporting sessions (during normal contact center hours of operation)</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>
### Table A-1 Server Capacities and Limits (continued)

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Cisco MCS-7845H1-CC1 and MCS-7845I1-CC1 (dual CPU using Win2003 Advanced Server OS) with call duration &gt; 3 minutes</th>
<th>Same server described to the left with call duration ≤ 3 minutes</th>
<th>Dual CPU Server using Win2003 Advanced Server (except MCS-7845H1-CC1 and MCS-7845I1-CC1) with call duration &gt; 3 minutes</th>
<th>All Other Supported Servers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of simultaneous recording or playback sessions</td>
<td>32 (or 80 with Unified CCX Premium edition and the recording component installed on a separate expansion server)</td>
<td>32</td>
<td>32</td>
<td>16</td>
</tr>
<tr>
<td>Number of simultaneous silent monitoring sessions</td>
<td>32</td>
<td>30</td>
<td>32</td>
<td>10</td>
</tr>
</tbody>
</table>

**Note**

Systems that use SPAN ports and have more than 200 agents require an expansion server for recording and monitoring.

Use the Configuration and Ordering tool as the final authority on limits for the servers. The Cisco Unified Contact Center Express Configuration and Ordering Tool, is available online at

Voice Over IP Monitoring

Monitoring and recording of agent calls can be supported by two different methods in this release of Unified CCX:

- The traditional VoIP monitor Service: captures packets directly from an IP network switch via the switch's Switched Port Analyzer (SPAN) configuration. Design considerations for the traditional SPAN-based VoIP monitor Service are provided at the end of this appendix (see Design Considerations for SPAN-Based Services, page B-1).

- The Cisco Agent Desktop, also known as Endpoint monitoring or the Desktop Monitoring Service: The agent's IP phone repeats RTP packets to the agent's PC. When a supervisor wants to monitor/record the agent, the supervisor application sends a message to the agent desktop to forward the RTP packets to the supervisor, who can then monitor the agent/caller conversation via the sound card on his or her PC. This method requires the agent to use the Cisco Agent Desktop (not the IP Phone Agent) and a 7940, 7960, or 7970 IP Phone. Design considerations for the new Desktop (Endpoint) Monitoring Service are provided in Chapter 6, “Bandwidth, Security, and QoS Considerations.”

Design Considerations for SPAN-Based Services

Traditional SPAN-based VoIP services allows the IP traffic from one or more ports to be copied and sent to a single destination port.

Be aware if these factors when configuring traditional SPAN-based VoIP monitor services:

- The following switches do NOT support SPAN sessions: 1700, 2100, 2800, 2948G-L3, 4840G.

- Local SPANs (LSPANs) are SPANs where all the source ports and the destination port are physically located on the same switch. Remote SPANs (RSPANs) can include source ports that are physically located on another switch. The following switches do NOT support RSPAN (although they may be an intermediate switch in an RSPAN configuration): 1200, 1900, 2550, 2820, 2900, 2900XL, 2926GS, 2926F, 2926GL, 2926T, 2948G, 2950, 2980G, 3000, 3100, 3200, 3500XL, 3524-PWR XL, 3508GL XL, 5000, 5002, 5505, 5505.

- Some switches do not allow the destination port of a SPAN configuration to act as a normal network connection. The only traffic that can flow through this port is the traffic copied from the SPAN source ports; this requires the computer running the VoIP monitor service to have two network connections (NICs) to function properly. The following switches do NOT support normal network traffic on SPAN destination ports: 2950, 3000, 3100, 3200, 3550.
In some configurations, the VoIP Monitor service can receive duplicate voice packets, which causes poor speech quality. To avoid this, only Ingress packets to a port are sent to the VoIP monitor service. This is a setting for SPAN, which the following switches do NOT support: 1900, 2820, 2900, 2900XL, 3000, 3100, 3200, 3500XL.

In some switches, SPAN cannot use VLANs as sources, which is known as VSPAN. In that case, SPAN must designate individual ports to use for monitoring. The following switches do NOT support VSPAN: 1200, 1900, 2820, 2900XL, 2950, 3000, 3100, 3200, 3500XL, 3524-PWR XL.

For more information, refer to the *Voice Over IP Monitoring Best Practices Deployment Guide*.

