Cisco IPCC Express Solution Reference Network Design

Cisco IPCC Express, Release 3.5
April 2004

Corporate 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: 408 526-4100

Customer Order Number: 9560890308

Preface

Purpose

This document provides system-level best practices and design guidance for the Cisco IP Contact Center (IPCC) Express Edition Release 3.5. With proper planning, design, and implementation, Cisco IPCC Express 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 Cisco IPCC Express.

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

- Cisco CallManager Administration
- Cisco IPCC Express and Cisco 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 Cisco IPCC Express 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 Cisco IPCC Express, IP IVR, and Agent Desktop. For more information about these Cisco products, refer to the online product documentation available at Cisco.com.
- Cisco IP IVR and Cisco QM programming guidelines. IPCC Express is a packaged solution built upon a Cisco software platform called Customer Response Solutions (CRS). The CRS platform supports other solution packages—IP IVR and IP Queue Manager (QM). IP IVR and IP QM are primarily used with IPCC Enterprise. Unlike IPCC Express, the IP IVR and IP QM solutions do not provide ACD and CTI functions. In IPCC Enterprise deployments, the ACD and CTI functions are provided by the Intelligent Contact Management (ICM) software. ICM software, combined with either IP IVR or IP QM and CallManager, make up the IPCC Enterprise Solution. Both IP IVR and

IP QM contain a module of software which allows it to interact with the ICM software. IPCC Express does not contain the ICM interaction module of software. A single physical server can run only one of the CRS packages, either IPCC Express, IP IVR, or IP QM. A CallManager cluster allows multiple servers of different types to interoperate with the cluster.

- Best practices for Contact Service Queues (CSQs) and priority queuing of IPCC Express.
- Design guidelines for Cisco IP Telephony common infrastructure and call processing. For information on Cisco IP Telephony design, refer to the *Cisco IP Telephony Solution Reference Network Design* documentation available online at

  http://www.cisco.com/go/srnd.
- IPCC Express Voice Browser (using VoiceXML), automatic speech recognition (ASR), and text-to-speech (TTS) best practices. For specific information on these topics, refer to the Nuance Communications Inc. website at

  http://www.nuance.com
- The call sizing guidelines in this document are intended only to illustrate concepts in providing high-level sizing of call center resources. This document is not intended to be an all-inclusive guide to designing and sizing contact centers. Each deployment will be different and specific to your system requirements.

Software Releases

Unless stated otherwise, the information in this document applies specifically to Cisco IPCC Express Edition Release 3.5. Software releases are subject to change without notice, and those changes may or may not be indicated in this document. Refer to the IPCC Express release notes for the latest software releases and product compatibility information.

Document Structure

This guide contains the following chapters:

- Chapter 1 describes the packaging of the IPCC Express software.
- Chapter 2 describes the deployment models that are supported for IPCC Express.
- Chapter 3 provides some details on the architecture of the software and how the different components interact with one another.
- Chapter 4 discusses high availability design considerations.
- Chapter 5 discusses call center sizing.
- Chapter 6 provides help with using the IPCC Express configuration and ordering tool to determine the number and type of servers needed for a deployment.
- Chapter 7 discusses the performance impact to CallManager software resulting from IPCC Express.
- Chapter 8 discusses bandwidth, security, and quality of service (QoS) considerations for an IPCC Express deployment.
- Appendix A provides server capacities and limits.
- Appendix B provides information about the maximum SPAN sessions allowed on specific Catalyst switches.
- The Index helps you find information in this guide.

Revision History

The following table lists the revision history for this document.

Revision Date	Comments
April 19, 2004	Initial draft.
April 23, 2004	Final draft.

Obtaining Documentation

Cisco provides several ways to obtain documentation, 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 on the World Wide Web at this URL:

http://www.cisco.com/univercd/home/home.htm

You can access the Cisco website at this URL:

http://www.cisco.com

International Cisco websites can be accessed from this URL:

http://www.cisco.com/public/countries_languages.shtml

Documentation CD-ROM

Cisco documentation and additional literature are available in a Cisco Documentation CD-ROM package, which may have shipped with your product. The Documentation CD-ROM is updated regularly and may be more current than printed documentation. The CD-ROM package is available as a single unit or through an annual or quarterly subscription.

Registered Cisco.com users can order a single Documentation CD-ROM (product number DOC-CONDOCCD=) through the Cisco Ordering tool:

http://www.cisco.com/en/US/partner/ordering/ordering_place_order_ordering_tool_launch.html

All users can order annual or quarterly subscriptions through the online Subscription Store:

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

Ordering Documentation

You can find instructions for ordering documentation at this URL:

http://www.cisco.com/univercd/cc/td/doc/es_inpck/pdi.htm

You can order Cisco documentation in these ways:

- Registered Cisco.com users (Cisco direct customers) can order Cisco product documentation from the Networking Products MarketPlace:

  http://www.cisco.com/en/US/partner/ordering/index.shtml

- Nonregistered Cisco.com users can order documentation through a local account representative by calling Cisco Systems Corporate Headquarters (California, USA) at 408 526-7208 or, elsewhere in North America, by calling 800 553-NETS (6387).

Documentation Feedback

You can submit comments electronically on Cisco.com. On the Cisco Documentation home page, click **Feedback** at the top of the page.

You can send your comments in e-mail 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.

Obtaining Technical Assistance

For all customers, partners, resellers, and distributors who hold valid Cisco service contracts, the Cisco Technical Assistance Center (TAC) provides 24-hour, award-winning technical support services, online and over the phone. Cisco.com features the Cisco TAC website as an online starting point for technical assistance.

Cisco TAC Website

The Cisco TAC website (http://www.cisco.com/tac) provides online documents and tools for troubleshooting and resolving technical issues with Cisco products and technologies. The Cisco TAC website is available 24 hours a day, 365 days a year.

Accessing all the tools on the Cisco TAC website requires a Cisco.com user ID and password. If you have a valid service contract but do not have a login ID or password, register at this URL:

http://tools.cisco.com/RPF/register/register.do

Opening a TAC Case

The online TAC Case Open Tool (http://www.cisco.com/tac/caseopen) is the fastest way to open P3 and P4 cases. (Your network is minimally impaired or you require product information). After you describe your situation, the TAC Case Open Tool automatically recommends resources for an immediate solution. If your issue is not resolved using these recommendations, your case will be assigned to a Cisco TAC engineer.

For P1 or P2 cases (your production network is down or severely degraded) or if you do not have Internet access, contact Cisco TAC by telephone. Cisco TAC engineers are assigned immediately to P1 and P2 cases to help keep your business operations running smoothly.

To open a case 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 listing of Cisco TAC contacts, go to this URL:

http://www.cisco.com/warp/public/687/Directory/DirTAC.shtml

TAC Case Priority Definitions

To ensure that all cases are reported in a standard format, Cisco has established case priority definitions.

Priority 1 (P1)—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.

Priority 2 (P2)—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.

Priority 3 (P3)—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.

Priority 4 (P4)—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.

- The Cisco Product Catalog describes the networking products offered by Cisco Systems, as well as ordering and customer support services. Access the Cisco Product Catalog at this URL:

  http://www.cisco.com/en/US/products/products_catalog_links_launch.html

- Cisco Press publishes a wide range of networking publications. Cisco suggests these titles for new and experienced users: Internetworking Terms and Acronyms Dictionary, Internetworking Technology Handbook, Internetworking Troubleshooting Guide, and the Internetworking Design Guide. For current Cisco Press titles and other information, go to Cisco Press online at this URL:

  http://www.ciscopress.com

- Packet magazine is the Cisco quarterly publication that provides the latest networking trends, technology breakthroughs, and Cisco products and solutions to help industry professionals get the most from their networking investment. Included are networking deployment and troubleshooting tips, configuration examples, customer case studies, tutorials and training, certification information, and links to numerous in-depth online resources. You can access Packet magazine at this URL:

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

- iQ Magazine is the Cisco bimonthly publication that delivers the latest information about Internet business strategies for executives. You can access iQ Magazine at this URL:

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

- 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/en/US/about/ac123/ac147/about_cisco_the_internet_protocol_journal.html

- Training—Cisco offers world-class networking training. Current offerings in network training are listed at this URL:

  http://www.cisco.com/en/US/learning/index.html

IPCC Express Architecture and Capabilities

This chapter describes the basic architecture and capabilities of IPCC Express and explains how to match those capabilities to your system requirements. This chapter contains the following sections:

IPCC Express Overview

Cisco IPCC Express 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 200 agents. The CTI function allows call data to be “popped” onto the agent desktop.

The IPCC Express software runs on approved Cisco MCS, HP, or IBM servers and requires interaction with Cisco CallManager. The IPCC Express software can run on the same server with Cisco CallManager (co-resident) or on a separate server. For larger deployments requiring large amounts of historical reporting, silent monitoring, recording, ASR, or TTS, multiple servers might be required. A major purpose of this design guide is to help system designers to determine the number and type of servers required for an IPCC Express deployment.

CallManager provides the functionality typically associated with a PBX—call setup, teardown, and transition (transfer or conference). For calls requiring intelligent routing and queueing, CallManager interacts with IPCC Express. Within CallManager, a logical device called a CTI port is defined. Each CTI port on CallManager correlates to a logical IVR port on the IPCC Express server. When a new call arrives at CallManager, if the dialed number is associated with the IPCC Express server, CallManager will ask the IPCC Express server which CTI port to route the call to. After the IPCC Express server selects an available IVR port, CallManager sets up a VoIP data stream between the logical IVR port and the IP endpoint that made the call (either a Voice Gateway port or an IP Phone). At that point, the IPCC Express server begins a workflow that defines the call treatment to give the caller. Typically, the workflow will begin with something like “Thank you for calling...” The announcements to be played to a caller are stored on the disk of the IPCC Express server. Users interact with the IVR port by way of DTMF or speech input.

At some point in the workflow, it is possible to initiate a transfer of the call to an agent. Using agent skill information, the IPCC Express server will select an available agent and instruct CallManager to transfer the caller to the agent’s phone. If there are no agents available, the IPCC Express server plays queue announcements to the caller until an agent becomes available. When an appropriately skilled agent

becomes available, the IPCC Express server then instructs CallManager to transfer the call to the selected agent's phone. While the call is being transferred, the IPCC Express server sends call data to the agent desktop in the form of a screen pop.

IPCC Express Packaging

Cisco IPCC Express provides three primary functions—IVR, ACD, and CTI. Within the IPCC Express 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 IPCC Express licensed packages—Standard, Enhanced, and Premium.

The following table shows at a high level what functionality is included within each IPCC Express package. Details about each function are included in the sections that follow.

Functionality	Standard Package	Enhanced Package	Premium Package
Basic IVR (prompt & collect)	Yes	Yes	Yes
Advanced IVR	No	No	Yes
Basic ACD	Yes	Yes	Yes
Advanced ACD	No	Yes	Yes
Basic CTI	Yes	Yes	Yes
Advanced CTI	No	Yes	Yes

Basic IVR Functionality

All IPCC Express 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...). There is a maximum system limit, a maximum limit for a given hardware server, and a practical limit based on the number and kinds of features deployed on a given server for a given deployment. The basic IVR features are not licensed separately. The cost for the basic IVR functionality is included in the server and seat licensing costs.

All IPCC Express Edition packages support the ability to read from web-based documents, including HTTP and XML. Data obtained from these documents may be used in support of routing our screen pop.

Basic ACD Functionality

All IPCC Express packages include basic ACD functionality, such as agent and supervisor desktops and call routing and queueing.

Agent Desktop

The IPCC Express Basic ICD 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 IPCC Express software will ensure that the current call state for the IP Phone and agent desktop application are kept in synch.
- **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 by way of text chat.
- **Reason Codes**. Agents can be configured to enter reason codes for not ready and logout.
- **Media Termination Option**. Agent desktops can be installed with media termination software. This removes the requirement for using an IP Phone. Instead the sound capabilities of the agent's workstation are used. Note that the media termination option is for a single line or extension only. Therefore, if an agent needs a separate DID number for personal calls and voice mail, then an IP Phone is required. Prior to IPCC Express 3.5, the media termination option was a separately licensed option. This is no longer the case in the 3.5 release. IPCC Express software is now licensed based upon the maximum number of simultaneously logged in agents.
- **Hot Desking**. Agents can hot desk using the CallManager Extension Mobility feature. Hot desking allows agents to log in from any IP Phone registered with the same CallManager cluster. Hot desking for agents using media termination is NOT supported. Hot desking for agents using the IP Phone Agent option is supported.
- **Basic CTI**. Agent desktops provide an enterprise data window that is "popped" upon call ringing. See the section Basic CTI.