Table B-1 gives the limits to the number of SPAN and RSPAN sessions that can exist on a switch:

<table>
<thead>
<tr>
<th>Switch Model</th>
<th>Maximum SPAN Sessions Allowed</th>
</tr>
</thead>
<tbody>
<tr>
<td>1200</td>
<td>1</td>
</tr>
<tr>
<td>1900</td>
<td>1</td>
</tr>
<tr>
<td>2820</td>
<td>1</td>
</tr>
<tr>
<td>2900</td>
<td>1</td>
</tr>
<tr>
<td>2900XL</td>
<td>1</td>
</tr>
<tr>
<td>2926GS</td>
<td>5</td>
</tr>
<tr>
<td>2926GL</td>
<td>5</td>
</tr>
<tr>
<td>2926T</td>
<td>5</td>
</tr>
<tr>
<td>2926F</td>
<td>5</td>
</tr>
<tr>
<td>2948G</td>
<td>5</td>
</tr>
<tr>
<td>2950</td>
<td>1</td>
</tr>
<tr>
<td>2980G</td>
<td>5</td>
</tr>
<tr>
<td>3000</td>
<td>1</td>
</tr>
<tr>
<td>3100</td>
<td>1</td>
</tr>
<tr>
<td>3200</td>
<td>1</td>
</tr>
<tr>
<td>3500XL</td>
<td>1</td>
</tr>
<tr>
<td>3524-PWR XL</td>
<td>1</td>
</tr>
<tr>
<td>3508GL XL</td>
<td>1</td>
</tr>
<tr>
<td>3550</td>
<td>2</td>
</tr>
<tr>
<td>4003</td>
<td>5</td>
</tr>
<tr>
<td>4006</td>
<td>5</td>
</tr>
<tr>
<td>4912G</td>
<td>5</td>
</tr>
<tr>
<td>5000</td>
<td>5</td>
</tr>
<tr>
<td>5002</td>
<td>5</td>
</tr>
<tr>
<td>5500</td>
<td>5</td>
</tr>
<tr>
<td>5550</td>
<td>5</td>
</tr>
<tr>
<td>5505</td>
<td>5</td>
</tr>
<tr>
<td>5509</td>
<td>5</td>
</tr>
<tr>
<td>6006</td>
<td>30</td>
</tr>
</tbody>
</table>
## Table B-1  SPAN AND RSPAN Switch-Based Session Limits

<table>
<thead>
<tr>
<th>Switch Model</th>
<th>Maximum SPAN Sessions Allowed</th>
</tr>
</thead>
<tbody>
<tr>
<td>6009</td>
<td>30</td>
</tr>
<tr>
<td>6506</td>
<td>30</td>
</tr>
<tr>
<td>6509</td>
<td>30</td>
</tr>
<tr>
<td>6513</td>
<td>30</td>
</tr>
</tbody>
</table>
Cisco Unified Contact Center Express (Unified CCX) uses directory service to store configuration information and access user and resource information. Lightweight Directory Access Protocol (LDAP) is a standard to access and modify the information in a directory. This appendix talks about some best practices when using LDAP directory with Unified CCX.

Currently Unified CCX supports three types of directory services, Cisco Unified CallManager Data Connection (DC) directory, Microsoft Active Directory (AD) and Netscape Directory (ND).

You can configure the LDAP server for user information under Cisco Unified CallManager LDAP Server Information page of CRS Administration. And the CRS configuration LDAP server can be configured under CRS LDAP Server Information page. Although the LDAP server configuration for these two types of information is separate, you should use the same LDAP server that is configured on Cisco Unified CallManager.

Since user authentication requires access to user database in LDAP server, if LDAP server is down or unavailable, you cannot access the CRS Administration web interface and agents cannot log in. Thus, be sure to install a redundant LDAP server to provide high availability.

If Microsoft Active Directory is used, special users for JTAPI and RmCm subsystems are created, and the password can be entered in CRS Administration. However, the password must also be entered in AD using tools provided by the supplier (for example, Microsoft ADSI Edit for Active Directory). The password entered must match the password entered in CRS Administration. In release 4.1, CRS supports child domain for agents.

Unified CCX releases before release 4.1 allowed the sharing of applications among CRS systems. In release 4.1, applications are stored in the cluster's database repository, so applications cannot be shared between clusters any more.