Supervisor Desktop

The IPCC Express Basic ACD functionality provides a separate supervisor desktop application from the agent desktop application. If a supervisor wishes to handle calls, then they use the agent desktop application in addition to the supervisor desktop application. Supervisors are not licensed separately. They 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.
- **Integrated Text Messaging**. Supervisors can sent text messages to one or more agents.
- **Marquee Messages**. Supervisors can send a scrolling marquee (broadcast) message to all agents.
- **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 Guide* for more details on statistics available through the supervisor desktop application.
- **Historical Reporting**. Supervisors can view historical reporting statistics for the entire contact center. See the *Cisco CRA Historical Reports User Guide* for more details on reporting details available through the Historical Reporting Application.

Call Routing and Queuing

The IPCC Express Basic ACD functionality provides the following call routing and queuing capabilities:

- **Conditional Routing**. IPCC Express supports routing based upon caller input to menus, real-time queue statistics, time of day, day of week, ANI, and dialed number.
- **Agent Selection**. IPCC Express supports longest available, linear, and circular agent selection algorithms. With Basic ACD functionality, agents are associated with one skill only.
- **Customizable Queuing Announcements**. IPCC Express 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.

Basic CTI Functionality

All IPCC Express 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 caller input (account number...), 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 IPCC Express Premium Package includes both basic and advanced IVR functionality. Cisco provides no charge licenses for two advanced IVR ports for every licensed IPCC Express seat.

The IPCC Express 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 IPCC Express Premium packaging.

In addition to the functionality discussed above in the section Basic IVR Functionality, page 1-2, the advanced IVR functionality includes the following:

- **Database Integration**. The IPCC Express server can interoperate with any ODBC-compliant database. Data retrieved from databases can then be used with the conditional routing capabilities to provide customer profile-based routing and queuing. Database integration also provides the ability to offer complete self-service applications to callers.
- **HTTP Triggers.** The IPCC Express 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, etc.) can be passed to the IPCC Express workflow to allow customer profile-based routing and a data-rich screen pop.
- **E-mail Generation**. The IPCC Express server can generate and send e-mails for things such as order confirmation.
- **Voice XML Support**. IPCC Express supports use of Voice XML in support of ASR and TTS.
- **Java Support**. The IPCC Express server can support logic defined using Java. Java support allows for logic from existing web and Java applications to be reused.
- **Voice Recording**. The IPCC Express 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.

- **Automatic Speech Recognition (ASR)**. ASR is an optional licensed software component. ASR allows callers to interact with IPCC Express IVR by using speech commands in addition to DTMF. This feature is useful for callers whose hands might be busy, such as callers using a mobile phone while they are driving. ASR allows a higher percentage of calls to be serviced in an automated fashion, thus saving expensive agent time. The ASR feature uses Nuance software and provides "speaker independent" speech recognition, meaning there is no training required for callers to use this feature. Numerous languages are supported.
- **Text-to-Speech (TTS)**. TTS is an optional licensed software component. TTS allows strings of text to be dynamically converted into a voice announcement. TTS is used where it is difficult to record all possible phrases or text strings. A good example is for use with the playback of street addresses or city names. The TTS feature uses Nuance software. Numerous languages are supported.

Advanced ACD Functionality

The IPCC Express 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 includes agent and supervisor desktops, and call routing and queuing.

Agent Desktop

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

- **Application Integration.** The agent desktop can be configured to allow call data to be passed to other desktop 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.
- **Workflow Buttons.** The agent desktop 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.
- **Dynamic Call Recording**. The agent desktop can be configured to allow clicking a single button to start and stop call recording dynamically.

Supervisor Desktop

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

- **Silent Monitoring**. The supervisor desktop allows a supervisor to silently monitor agent calls. Agents are not aware that they are being monitored.
- **Barge-in**. The supervisor desktop allows a supervisor to barge in on an agent call. The barge-in feature essentially enters the supervisor, the agent, and the caller into a three-way conference. The agent is aware when the supervisor barges in.
- **Intercept**. The supervisor desktop allows a supervisor to intercept an agent call. This essentially transfers the call to the supervisor. As the call releases from the agent desktop and phone, the agent is obviously aware when this occurs. The agent is then available to take another call.
- **Dynamic Agent Call Recording**. The supervisor desktop allows a supervisor to dynamically start and stop recording agent calls. Agents are not aware that they are being recorded.
- **Call Recording Playback**. The supervisor desktop allows a supervisor to play back calls which were recorded.

Call Routing and Queuing

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

- **Agent Skill Competency-Based Routing**. Agents can be configured with multiple skills (up to 50), each with a different competency level (up to 10). Customer contacts can be configured as requiring multiple skills (up to 50), each with a different minimum skill competency required (up to 10). The IPCC Express routing logic will then match the caller requirements with agent skills to find the optimal match. This functionality provides a better match between customer needs and agent skills.
- **Prioritized Queuing**. Customer contacts can be prioritized (up to 10 levels) based upon call or customer data.
- **Voice Mail and Callback Routing**. When queue times are long, callers can be given the option to leave a voice mail and request a callback. After the caller has recorded their message and left a callback number, the customer call is released and the voice mail and callback request are put into queue. When the agent is selected, the agent's phone will ring and a screen pop with the call data from the original call will appear. After the agent answers the phone, the agent is prompted to press 1 to hear the voice mail. After listening to the voice mail, the agent is prompted to press 1 to have a call automatically placed to the callback number. Shortly after pressing 1, the agent should hear the customer phone ringing.

Advanced CTI Functionality

The IPCC Express 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 1-4, the advanced CTI functionality allows call data to be passed to other Windows-based desktop 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.

IPCC Express in Cisco CallManager Deployment Models

This Chapter discusses the design implications of where IPCC Express is located in your network with respect to call processing resources. On a systems level, how IPCC Express is deployed can affect its performance, possibly even compromising functionality. This section addresses the importance of the IPCC Express location within a Cisco CallManager deployment.

Cisco CallManager Deployment Models

The IPCC Express design considerations in this section focus on the following main IP Telephony deployment models:

- Single-Site Deployment, page 2-5
- Multi-Site WAN Deployment with Centralized Call Processing, page 2-5
  - IPCC Express Located at the Central Site, page 2-6
  - IPCC Express Located at the Remote Site, page 2-6
- Multi-Site WAN Deployment Distributed Call Processing, page 2-8

This section assumes that you are already familiar with the Cisco Architecture for Voice, Video, and Integrated Data (AVVID) network infrastructure and the Cisco CallManager cluster design considerations for each deployment model. For more information on the infrastructure and design models, refer to the *Cisco IP Telephony Solution Reference Network Design* documentation, available online at

http://www.cisco.com/warp/public/779/largeent/it/ese/srnd.html

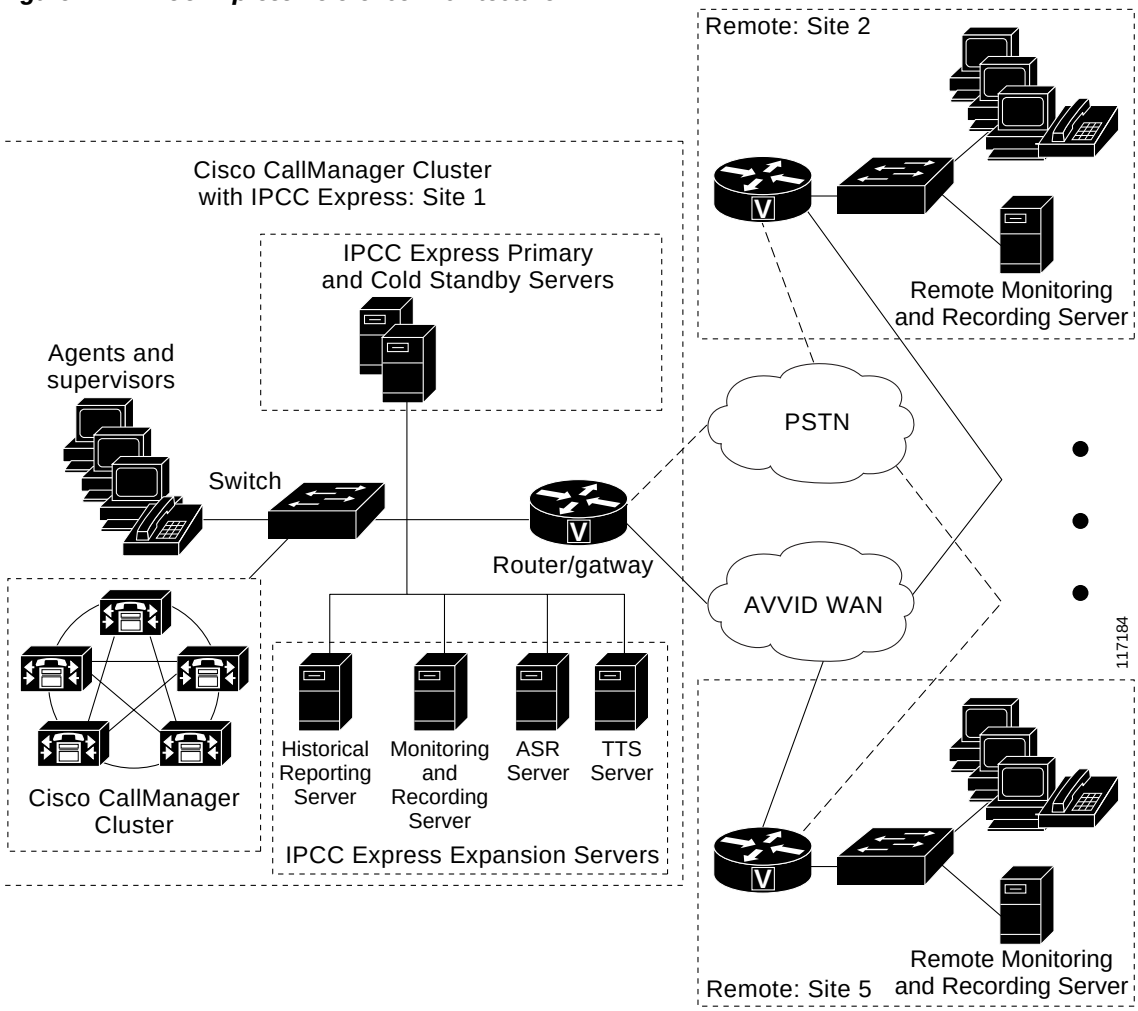
Reference Architecture

The IPCC Express reference architecture shown in Figure 2-1 represents a single IPCC Express contact center as it can be maximally deployed in Release 3.5[1]. The deployment consists of one central site and four remote sites. At the central site, an IPCC Express primary server, its (optional) cold standby server, and four IPCC Express Expansion Servers are deployed. The deployment as shown consists of the maximum number of servers supported for IPCC Express at its primary server site. Each remote site, in

[1.] Release 3.1 supports the same reference architecture; however, Release 3.0 does not. Specifically, Release 3.0 does not support either a historical reporting server nor any local or remote monitoring and recording servers

addition to the agents and supervisors deployed, has deployed a remote monitoring and recording server. The maximum number of servers supported in the multiple server model for IPCC Express Edition 3.5 is ten—one IPCC Express primary server, one cold standby server, and seven expansion servers, four of which must be remote monitoring and recording servers—one at each of the remote sites.

Figure 2-1 IPCC Express Reference Architecture



Single IPCC Express Contact Center

A single IPCC Express contact center deployment is a deployment with at most one IPCC Express primary server. A single IPCC Express deployment provides for one contact center with a single point of administration and is served by one ACD and one IVR. Deployments with multiple IPCC Express primary servers must be treated as multiple single IPCC Express deployments. Multiple IPCC Express primary server deployments can never share administration, agents, supervisors, or any features and functions and for all intents and purposes must be treated as separate contact centers.

The reference deployment architecture for a single IPCC Express deployment includes models for both single and multiple server deployments. Multiple server models include dedicated Express expansion servers, each dedicated to a single system function; in certain cases, a single expansion server may share select system functions.

Single Server Model

In a single server deployment model, all IPCC Express features and functions run on a single server including:

- The CRS Engine. The CRS engine provides all services for ACD and IVR functions.
- Desktop Services. All services supporting desktop clients including Cisco Agent and Supervisor Desktops and the IP Phone Agent XML server that supports clients running on Cisco 7940 or 7960 phones.
- CTI Services.

The IPCC Express single server model always supports deployment of all ACD, IVR, and CTI features as well as optional features such as ASR and TTS. The scale of these features in a single server model will be constrained (in some cases considerably) by the performance and capacity of the server chosen. The actual deployment capacities for any given configuration for any given server can only be determined by using the IPCC Express Configuration and Ordering Tool.

Note Note that this single server model is a conceptual model, not the actual implementation model.

Special Case for Single Server Model

IPCC Express may be deployed co-resident on the server on which the Cisco CallManager resides. When the IPCC Express primary server is deployed co-resident with Cisco CallManager no expansions servers may deployed. In addition, there are further constraints on the scalability of certain features. Finally, only certain Cisco MCS and Cisco partner (HP & IBM) servers are supported for an IPCC Express/Cisco CallManager co-resident deployment. In order to determine whether or not a given co-resident configuration will be supported you must use the IPCC Express Configuration & Ordering Tool to make that determination. Additional information on co-resident scenarios is provided in Chapter 6 of this Design Guide.

Multiple Server Model

In the multiple server model the IPCC Express primary server is augmented by one or more expansion servers. Expansion servers can be deployed only in support of the following system functions:

- Historical Reporting Contact Call Detail Records (CCDR) Database. This database is used for all historical records kept by the system and is the data repository for the Express Historical Reporting clients. From a configuration cost point of view, historical reporting is the single most expensive operation in the system and for deployments requiring multiple simultaneous historical reporting sessions this function should be moved to a dedicated expansion server. The database used for these records is, by default, the Microsoft MSDE database. The MS MSDE database is constrained in terms of the maximum size (maximum number of bytes) supported. Large deployments, deployments supporting multiple shifts or 24/7 operations, or deployments requiring a substantial period of time in which the system must provide historical data may require deployment of the optional Microsoft SQL Server database.
- Silent Monitoring and/or Recording Server. Silent monitoring and recording may require one or more dedicated expansion servers either because of the load that the number of simultaneous recording sessions puts on the system and/or to support silent monitoring and/or recording at remote WAN sites. Each remote LAN segment requires a dedicated monitoring and/or recording server. Silent Monitoring and on demand recording are features available only in IPCC Express Enhanced and Premium and are not supported in IPCC Express Standard deployments.

- Automatic Speech Recognition. A dedicated ASR server is often needed for any non-trivial number of ASR ports. ASR is only available in IPCC Express Premium and is not supported in IPCC Express Standard and Enhanced.
- Text-to-Speech. A dedicated TTS server is often needed for any non-trivial number of TTS ports.

The following functions can be deployed co-resident on the same expansion server:

- A single instance of the historical reporting database and a single monitoring and/or recording service can be deployed co-resident on the same expansion server on the same VLAN segment on which the IPCC Express primary server resides.
- A single instance of ASR and/or TTS can be deployed co-resident on the same server on the same LAN segment on which the IPCC Express primary server resides.

Note Note: Every deployment model described below in this chapter that shows support for a single IPCC Express primary server supports either the single or multiple server deployment models described above.

Required Switched Port Analyzer (SPAN) Port Configuration

Voice monitoring and recording capabilities are not built into IPCC. The Voice over IP (VoIP) monitoring and recording server accomplishes these functions by “sniffing” voice packets sent to and from IP phones. Because network switches do not normally deliver packets to Ethernet ports other than the destination port (in this case, an IP phone), the switch must be configured to perform this function.

To accomplish this, you must configure the Ethernet port for the VoIP Monitor server to monitor the Ethernet ports for all agent IP phones. If the voice packets going to and from an agent’s IP phone are not sent to the VoIP Monitor server’s port for any reason, that conversation will not be available to the supervisor.

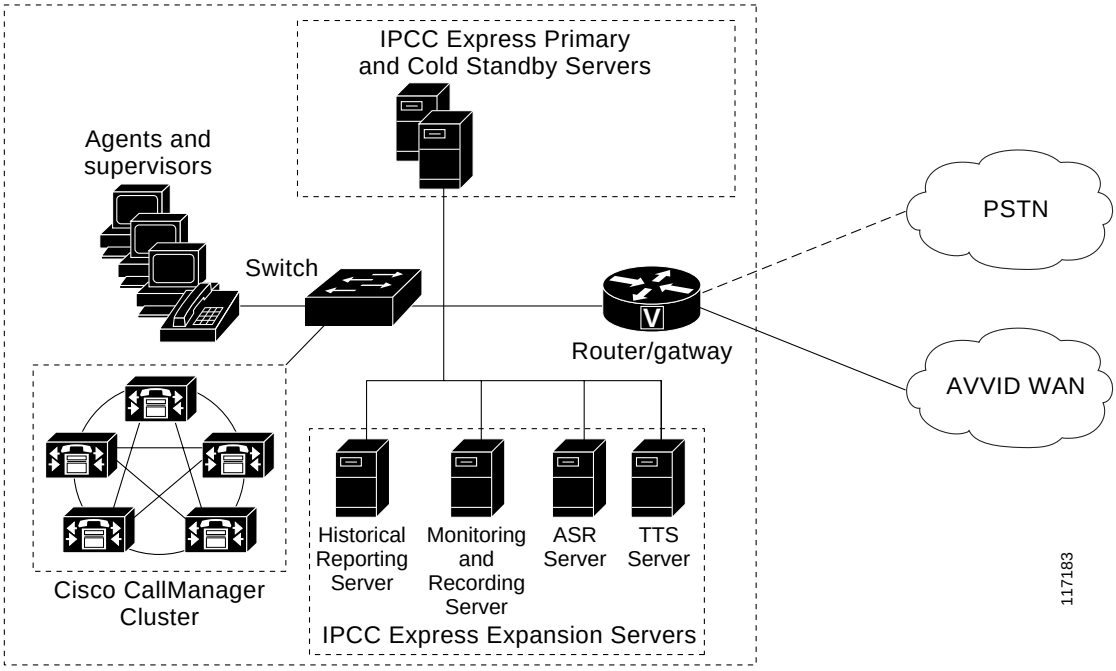
Having the VoIP Monitor server monitor a port that all voice traffic goes through (for instance, the Ethernet port to which a gateway to the PSTN is connected) is not sufficient. It must monitor the Ethernet ports to which the IP phones are directly connected. The reason for this is that the server identifies packets by the IP phone’s media access control (MAC) address. The packet’s MAC address changes as the packet moves around the network. There must not be a router between the IP phone and the port the server is monitoring. The port-monitoring feature on Cisco Catalyst switches is called Switched Port Analyzer (SPAN). For detailed information on SPAN, see Configuring the Catalyst Switched Port Analyzer (SPAN) Feature, available online at http://www.cisco.com/warp/public/473/41.html. You can also consult Appendix B of this Design Guide.

Each SPAN port must be connected either to the primary IPCC Express server or an IPCC Express Remote Monitoring and Recording server. Up to five (5) Remote Recording and Monitoring servers (one at the central site and one each at each of the remote sites) are supported per IPCC Express primary server.

Single-Site Deployment

In a single-site deployment, the Cisco CallManager cluster, its supporting IP Telephony application servers, IPCC Express agents, and their IP phones, are all located on a single central campus. Figure 2-2 shows an example IPCC Express configuration of a single-site campus.

Figure 2-2 IPCC Express in a Single-Site Deployment



In Figure 2-2, the IPCC Express primary server telephony subsystem connects to one of the Cisco CallManager servers in the cluster. This Cisco CallManager server also runs the CTI Manager service that handles the CTI call processing requests from IPCC Express.

Multi-Site WAN Deployment with Centralized Call Processing

This deployment model supports two configuration options:

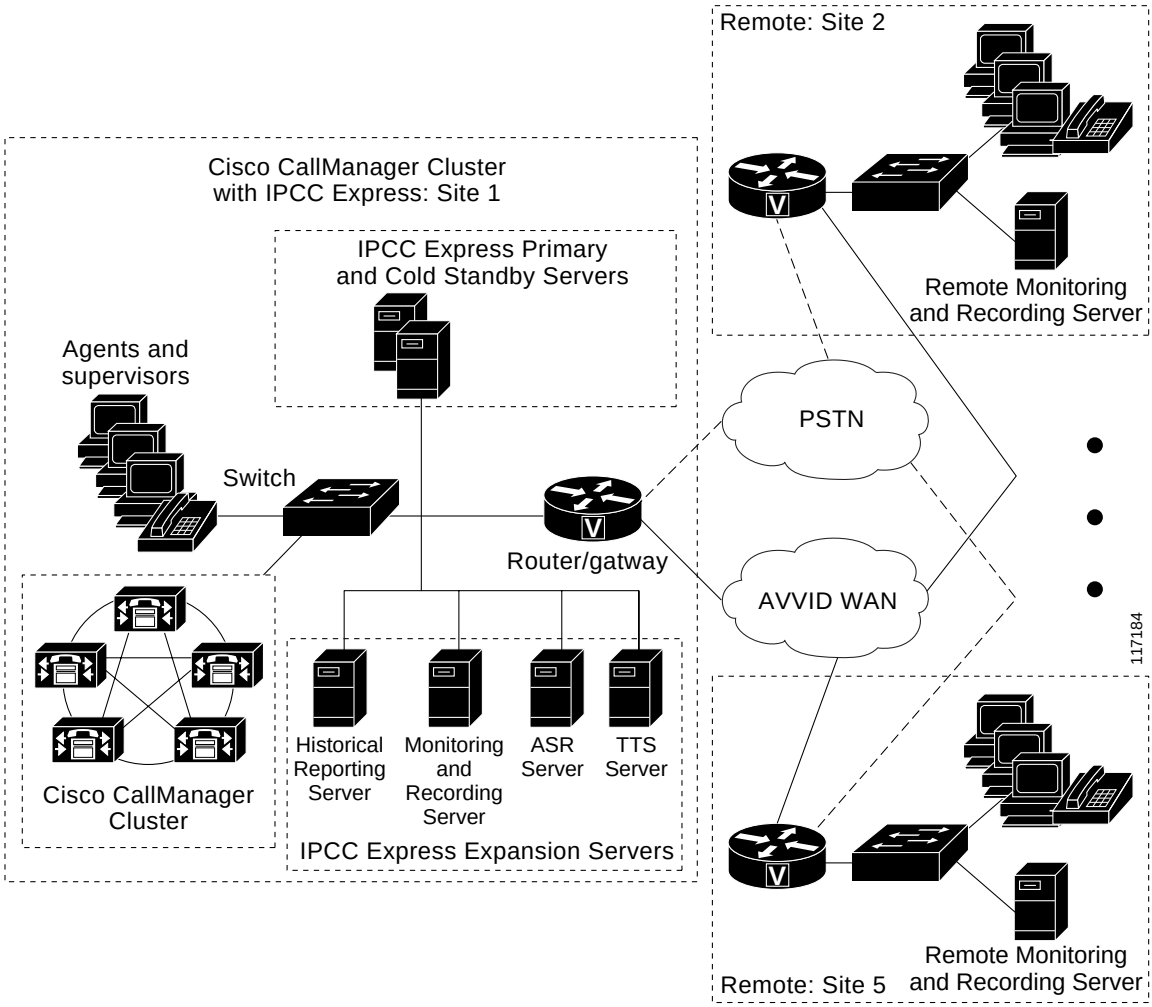
- IPCC Express Located at the Central Site, page 2-6
- IPCC Express Located at the Remote Site, page 2-6

Bandwidth considerations for both of these deployment models can vary based on the location of the IPCC Express primary server. See Chapter 8 for more details.

IPCC Express Located at the Central Site

In this deployment model, all of the call processing and IPCC Express application servers are located at the central site. Phones and IPCC Express agents are distributed at remote branches. Phones and other call processing endpoints interface to Cisco CallManager over an IP WAN link (for example, Frame-Relay). CTI, Skinny Client Control Protocol (SCCP), and RTP traffic pass over the IP WAN link between the central and remote sites. Figure 2-3 shows an example of this deployment model.