In scenarios where multiple CRS clusters are referencing the same LDAP instance, each cluster has visibility to all Resources. Any LDAP user that has been assigned an ICD extension will be listed as a Resource in CRS Administration for each CRS cluster. Since a Resource can only be associated with one CRS cluster, this requires that the Administrator be aware of which resources are associated with each cluster. The Administrator can mitigate the confusion by having a unique naming convention for Resources associated with a particular CRS cluster.

Since CRS synchronizes with LDAP for user information every 10 minutes, a directory with a large number of agents may take a long time to synchronize and this can impact the network and server performance.
For the provisioning instructions of LDAP Server on Unified CCX and specific information on directory configuration, including migration between DC, AD, and ND, consult Chapter 8 of the *Cisco Customer Response Solutions Administration Guide*, Release 4.1(x).


For more information on directory access and integration, please refer to Chapter 14 of *Cisco IP Telephony Solution Reference Network Design (SRND) for Cisco Unified CallManager 4.3* at:

INDEX

A

A2Q bid assurance 5-1
accessing corporate data stores 6-12
Active Directory C-1
Advanced ACD features 1-9
Advanced CTI features 1-11
Advanced IVR features 1-7
agents 6-1
   service level goal 4-2
AHT 4-2
application integration
   with CAD 1-10
Architecture 2-1
ASR
   integration 1-8
   resource allocation 2-11
   servers 2-11
audience for this guide vii
automatic failover
   with CAD 1-10
   with CSD 1-11
Automatic Speech Recognition (ASR) 2-11
average handle time (AHT) 4-2
AVVID network 2-14

B

bandwidth 6-1
   by Codec 6-5
   estimating 6-1
for desktop applications 6-8
for VoIP Monitor Service 6-6

IP Phone call usage 6-4
   per-call packet size requirements 6-4
   silent monitoring usage 6-2
Barge-in
   with CSD 1-11
Basic ACD features 1-4
Basic CTI features 1-7
Basic IVR features 1-4
BHCA 4-2
bid assurance, A2Q requirements 5-1
blocked calls 4-2
Busy Hour Call Attempts (BHCA) 4-2

C

CAC 6-13
   call admission control (CAC) 6-13
   call blockage 4-2
   call center sizing 4-1
      principal steps and design considerations 4-4
   call flow example 2-3
   call recording playback and exports
      with CSD 1-11
   call routing and queueing 1-9
   capacities of servers A-1
   Catalyst switches
      max SPAN sessions allowed B-2
Cisco QM vii
Citrix
   support for CAD 2-10
   terminal services 2-10
   classification of traffic 6-13
   cluster
difference from Unified CallManager 2-14
maximum size 2-6
codec
  G.711 6-5
  G.729 6-5
cold standby support 2-17
Committed Access Rate (CAR) 6-13
compatibility 2-18
Configuration and Ordering Tool 4-5
contact center
  representation 6-3
  corporate data stores 6-12
CRS Engine
  functions 2-6
CSD Record Viewer 2-8
CTI
  port group 2-14
CTI Manager
  Unified CallManager subsystem 2-14
two-server HA 3-6
desktops
  CAD features 1-5, 1-10
  Cisco Agent Desktop 2-5
  Cisco Supervisor Desktop 2-5
CSD features 1-7, 1-10
destination ports
  SPAN configuration B-1
directories
  DC 2-3
    Microsoft Active Directory 2-3
    Netscape iPlanet 2-3
  supported 2-3
directory services
  Netscape Directory C-1
  Unified CallManager DC Directory C-1
DNS lookups 6-13
document structure viii
DSCP markings 6-13
e-mail generation 1-8
end-to-end queuing 6-13
Erlang calculators 4-3, 4-5
examples
  Sizing Unified CCX Standard Application with 25 Agents 4-5
  Standard Resource Calculator Basic Example 4-6
  Standard Resource Calculator Call Treatment Example 4-8
fault tolerance 2-14
cold standby 2-17
CRS Engine redundancy 2-14
database redundancy 2-15
monitoring and recording redundancy 2-16
G

grade of service 4-2

H

High Availability 2-14
historical reporting 1-11
historical reports 2-5
history of this guide ix
HTTP triggers 1-8