Figure 2-3 Centralized Call Processing with IPCC Express at the Central Site



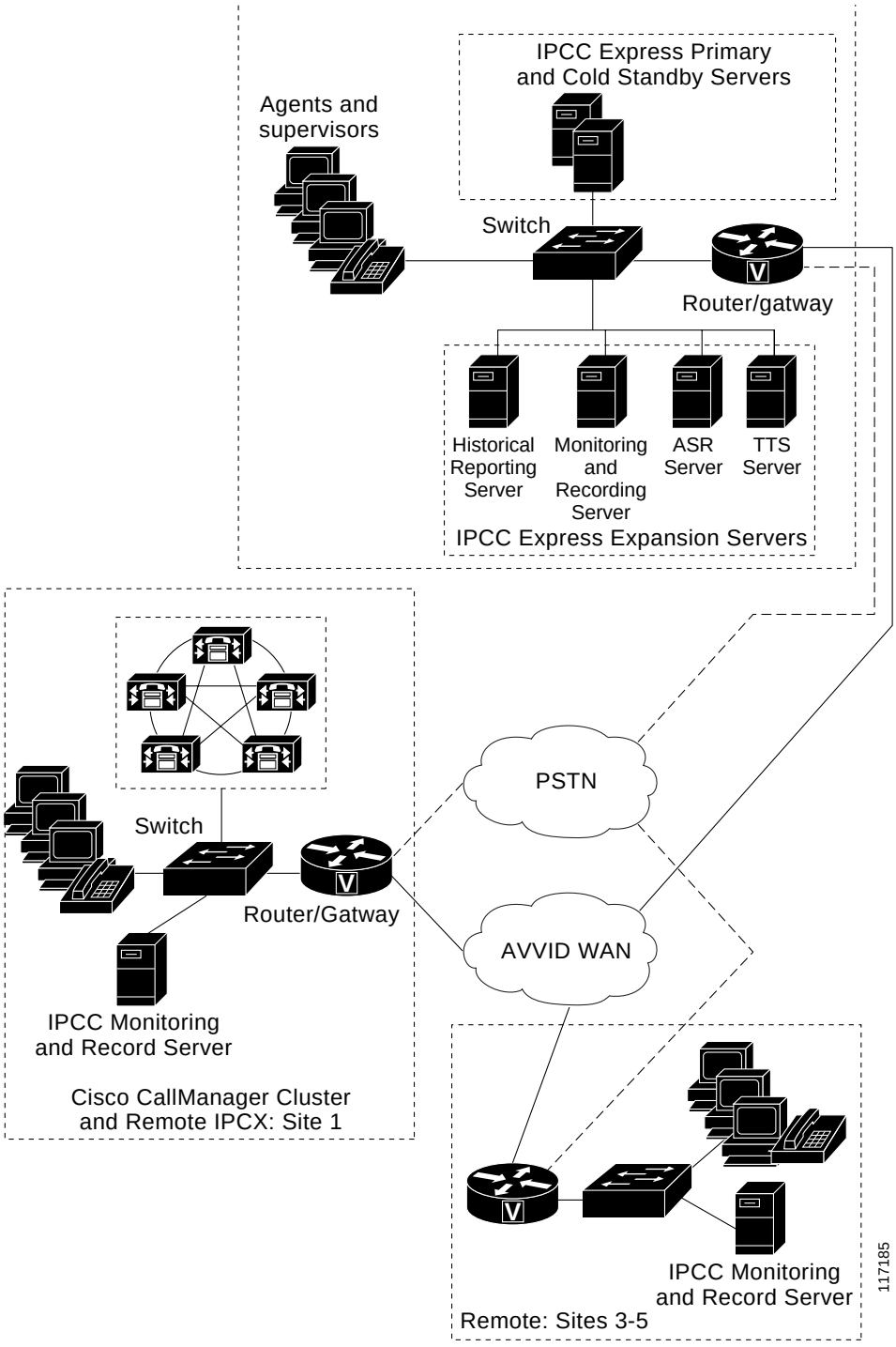
The disadvantage of this model is that any IP WAN link failure prevents remote agents from accepting incoming calls; however, IPCC Express agents at the central site can continue to accept incoming calls.

This is the recommended deployment model for multi-site WAN deployment with centralized call processing. In this model, the entire virtual contact center can fail with the failure of the IPCC Express primary server. Individual remote sites might fail with the failure of the WAN components required for establishment of the WAN link to the remote site.

IPCC Express Located at the Remote Site

As an alternative configuration, you can install IPCC Express at the remote site while leaving the Cisco CallManager cluster at the central site, as shown in Figure 2-4.

Figure 2-4 Centralized Call Processing with IPCC Express at the Remote Site



In general this configuration is not recommended except in the case where the gateway handling the calls is at the remote site. In this case, the primary benefit of this configuration is that it saves backhauling of the call from the central site.

A significant drawback of this deployment model is the potential loss of WAN connectivity from the CallManager site to the IPCC Express primary server site, which exposes IPCC Express to the following problems:

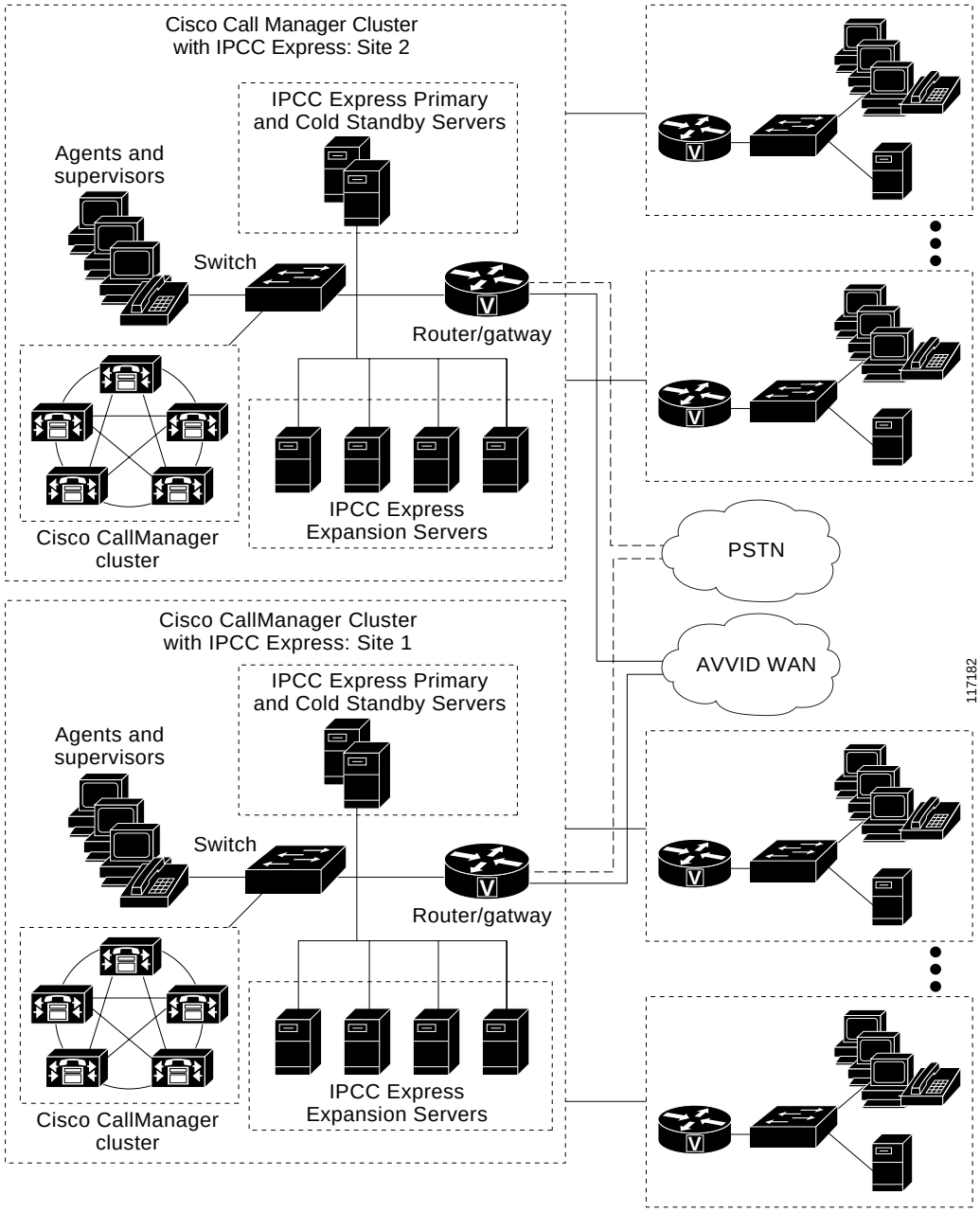
- Unavailability of a directory server for Lightweight Directory Access Protocol (LDAP) authentication. (For example, agents, supervisors, and administrators are logged off, and the administrator is unable to log in to IPCC Express to make configuration changes. The JTAPI subsystem therefore becomes unavailable.)
- No IPCC Express agents can log in.
- Loss of all call processing services from CallManager to the IPCC Express primary server – therefore no ACD routing or IVR call treatment services are possible.

Since it may well be the case that an actual deployment may support multiple remote sites, a failure of the WAN link between the central CallManager site and the branch site at which the IPCC Express server resides will result in a failure of the entire virtual contact center. Contrast that with the deployment model where both the Cisco CallManager and IPCC Express are both at the central site. In this case, a single WAN link failure impacts only a single remote branch.

Multi-Site WAN Deployment Distributed Call Processing

In a distributed call processing deployment, each site has its own Cisco CallManager server clusters. Figure 2-5 shows an example of this type of configuration.

Figure 2-5 IPCC Express in a Distributed Call Processing Deployment



In this model, the contact center at each remote site is treated as a separate single-site deployment. Therefore, you can apply single-site design considerations to each site in this model.

The major consideration with call treatment is bandwidth provisioning for call admission control at the gatekeeper. For more details call admission control in a distributed call processing deployment, refer to the *Cisco IP Telephony Solution Reference Network Design* documentation, available online at http://www.cisco.com/warp/public/779/largeent/it/ese/srnd.html.

Special Considerations in Deployment Model Design

Please refer to the section Reference Architecture, page 2-1 when reading the following discussion.

Expansion Servers

In IPCC Express, all ACD, CTI and IVR processing takes place on the IPCC Express primary server. In particular, this means that the performance and capacity for agents, supervisors and IVR ports are constrained by the physical server performance and capacities. The addition of certain features, particularly the number of simultaneous historical reporting sessions, the number of simultaneous recording and/or monitoring sessions, and the number of ASR and/or TTS ports will significantly reduce the scalability of ACD and IVR capabilities.

In almost all the cases, the decision to deploy expansion servers is driven by the performance and capacity loads put on the physical server by these features. Unless these features are moved to an expansion server, the number of agent and/or supervisor positions and the number of prompt & collect or IVR ports will be constrained, oftentimes below the number required to meet the customer's requirements.

Using the IPCC Express Configuration & Ordering Tool makes it easy to see the effect on ACD and IVR capacities of deploying one or more of these features on one or more expansion servers.

ACD/CTI and IVR on Separate Servers

It is possible to deploy the ACD/CTI capabilities on one IPCC Express primary server and to deploy an IP IVR server to handle the IVR ports. However, doing so prevents the passing of call-associated data to the ACD for use in routing decisions or in support of agent screen pop. For example, call ANI or DNIS might be required to make routing decisions. Caller-entered data may be required to be popped to an agent or to be used as a key for a database dip in support of screen pop or third party application integration.

Deploying ACD/CTI on one server and IP IVR on another server is a viable deployment model only when there will never be a need to share information associated with a call arriving at the IVR, or caller data collected by the IVR, with the ACD or CTI services on the IPCC Express server.

Meeting Capacities in Excess of a Single IPCC Express System

When customer requirements for system capacities for agent positions and IVR ports exceed the capabilities of a single IPCC Express system the only alternatives are to deploy larger servers. In the case where even the largest server fails to meet capacity requirements the only option is to deploy multiple single IPCC Express systems.

As previously discussed, this deployment model does not result in a single larger IPCC Express system but rather in multiple separate systems. Customers typically perceive the following issues as problems:

- Each system must be separately administered.
- Skill groups must be separate between system ACDs.
- Caller data cannot be shared between system IVRs.
- Historical and real-time reports reflect only the separate single systems.

IPCC Express System Design Considerations

This chapter addresses system design consideration for integrating IPCC Express with a Cisco IP Telephony network, and it contains the following major sections:

- Mapping IPCC Express to Cisco CallManager Devices, page 3-1

  Describes how physical and logical devices on IPCC Express map to Cisco CallManager devices (together with a description of a typical IPCC Express call flow), and it explains how to provision these Computer Telephony Integration (CTI) devices.

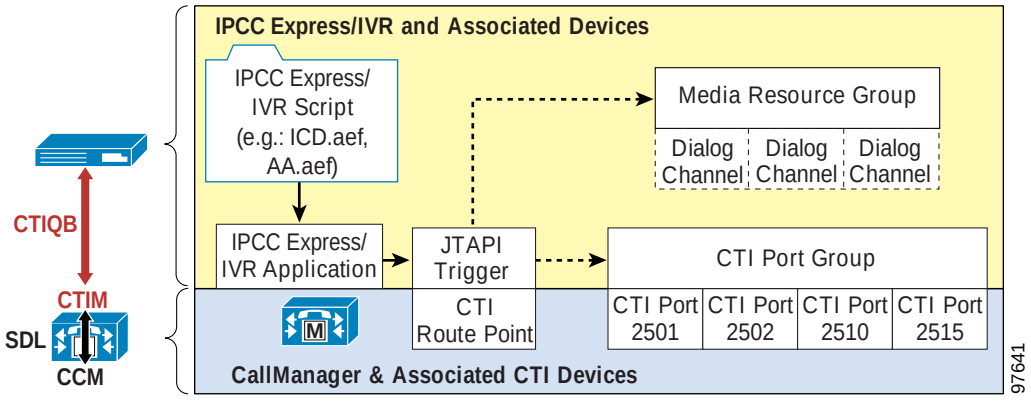
- Provisioning Cisco CallManager Resources, page 3-3

  Shows how to determine the number of Cisco CallManager servers needed to support agent phones, gateways, and IVR applications.

Mapping IPCC Express to Cisco CallManager Devices

To scale end-to-end IP Telephony call processing resources accurately, it helps to understand how the IPCC Express telephony (JTAPI) and subsystems (RMCM) interface with Cisco CallManager. Figure 3-1 illustrates how Computer Telephony Integration (CTI) serves as an interface for the IPCC Express, CTI Manager, and Cisco CallManager services, and how the physical and logical devices of IPCC Express map to Cisco CallManager and associated CTI devices.