I

Intercept
with CSD 1-11
interfaces 6-14
IPCC Express Gateway PG 2-11
IP Phone Agent (IPPA)
features 1-6, 1-10
SPAN port server 2-9
IVR
average port usage time 4-2
ports 4-2, 4-5
sizing ports 4-5
IVR port call recording 1-8

J

Java support 1-8
JTAPI
defined 2-2
trigger 2-2
user 2-3

L

LDAP directory 2-3
LDAP server
integrating Unified CCX with C-1
licensing 1-3
IVR ports by package 1-3
product components by seat license 1-4
LSPAN B-1

M

managing the Unified CCX system 2-5
Media Resource Control Protocol (MRCP) 2-11
monitoring
desktop 2-7
remotely 1-9
Monitoring and Recording redundancy 2-16
Monitoring Server 2-6

N

NAT 6-13
Network Address Translation (NAT) 6-13
network partitioning 2-16
NIC B-1
processing packets 2-7

O

On-Demand Agent Call Recording
with CSD 1-11
On-Demand Call Recording
with CAD 1-10

P

packaging
Advanced ACD 1-9
Advanced CTI 1-11
Advanced IVR 1-7
Basic ACD  1-4
Basic CTI  1-7
Basic IVR  1-4
  functionality included  1-2
  primary functions  1-2
percent blockage  4-2
performance criteria
  categories  5-2
  effects of  5-2
Ping
  used by CAD  6-13
Policy Based Routing (PBR)  6-13
Port Address Translation (PAT)  6-13
ports
  gateway  4-1
  interfaces  6-14
IVR  4-2
protocols  6-14
PSTN  4-1
queue  4-2
sizing  4-5
utilization  6-12
preface  vii
preliminary information  4-2
protocols  6-13, 6-14
provisioning (See sizing)
PSTN  4-1
purpose of this guide  vii

R
real-time reports  2-5
Recording Server  2-7
  redundant  2-8
releases of software  viii
remote agent  6-1
Remote Agent Traffic Profile  6-1
remote silent monitoring  1-9
remote supervisory monitoring  2-9
reports
  historical  2-5
  real-time  2-5
resource requirements  4-5
revisions history  ix
routing
  agent  1-9
  agent skill  1-9
  competency based  1-9
RSPAN  8-1
RTP streams  2-8
rules for design  3-3

S
scope of this guide  vii
security  6-1, 6-12
servers
  capacities  A-1
  limits  A-1
  sizing  5-1, A-1
serviceability  6-12
service level  4-2
silent monitoring  1-10
  bandwidth usage  6-2
  providers  6-2
  remote  1-9
  requestors  6-2
  sizing
QoS  6-1, 6-13
QoS mechanisms
  for WAN bandwidth use  6-7
Quality of Service (QoS)  6-1, 6-13
queue ports  4-2
queueing
  prioritized  1-9
call center 4-1
Configuration and Ordering Tool 4-5
document 4-2
example calculations 4-5
IVR ports 4-5
servers 5-1, A-1
Unified CCX server 5-1
sizing tools
  Configuration & Ordering Tool 5-1
descriptions 5-1
software components 2-5
  CRS Engine 2-6
  Database 2-6
  Monitoring 2-6
  Recording 2-6
software releases (versions) viii
source ports
  SPAN configuration B-1
SPAN B-1
  port monitor server 2-9
  port monitoring SPAN 2-16

T
TCP ports 6-13
terminology 4-1
Text-to-Speech (TTS) 2-11
traffic classification 6-13
Traffic Shaping 6-13
TTS
  integration 1-9
  resources 2-11

U
Unified CallManager
  cluster approach 2-14
Unified CCX
  Configuration and Ordering Tool 4-5
deployment models 3-1
interfaces 6-14
overview 1-1
protocols 6-14
sizing the server 5-1
solution architecture 2-1
terminology 2-2
Unified CM
  sizing tool 5-2
  system performance 5-2
Unified ICME software
  integration with Unified CCX 2-11

V
versions of software viii
Voice Gateway 2-14
Voice over IP monitoring B-1
Voice XML 1-8
VSPAN B-2

W
Wireless 7920 phone support
  roaming 2-9
Wireless Access Point (WAP) range 2-9
workflow buttons
  with CAD 1-10
work mode 1-10
wrap-up mode 1-10