Figure 3-1 IPCC Express Resources in Relation to Cisco CallManager Devices



As shown in Figure 3-1, IPCC Express interfaces with the Cisco CallManager server primarily through its JTAPI subsystem. CTI Manager (CTIM) is a service running on a Cisco CallManager server. CTIM acts as an application broker for IPCC Express, and it abstracts the physical binding of the application to a particular Cisco CallManager server to handle call control for its CTI resources. The CTI Manager handles CTI requests from IPCC Express by way of the CTI Quick Buffer Encoding (CTIQBE) protocol. CTIM then passes those requests to the Cisco CallManager service for call processing. In turn, the CTI Manager is aware of secondary and tertiary Cisco CallManager (CCM.EXE) services that may be prioritized to handle these call processing requests.

From the perspective of IPCC Express, an IVR application script is associated with a JTAPI trigger in order to handle an incoming call routed from the Cisco CallManager server. To Cisco CallManager, this JTAPI trigger (configured in IPCC Express) maps to an associated Route Point configured in Cisco CallManager. IPCC Express then determines if there are available CTI ports within one of its CTI port groups that can handle the call session. If prompting is required, IPCC Express uses a configured resource from its media port group.

Typical IPCC Express Call Flow

With a basic understanding of the CTI Architecture, the first step is to determine what types of CTI resources are required by IPCC Express for configuration on the Cisco CallManager server. As an example, we will use a typical IPCC Express call flow, as illustrated in Figure 3-2.

Figure 3-2 Typical IPCC Express Call Flow

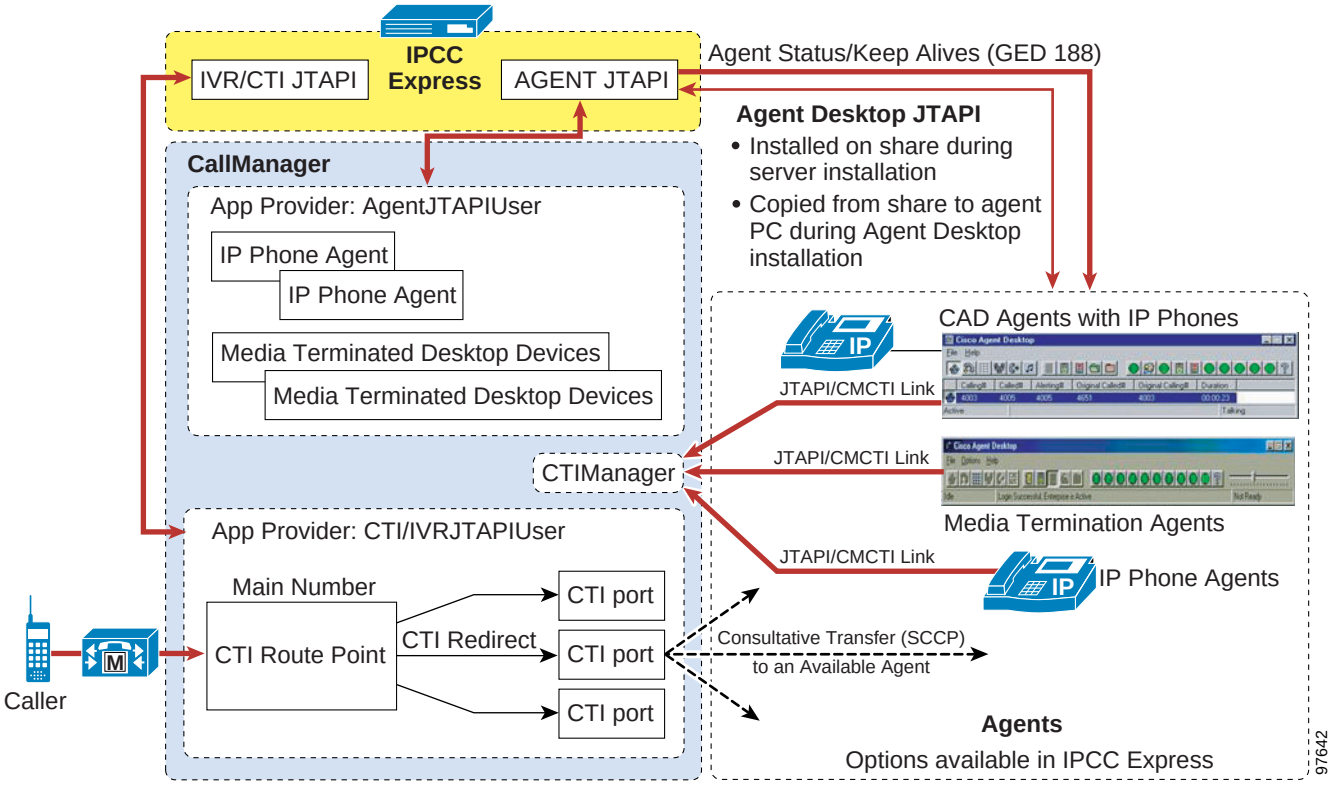


Figure 3-2 shows an incoming call routed through the Cisco CallManager server to a CTI route point assigned to the IPCC Express JTAPI trigger. Route points and CTI ports are configured in IPCC Express such that all calls received at the main directory number are routed through the JTAPI subsystem to the IPCC Express script.

In Figure 3-2, the dotted line around the Cisco CallManager and IPCC Express servers represents logical instances of the CTI route point and CTI ports. This depiction shows that the IPCC Express and Cisco CallManager servers both have handles to these route points and CTI ports, but each uses these handles differently. In particular, IPCC Express does not process the internal CTI route point and CTI port messages. Instead, IPCC Express makes CTI requests for Cisco CallManager to perform the call routing. Cisco CallManager then processes the CTI Redirect message from the route point to an available CTI port. Finally, the CTI port performs a consultative transfer to an available agent.

Table 3-1 maps IPCC Express logical devices (application providers) to their equivalent Cisco CallManager devices

Table 3-1 Mapping IPCC Express Application Providers to Cisco CallManager CTI Resources

IPCC Express Application Provider	Cisco CallManager CTI Resources	Description
IPCC Express IVR/CTI JTAPI Provider (CTI components)	CTI route points	Main directory number fronting incoming calls to open call handling sessions
	CTI ports	Call handling sessions for incoming calls
IPCC Express Agent JTAPI Provider (Agent Desktop components)	CTI ports	IPCC Express agent softphones
	Third-party control phones	Monitoring of: • Cisco IP Phone (not media termination) • Media-terminated desktop devices
Cisco Agent Desktop Client Machine JTAPI Provider	IP Phone extensions	The JTAPI Client on the Cisco Agent Desktop authenticates with the agent's UserID. In Cisco CallManager, the agent's UserID has devices associated with it. The JTAPI Client takes control of the agent's device line for call control capabilities.

Note that there are no equivalent media termination ports mapped to Cisco CallManager because the IPCC Express media termination ports, from the perspective of Cisco CallManager, are the CTI ports themselves. IPCC Express imposes a media port license that blocks dialog channels from being able to establish Real-Time Transport Protocol (RTP) streaming connections. Therefore, when provisioning, make sure the number of media ports does not exceed the total number of provisioned CTI ports per group, as indicated in the following equation:

Total number of media ports ≤Total number of CTI ports

Provisioning Cisco CallManager Resources

After determining how many IPCC Express resources are required for your contact center or IVR solution, the next step is to determine which Cisco CallManager resources (if any) are needed for your IPCC Express solution.

As indicated in Table 3-1, IPCC Express uses the following logical devices, or application providers:

- CTI and IVR JTAPI provider — Consists of route points and CTI ports per caller session.

- Agent JTAPI provider — Consists of all phones that can be used by logged-in agents.
- Cisco Agent Desktop Client Machine JTAPI providers — Each agent that uses Cisco Agent Desktop also requires one application provider per Cisco CallManager CTI connection.

Each application provider function maps to a related Cisco CallManager CTI resource.

Note Currently, Cisco strongly recommends that you configure both the CTI/IVR JTAPI and Agent JTAPI application providers on the same Cisco CallManager and CTI Manager server. The backup Cisco CallManager and CTI Manager server should be used as a hot standby. Ideally, at least one backup CTI Manager should be configured in the CTI/IVR JTAPI and Agent JTAPI (ICD) subsystems. Cisco recommends that you do *not* use the database publisher as the primary CTI Manager because the publisher would then become a single point of failure. Therefore, if you have two active Cisco CallManager servers for your IPCC Express, place the primary CTI Manager on the subscriber and the backup CTI Manager on the publisher.

Provisioning IPCC Express Agents

You must provision IPCC Express agents in two places:

- IPCC Express Agent (ICD) subsystem
- Cisco CallManager device pools (as part of the Cisco CallManager CTI design)

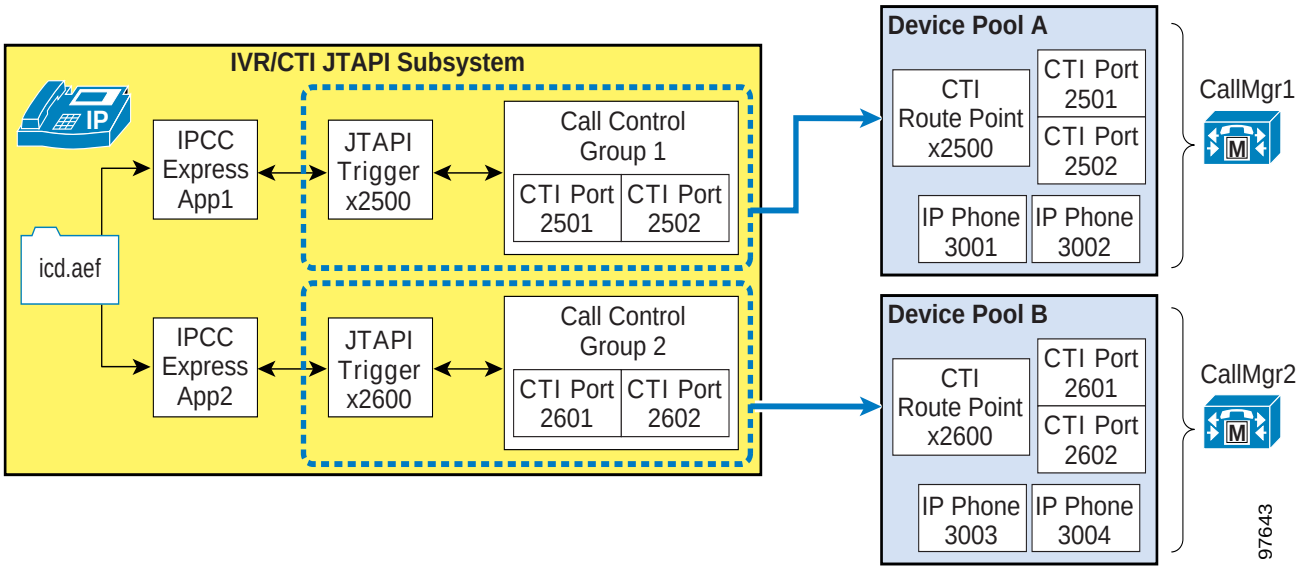
The following agent provisioning configurations are possible:

- One pool of agents shared among multiple IPCC Express scripts.
- One pool of CTI ports shared among multiple IPCC Express or IP IVR scripts. (For example, all CTI ports in Cisco CallManager device pool X can be assigned to multiple CTI port groups in the IPCC Express JTAPI subsystem.)
- NxN mesh of agents and ports shared among N scripts. (For example, if the agents and CTI ports are in device pools X and Y, the agents can be assigned to one or more IPCC Express resource groups, and the CTI ports can be assigned to one or more CTI port groups.)

Provisioning CTI Port Groups

In IPCC Express, CTI devices can be pooled into CTI port groups (also known as call control groups) and assigned to triggers. As previously mentioned, CTI ports can be related to individual IVR or IPCC Express sessions. Figure 3-3 illustrates an example of how you can distribute IPCC Express resources across different Cisco CallManager device pools and CTI port groups.

Figure 3-3 Grouping IPCC Express Agents and CTI Ports



Within the IPCC Express, call control groups (CTI port groups) and dialog groups are assigned to triggers. Triggers are assigned to an application, and the application also has an associated script, such as icd.aef. A single application can have multiple triggers and, depending on the trigger configuration, could be associated dynamically with ports from different call control groups and dialog groups. Agents within a single resource group can also be distributed among multiple Cisco CallManager device pools and groups. This flexibility in distributing JTAPI triggers, call control groups, and agent resource groups can be beneficial for redundancy purposes if there is a Cisco CallManager or CTI Manager failure within a Cisco CallManager cluster. For more information on redundancy considerations, see the chapter on Design Considerations for High Availability, page 4-1.

Design Considerations for High Availability

This chapter presents design considerations to help ensure that your IPCC Express applications remain available for use even under certain fault conditions. Various system failure scenarios are presented, along with the expected effects on IPCC Express agents and calls in those scenarios.

Designing for Fault Tolerance

IPCC Express applications predominantly focus on providing some type of customer service (for example, bank transactions, customer order status, call center forwarding, and so forth), where unavailability of the system could have a direct impact on business revenue. Therefore, the IPCC Express applications must maintain a high level of availability to ensure reliable service for business customers. This includes ensuring a seamless transition to backup systems during a failure scenario.

IPCC Express applications can be configured for CTI redundancy across Cisco CallManager servers in a single cluster. The IPCC Express JTAPI client connects to the primary CTI Manager (CTIM) and is aware of secondary and tertiary CTIMs on other Cisco CallManager servers. CTIM uses the same intra-cluster communication Signal Distribution Layer (SDL) mechanism that Cisco CallManagers in the cluster use to communicate with each other.

At a minimum, perform the following steps to enable CTIM fault tolerance in an IPCC Express solution:

- Set up a Cisco CallManager cluster that includes at least one publisher and one subscriber, then configure a Cisco CallManager group that includes the publisher and primary subscriber servers (plus any other secondary subscribers). All Cisco CallManager servers configured for IPCC Express should be running CTIM. (The CTI Manager Service is installed automatically during Cisco CallManager installation, and it is configured to start automatically with Windows 2000 Server, unless otherwise specified by the user.)
- For the IPCC Express JTAPI and Integrated Contact Distribution (ICD) subsystems, configure more than one CTIM.
- Ensure that the JTAPI preferences for each Cisco Agent Desktop point to a primary and a redundant CTIM.

The following sections describe failure scenarios for CTI redundancy:

- Cisco CallManager and/or CTI Manager Fails, page 4-2
- IPCC Express Server Fails, page 4-4

Furthermore, the following single points of failure exist in an IPCC Express solution under the specified conditions:

- Cisco CallManager publisher fails

  This failure is an issue only where the publisher is serving as the Domain Controller Directory (DCD) server. Primary CTIM failure is not of itself a single point of failure unless there is only one Cisco CallManager server or IPCC Express is a co-resident configuration.

- IP WAN fails in a centralized call processing model (with IPCC Express at the central site)

  Agents at the remote sites lose connectivity to either the IPCC Express application or to the Cisco CallManager cluster.

- IPCC Express Server fails

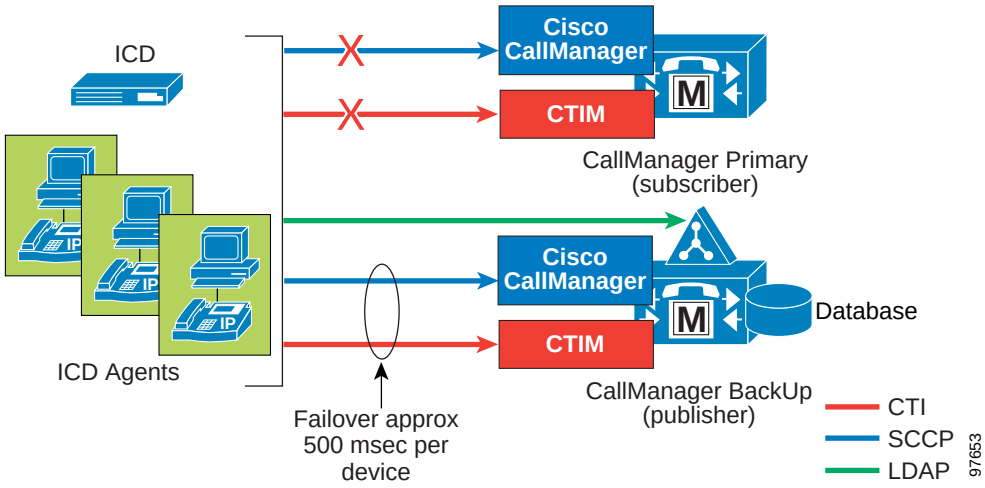
  If there is an IPCC Express application failure, all IPCC Express subsystems and application functions are unavailable, even though CTIM and Cisco CallManager services are running normally.

In all three scenarios, the IPCC Express application's CTI connection to its configured primary CTIM fails, thereby affecting call routing and IPCC Express agent connections. The following sections focus on anticipated behavior of call routing as well as call and agent availability after the failure. You should review these failure cases in order to address and appropriately set user expectations in providing redundancy for your IPCC Express system.

Cisco CallManager and/or CTI Manager Fails

In this failure scenario, IPCC Express loses its CTI connection to the primary Cisco CallManager or CTI Manager configured for IPCC Express. (See Figure 4-1.)

Figure 4-1 IPCC Express Recovery from Cisco CallManager or CTIM Failure (with Two Servers)



For the example in Figure 4-1, the IPCC Express JTAPI provider is pointing to the Cisco CallManager primary (subscriber) server for both its primary Cisco CallManager and CTIM services.

The following IPCC Express behavior can be expected of IPCC Express if the connection to Cisco CallManager primary is lost:

- The JTAPI and Integrated Contact Distribution (ICD) subsystems reconnect to the backup CTIM, as it was configured in the IPCC Express JTAPI and ICD subsystem parameters.
- All of the application provider resource management gets transferred to the backup server, which in this example is the publisher.

CTI devices (CTI ports and route points) re-home to the backup Cisco CallManager server. (Cisco recommends that you do not configure any devices on the publisher in a Cisco CallManager cluster with multiple servers.) Re-registration takes approximately 500 msecs per single line device or CTI port. During this re-registration period, there could be 15 to 20 seconds of call center downtime, and this delay can vary with the number of CTI devices in use. Once all devices have been re-registered to the backup Cisco CallManager and CTIM, routing of new incoming calls will resume, and IPCC Express agents can log in again and accept incoming calls.

If the failed Cisco CallManager is the DCD server authenticating the IPCC Express (the Directory Host Name defined in IPCC Express Configuration and Repository, Directory Setup parameters), there is no IP application redundancy because the failure of a Cisco CallManager hosting DCD is a single point of failure for IP IPCC Express. The following behavior can occur:

- LDAP authentication errors will occur in IPCC Express Application Administration.
- All incoming calls to the IPCC Express route point will return reorder tone (fast busy).
- Agents will not be able to log in again.

For this reason, the subscriber server should be configured as the primary CTIM and Cisco CallManager. An option for reducing the downtime is to install a DCD on a standalone Cisco CallManager server in the cluster or on an external corporate directory server.

Call Survivability

During Cisco CallManager or CTIM failure, expect the following call survivability behavior:

- Calls in progress with IPCC Express agents prior to failure are not interrupted.
- Calls in queue during failure are dropped.
- The average time for failover for a single server with CTI devices ranges from 15 to 20 seconds. During this time, incoming calls receive a busy (reorder) tone until all CTI devices have re-registered to the backup Cisco CallManager and CTIM.

You can use the following option to provide call continuity during this failure scenario:

- Distribute the agents and CTI ports into multiple device pools on Cisco CallManager.
- Have route points, CTI ports, and agent devices in a non-failed device pool to continue queuing calls and routing them to agents.
- If the failed Cisco CallManager and CTIM does not host the IPCC Express application route point, agents in a non-failed device pool can continue to accept calls from the queue with no downtime.

Note Adding a second IPCC Express application with a forward-on-failure to a secondary route point does not appear to improve failover time. Testing has shown that the time for Cisco CallManager to detect that the application is no longer available (and forward to a second route point) is equal to the time it takes for Cisco CallManager to re-register CTI devices to the backup server.

IPCC Express Agent Impact

As mentioned previously, calls in progress to agents prior to failure will survive. All agents are automatically logged out of their Cisco Agent Desktop. Agents receive a message indicating that either their phone or Cisco CallManager is offline. Agents might also observe the following behavior:

- Agent's desktop might go into a Reserved state, and all of the hot buttons are grayed. In such a case, that agent cannot become available to accept calls.
- Agent's IP phone might display a TEMP FAIL message. This event is related to a Cisco CallManager survivability feature, Quiet Clear, which sends a DeviceUnregistered message from Cisco CallManager to the phone where a call state is active during a failover.

The Cisco Agent Desktop will re-register to the CTIM configured in the JTAPI preferences on the local PC. The agent will have to log back in twice, once after failover and once after failback. All agents without active calls who log in again will have their agent states reset to the default Not Ready state, and they will have to make themselves available again to accept incoming calls. Agents with calls in progress prior to the failure must hang up (that is, release control of the old call) before they can become ready to accept new calls.

Cisco Agent Desktop VoIP Monitor server uses information in the SQL server database on the Cisco CallManager publisher to monitor calls silently. It needs this information to begin a monitoring session, but it does not require access to Cisco CallManager after a monitoring session has begun. If the SQL server or the connection to it fails, the currently active voice monitoring sessions are not interrupted because the VoIP Monitor server does not realize that failover has occurred. However, the first attempt to start a voice monitoring session will fail. The failure may take up to one minute if the failure is because the Cisco CallManager's IP address is not accessible. Subsequent attempts to monitor will try to connect to other Cisco CallManagers (subscribers) in the cluster until a connection is made. This process can take up to five minutes, depending on how many Cisco CallManagers there are in the cluster and how many of them are running. For example, if there are five Cisco CallManagers and they are all down and inaccessible on the network, the VoIP Monitor server will try each in succession. Each attempt can take up to a minute, for a total of five minutes.

IPCC Express Server Fails

This scenario focuses on the impact if the IPCC Express application server fails.

IPCC Express Availability

Currently, there is no failover or hot standby mechanism for the IPCC Express server. If the IPCC Express server hosting the IPCC Express application fails, the following conditions apply:

- No automatic failover mechanism for IPCC Express agents
- No intra-cluster communication among IPCC Express servers that will maintain transitional states in the event of a IPCC Express application service failure

There is, however, an IPCC Express recovery mechanism that can be performed manually through hard disk mirroring and drive swapping over to an identical Cisco Media Convergence Server (MCS). This recovery method is generically referred to as *cold standby*. For configuration details on this manual system recovery process for IPCC Express, refer to the section on IPCC Express Server Recovery – Cold Standby Server Configuration, page 4-5. Future releases of IPCC Express will have a warm standby option.

Call Survivability

During failure of the IPCC Express application server, expect the following call survivability behavior:

- Calls in progress with IPCC Express agents prior to the IPCC Express failure are not interrupted.
- Incoming calls from either the PSTN or the local network will receive a reorder (fast busy) tone until the IPCC Express server has recovered.

IPCC Express Agent Impact

During failure of the IPCC Express application server, agents might observe the following behavior:

- All agents are automatically logged out of their Cisco Agent Desktop. Agents receive a message indicating that either their phone or Cisco CallManager is offline.
- Agents are not able to log back in until the IPCC Express server is restored. Any attempt to do so will return a licensing error message (which means that the share to the IPCC Express server is not available).
- Cisco Agent Desktop will not re-home if you simply move the network cable to an identically configured IPCC Express server (unless you use the recover method described in IPCC Express Server Recovery – Cold Standby Server Configuration, page 4-5). Agents will not be able to log in and will receive an error message. Similarly, if the IP address or subnet mask of the IPCC Express server is changed, IPCC Express will have to be reinstalled to remove these errors and allow agents to log in again.

IPCC Express Server Recovery – Cold Standby Server Configuration

An *identically configured cold standby* implies one of the following recovery methods:

- An image of the IPCC Express server's disk drive is placed on a backup server (the cold standby). You can create the image by using either Ghost or similar imaging tools, or by using the replication mechanism in RAID. To be identically configured, the cold standby must contain a periodic snapshot of a known working image of the primary IPCC Express server.
- The redundant disk drive in the failed IPCC Express server is moved into a backup server (the cold standby). To be identically configured, the cold standby must contain a mirrored drive that provides a backup copy of the failed IPCC Express server.

Perform the following steps to recover IPCC Express with a redundant disk and a cold standby server.

Step 1 Install and configure the primary IPCC Express server, with all agents up and running (Cisco Agent Desktops and/or Supervisor Desktops logged in and agents accepting calls from queue).

Step 2 Install a cold standby server in a location near the primary IPCC Express server. Leave the cold standby unplugged from the network in a powered-down state.

Step 3 If a failure of the primary IPCC Express server occurs, all callers in the queue will be disconnected. Agents with calls already in progress will be able to complete those calls; however, their desktops will be logged off. All agents should exit their desktops in preparation for the cold standby server to go online.

Step 4 Remove the redundant drive (for example, disk 1, or ID 1) with the good configuration from the primary IPCC Express server. The Cisco MCS-7835 supports hot-swappable drives, so you can remove the disk drive without turning off the system power.

Step 5 Unplug the failed primary server's network cable, and move that same cable to the Network Interface Card (NIC) on the cold standby server. Keep the network cable in the same VLAN and subnet on the switch port.

Step 6 Insert the redundant drive (for example, disk 1, or ID 1) into the cold standby's disk array. Insert the drive while the standby server's power is still off.

Step 7 Boot up the cold standby server with this disk. At the boot-up prompt, press F2 to declare the empty bay as failed. (The F1 option ignores the missing drive and continues.) Automatic data recovery is enabled, and the server will boot into an identical copy of the primary IPCC Express server.

According to *Configuring and Using Redundant Disks with Cisco MCS* (available online at http://www.cisco.com/warp/public/788/AVVID/disk_redundancy_mcs_9229.html), the time required for an MCS rebuild is approximately 15 minutes per GB (time for the drives to sync up). The actual rebuild time is dependent upon the rebuild priority set, the amount of I/O activity occurring during the rebuild operation, the number of drives in the array, and the disk drive speed.

Note The IPCC Express application will still be available during this sync-up period.

Step 8 Once the cold standby server is up and running, check that the JTAPI subsystem is in IN_SERVICE state on the Application Administrator Engine Status page. Agents should then be able to start their desktops, log back in, and accept calls.

Note It is important to verify that both the server and the Cisco Agent Desktops and Supervisor Desktops have started up after the failure because they do not have the ability to switch over automatically. Each client application has to re-discover and re-establish a share to the cold standby IPCC Express server.

Normal contact center operations are interrupted for only as long as it takes to move the drive to the cold standby's array, write the redundant drive's information to the cold standby disk, and log agents back in.

Step 9 Once the primary IPCC Express server is restored, power down the primary and move the network cable back to this server. If add or modify changes have occurred on the cold standby in the interim, it may be necessary to take the drive from the cold standby's array and place it back into the primary. Boot the primary with the redundant drive, as described in Step 7. (A general field practice is to remove all of the drives and boot the restored primary with the redundant drive, then reinstall the original drive back into the server and mirror it.) Verify that agents can log back in, transition to a ready state, and accept calls.

Failure Scenario Summary

Table 4-1 summarizes the impacts from failures of particular components in the IPCC Express solution. The estimated downtime varies based on the number of CTI devices in use.

Table 4-1 Impacts of IPCC Express System Failures

Failed Component	Estimated Downtime[1]	Impact on IPCC Express	Impact on Agent (Cisco Agent Desktop)
Cisco CallManager server (CM1), with CTI Manager and/or DCD	91 seconds for 183 CTI devices	Application provider resource management transferred to CM2. Re-registration takes 500 msecs per single line device or CTI port.	Calls in progress with agents are not interrupted. All logged-in agents are automatically logged out, unless CM1 is the DCD. If the DCD fails, agents cannot log in again until the DCD is restored.
Cisco CallManager server (CM2), with IPCC Express route point and one device pool (PoolA)	15-20 seconds	Calls will be routed to agents with devices in PoolB. Callers dialing the route point during the 15-20 second failover hear reorder (fast busy). CM3 takes on the additional device weight load of CM2.	Calls in progress with agents are not interrupted. PoolA agents are logged out and must wait for their phones to re-register to CM3 before logging in again.
Cisco CallManager server (CM3), with one device pool (PoolB)	None	Calls will be routed to agents with devices in PoolA. CM2 takes on the additional device weight load of CM3.	Calls in progress with agents are not interrupted. PoolB agents are logged out and must wait for their phones to re-register to CM2 before logging in again.

1. Estimated downtime depends on the number of CTI resources in use.

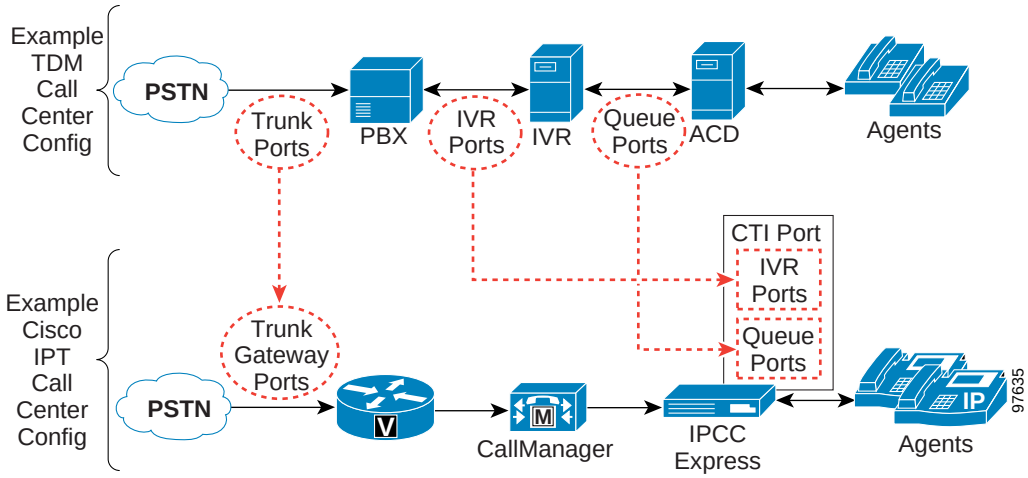
Basics of Call Center Sizing

This chapter introduces the basic concepts involved in call center sizing.

Terminology

Figure 5-1 illustrates the common port types and how they map to IPCC Express.

Figure 5-1 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 IPCC Express.
- **Prompt-and-collect (P&C) ports** — perform basic P&C and/or queuing only. P&C ports are IVR ports that have limited functionality, and they are available for both IPCC Express Standard and IPCC Express Enhanced. These ports perform basic call treatment of P&C only, and they do *not* provide the capability to do database dips, email, ASR/TTS, VoiceXML, or HTTP. Both IPCC Express Standard and IPCC Express Enhanced have basic IVR capability in the form of P&C ports, and you can have as many P&C ports as the hardware server supports. (See Server Capacities and Limits, page A-1.) These ports are included at no additional cost with IPCC Express Standard or

Enhanced, but they must be sized for proper capacity planning for the IPCC Express server. For sizing and capacity details, refer to the IPCC Express Configuration and Ordering Tool, available online at

http://www.cisco.com/en/US/partner/products/sw/custcosw/ps1846/prod_how_to_order.html

- **Queue ports** — are IVR ports that queue calls prior to transferring the caller to an available agent. These ports are included at no additional cost with IPCC Express Standard or Enhanced, but they must be sized for proper capacity planning for the IPCC Express server. Refer to the IPCC Express Configuration and Ordering Tool for more details.
- **IVR ports** — are full-featured IVR ports with all the capabilities found in the standalone Cisco IP IVR product, except that the IPCC Express IVR ports do not support Intelligent Contact Management (ICM) integration.
- **CTI ports** — are computer software devices used to handle telephony hardware devices such as IVR ports and queuing ports. CTI ports include all three types of IVR port functions for P&C, queuing, and IVR port options for database dips, ASR, and TTS. CTI ports are an important component in the IPCC Express architecture, but you do not have to purchase them separately.

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, then IPCC Express would need additional functionality in the IP IVR to perform these operations. Because the basic P&C ports cannot handle additional features such as these, you would have to purchase additional IVR port licenses. These IVR ports, in addition to functioning as P&C ports, will also perform database dips, VoiceXML, ASR, TTS, web operations, and so forth.

The goal of the system architect is to determine the appropriate number and types of IVR ports to provision for the IPCC Express system. However, as shown in Figure 5-1, the IPCC Express 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

Cisco recommends that your system designers create a sizing document to do the following:

- Scope out the preliminary configuration information for the IPCC Express 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 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 5-1.

Table 5-1 Call Center Sizing Metrics

Metric	Description
Average handle time (AHT)	Average duration (talk time) of a call plus after-call work time, which is the wrap-up time after the caller hangs up.
Average IVR port usage time	The total time for prompt playout and/or menu navigation (if any) in the IPCC Express script.
Service level goal for agents	Percentage of calls answered by agents within a specific number of seconds.
Busy Hour Call Attempts (BHCA)	Average number of calls received in a busy hour.
Grade of service (% blockage) for gateway ports to the PSTN	Percentage of calls that get a busy tone (no gateway trunks available) out of the total BHCA.

All of the metrics in Table 5-1 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 IPCC Resource Calculator available at: http://tools.cisco.com/partner/ipccal/index.htm.

The IPCC 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 IPCC Express system properly.

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

Note If the system being designed is a replacement for an existing IPCC Express or IP IVR system, you might not need all of the information listed above. You might be able to use the current agents, call flow, and historical reporting information from the existing system to size the new system (assuming there are no changes in the application, load, call flow routing, or service level desired).

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 5-2 illustrates the principal steps and design considerations for sizing a call center.

Figure 5-2 IPCC Express Design Process – Call Center Sizing

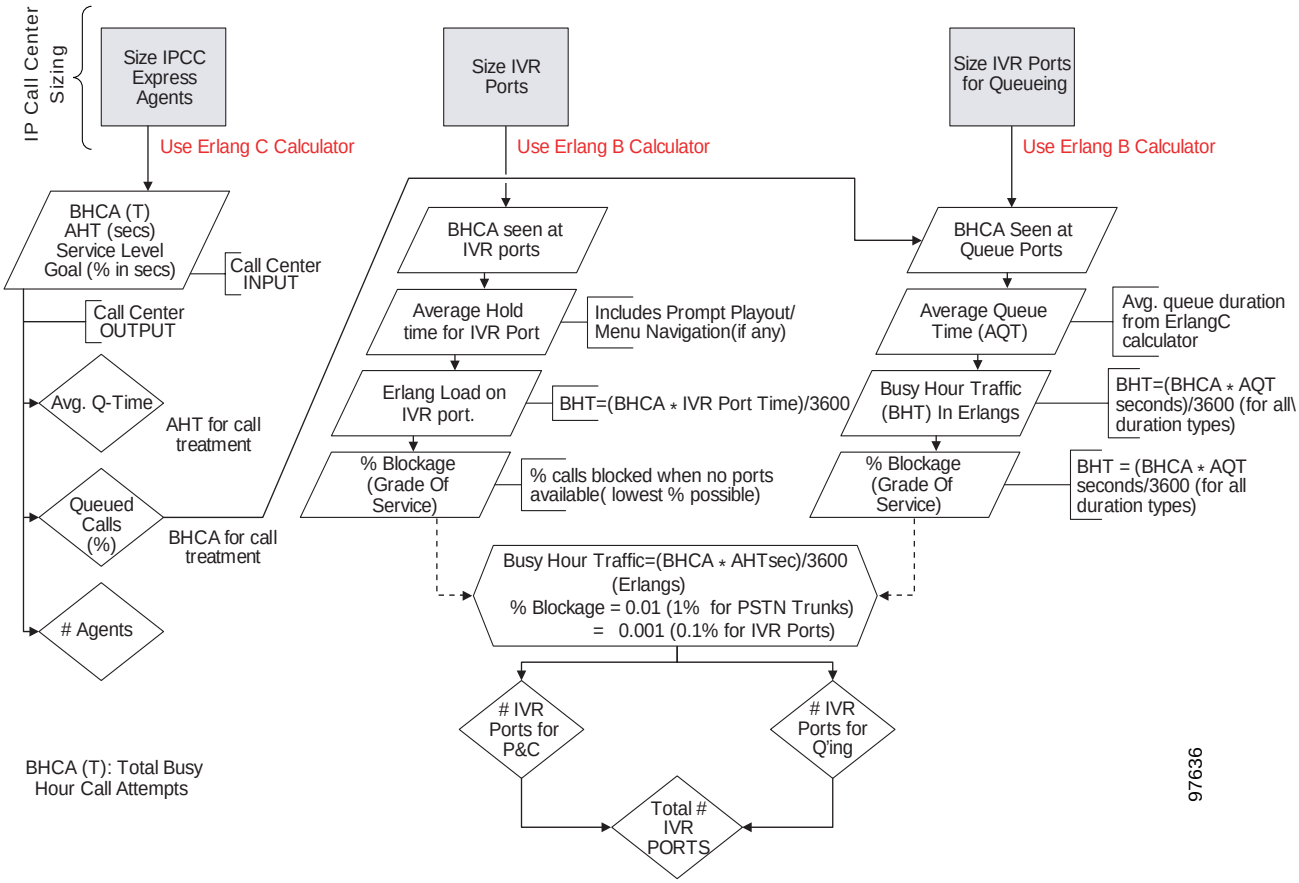


Figure 5-2 is a general overview of the design considerations for call sizing. For a detailed description of the call center sizing design process, refer to the section on sizing call center resources in the *Cisco IP Contact Center Solution Reference Network Design Guide*, available online at

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

There are similar basic call center sizing considerations and steps for IPCC Enterprise, and they also can be used in sizing a smaller contact center for IPCC Express. 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 IPCC Express, 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.
- All call center designs must be sized to the system correctly. Do not size a call center without using the IPCC Express Configuration and Ordering Tool to determine the required quantity of servers and gateway trunks.

- 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 IPCC Express 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.
- Running the Historical Reporting client on a co-resident Historical Reporting Database is a CPU-intensive process.
- 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 IPCC Express 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 IPCC Express system.

Note Not all of the IPCC Express system limits are available at the same time. In particular, the number of calls in queue can impact a system limit such as the maximum number of agents in a resource group.

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

http://www.cisco.com/en/US/partner/products/sw/custcosw/ps1846/prod_how_to_order.html

Planning Resource Requirements for Call Center Sizing

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

Example of Sizing IPCC Express 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.

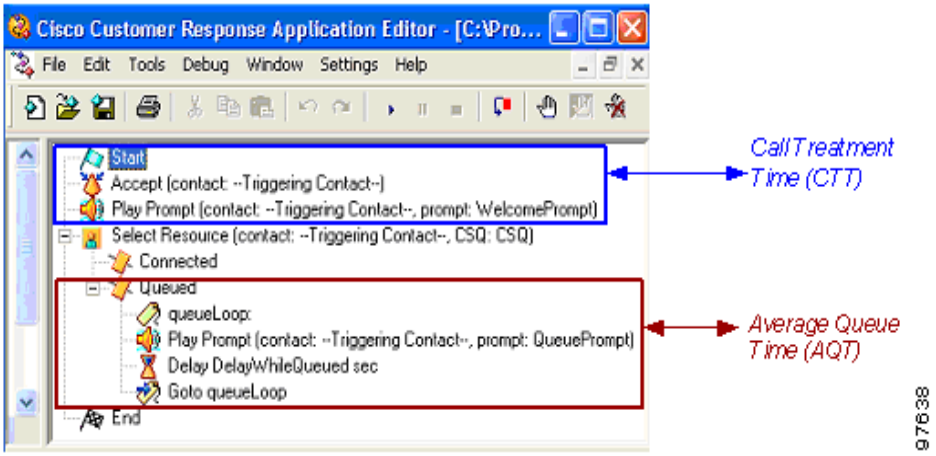
The following information applies to this example of IPCC Express Standard with 25 agents:

Metric	Metric Value
Busy Hour Call Attempts (BHCA)	800 calls in 60-minute interval
Service level goal	90% of all calls handled within 15 seconds
Average handle time (AHT)	90 seconds: • Average talk time = 90 seconds • Wrap-up time = 0 seconds
Wait before Abandon	120 seconds
Grade of service (% blockage) for gateway ports to the PSTN	1% (0.01)

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 IPCC Express Configuration and Ordering Tool indicates that all of these parameters fit within a single-server IPCC Express system.

The IPC Resource Calculator also uses Erlang B and C to calculate the number of IVR ports needed for call treatment (prompt/collect) and queuing. As an example of this calculation, we use the default icd.aef script logic that is available with all the Cisco IPCC Express packages. Note in Figure 5-3 how the script logic allows the application developer to insert various delays in the script; these delays must be included in Call Treatment Time, (Average IVR Delay) in put to the IPC Resource Calculator.

Figure 5-3 Application Processing Time for IPCC Express



The following steps detail the procedure for calculating IVR ports for our example IPCC Express 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 IPCC Express 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 IPCC Express 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 2 seconds into the IPC Resource Calculator, and notice that in this example five 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 six IVR ports required for queuing.

Step 3 Calculate the total number of IVR Ports required.

If the same IVR resources are shared for both Call Treatment and Queuing, then the IPC Resource Calculator can be used to determine that in this example a total of seven IVR ports are required. If different IVR resources are used for Call Treatment and Queuing, then simply use the answers calculated in Steps 1 and 2 for the different IVR Resources.

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, 31 trunks (DS0's) are required.

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.

For more gateway sizing guidance, refer to the *Cisco IP Telephony Solution Reference Network Design* documentation available online at

http://www.cisco.com/warp/public/779/largeent/it/ese/srnd.html

Sizing the IPCC Express Server

With the call sizing guidelines from the preceding chapters, we are able to calculate the number of agents and CTI ports (queuing and P&C IVR ports) required to deploy a contact center solution. What does this mean in terms of sizing the IPCC Express resources? When do we know that we have reached maximum capacity on the IPCC Express system and need to add another server? These are some of the questions addressed in this chapter.

To help the designer estimate overall scalability on the IPCC Express server, Cisco has conducted extensive performance testing, and the section on Impact of Performance Criteria on the IPCC Express Server, page 6-3, summarizes the critical factors and their impact on performance.

Configuration and Ordering Tool

The IPCC Express Configuration and Ordering Tool uses a points system to size the server automatically based on your specific configuration. You are not required to size the server manually using this points system, but Table 6-3 lists the point values for reference only. The Configuration and Ordering Tool for IPCC Express and IP IVR is available online at

http://www.cisco.com/en/US/partner/products/sw/custcosw/ps1846/prod_how_to_order.html

Note BHCA rate has an impact on scalability and is not explicitly called out in the IPCC Express Configuration and Ordering Tool; however, the impact of the BHCA is implicitly accounted for in the device quantities in the input section of that Tool.

The Configuration and Ordering Tool enables you to configure and order your system using step-by-step instructions in the following basic order (the exact steps may vary as the tool is enhanced for future releases):

Step 1 Select the deployment type (for example, standalone or co-resident, upgrade or new order). Also choose the type of hardware server on which you will deploy the software product you are configuring. Each type of hardware server has a specific “total available server points” value associated with it. As you go through the rest of the configuration, each feature and quantity you select will be deducted from the total available server points.

Then choose the specific product that you want to configure. You must select (or check for) correct values for this server configuration for each and every gray cell in the tool (Excel based). If you do not, the results might not be correct.

Step 2 Select the server, the SmartNet option (if applicable), and the feature(s) for the respective Feature Servers available. Based on your selections and completion of the remaining feature Tabs, the configured feature points will be off-loaded and allocated to the respective Feature Server from the main server. Some examples of Feature Servers include:

Server	Uses
Main Server	ACD, CTI, IVR, VoiceXML, Recording and Monitoring Server (RMS), Historical Reports Database Server (HRDB), ASR, TTS
Recording and Monitoring Feature Server	IPCC Express Call Recording and Monitoring Server (RMS1)
ASR/TTS Feature Server	Automatic Speech Recognition Server (ASR), Text-to-Speech Server (TTS)
Historical Reporting Feature Server	Historical Reports Database Server (HRDB)

The Configuration and Ordering Tool lists more server options.

Step 3 Configure IPCC Express (agents, supervisors, and so forth) and CTI ports (prompt-and-collect, full IP IVR ports, and so forth). Enter or check for correct values for the above server configuration you choose for each and every gray cell in the tool. If you do not, the results might not be correct.

In this step of the Tool, you are configuring for the type of deployment, hardware server, and product. The value shown for the server type currently being configured comes from the values you entered in Step 1.

Step 4 Configure the Call Monitor and Record features. Enter or check for correct values for the server configuration for each and every gray cell in the tool. If you do not, the results might not be correct.

The value shown for the server type currently being configured comes from the values you entered in Step 1 and/or Step 3.

Step 5 Configure the Historical Reporting feature. Enter or check for correct values for this server configuration for each and every gray cell in the tool. If you do not, the results might not be correct.

The value shown for the server type currently being configured comes from the values you entered in Step 1 and/or Step 3.

Step 6 Configure IVR ports, ASR, TTS, and VoiceXML. Enter or check for correct values for this server configuration for each and every gray cell in the tool. If you do not, the results might not be correct.

The value shown for the server type currently being configured comes from the values you entered in Step 1.

This last step will give you the status of your configuration and order, and will indicate whether it is approved or if it will require Bid Assurance approval.

Use the output of the Configuration and Ordering Tool to determine whether or not your IPCC Express system design is a valid configuration. If not, the options are to reduce or remove the use of features such as simultaneous recording sessions or historical reporting sessions, to reduce the number of agents or IPCC Express related devices, or (if appropriate) to move one or more of the functions to a separate feature server that accommodates the desired configuration.

Note The IPCC Express points system applies only to the IPCC Express server and does not consider Cisco CallManager scalability. Sizing Cisco CallManager servers requires the use of the latest Cisco CallManager capacity sizing tool. Check with your Cisco Systems Engineer (SE) or Partner for the latest Cisco CallManager capacity tool.

Impact of Performance Criteria on the IPCC Express Server

System performance criteria fall into two general categories:

- IPCC Express and IP IVR components — Applications, capabilities, and options 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 Cisco IPCC Express or IP IVR system. In general, the more Cisco IPCC Express or 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.

List of Performance Criteria

Table 6-1 lists the criteria that can affect the performance of system hardware, and in turn the performance of the Cisco IPCC Express system. Cisco conducted extensive testing to quantify the effect of many of the performance criteria, and a point value was assigned to each of these effects. Table 6-3 lists these point values.

Table 6-1 System Performance Criteria

Criterion	Effect on System
Busy-hour call completion rate	As the call rate increases, more resources are consumed on the Cisco Customer Response Solutions (CRS) server and on the Cisco CallManager server.
Average call duration	Longer average call duration means a lower busy-hour call completion rate, which lowers CPU usage.
Trace enabled	CPU resource consumption: Varies depending on trace level enabled. Cisco CallManager tracing can consume significant additional CPU resources on a co-resident system.
Number of IVR ports	CPU resource consumption: Light.
IVR software redundancy option	No effect.
Number of IP Queue Manager ports	CPU resource consumption: Light.
HTTP traffic	CPU resource consumption: Light.
Number of TTS ports	CPU resource consumption: Moderate.

Table 6-1 System Performance Criteria (continued)

Criterion	Effect on System
ASR or TTS with VoiceXML	CPU resource consumption: • Heavy • Varies depending on the grammar used
Number of ASR ports	CPU resource consumption: • Heavy • Varies depending on the language and the complexity of the grammar
ASR grammars	CPU resource consumption: Varies depending on the complexity of the grammar.
Languages	CPU resource consumption: • CPU usage varies depending on the language. • Memory usage is heavy for each additional language.
On-demand recording session	CPU resource consumption: • Heavy for G.711 codec. • Very heavy for G.729 codec. Each minute of recording takes approximately 1 MB of disk space. **Note** To prevent on-demand recording sessions from consuming CPU resources on the Cisco IPCC Express or IP IVR server, you can set up a dedicated Call Statistics, Recording, and Monitoring Server. With this configuration, on-demand recording sessions will consume resources on the dedicated server but will not consume resources on the Cisco IPCC Express or IP IVR server. (For information about setting up a Call Statistics, Recording, and Monitoring Server, refer to *Getting Started with Cisco Customer Response Applications*, available online at http://www.cisco.com/univercd/cc/td/doc/product/voice/sw_ap_to/apps_3_5/english/admn_app/index.htm.)
Number of IPCC Express agents (equivalent to an IP Phone agent)	CPU resource consumption: Moderate.
Media termination (agent without a hardware IP Phone)	No effect.
Number of supervisors	CPU resource consumption: Moderate.
ACD	Because an ACD system is a collection of many features, you must analyze the effect of each feature separately.

Table 6-1 System Performance Criteria (continued)

Criterion	Effect on System
Silent monitoring	CPU resource consumption: • Heavy for G.711 codec. • Very heavy for G.729 codec. **Note** To prevent silent monitoring sessions from consuming CPU resources on the Cisco IPCC Express or IP IVR server, you can set up a dedicated Call Statistics, Recording, and Monitoring Server. You can also set up one or more dedicated Call Monitoring Servers. With these servers, silent monitoring sessions will consume resources on the dedicated server but will not consume resources on the Cisco IPCC Express or IP IVR server. (For information about setting up these dedicated servers, refer to *Getting Started with Cisco Customer Response Applications*, available online at http://www.cisco.com/univercd/cc/td/doc/product/voice/sw_ap_to/apps_3_5/english/admn_app/index.htm.)
Historical Reporting session	CPU resource consumption: Very heavy (increases as the size of the Cisco IPCC Express and IP IVR databases increase). **Note** If you run historical reporting sessions on the Cisco IPCC Express or IP IVR server, Cisco recommends that you run these sessions during off-peak hours so that the generation of historical reports does not compete for server resources. **Note** To prevent historical reporting sessions from consuming CPU resources on the Cisco IPCC Express or IP IVR server, you can set up a dedicated Historical Reports Database Server. With this configuration, historical reporting sessions will consume resources on the dedicated server but will not consume resources on the Cisco IPCC Express or IP IVR server. (For information about setting up an Historical Reports Database Server, refer to *Getting Started with Cisco Customer Response Applications*, available online at http://www.cisco.com/univercd/cc/td/doc/product/voice/sw_ap_to/apps_3_5/english/admn_app/index.htm.)
Size of IPCC Express and IP IVR databases	CPU resource consumption: Increases as database size increases.

Supported Servers

Table 6-2 shows the point value for the maximum acceptable CPU usage for each type of supported server. It also shows the minimum memory each server requires in a deployment scenario, and the number of IP phones that each server supports.

Table 6-2 Point Values and Memory Requirements for Supported Servers

Server Model	Equivalent Model	Maximum Points Supported	Minimum Memory Required
Cisco MCS-7815-1000	IBM xSeries 200	1,000	1 GB[1]
Cisco MCS-7815 I-2.0-CC1	IBM 205-2000	1,000	1 GB[1]
Cisco MCS-7825-800	Compaq DL320	800	1 GB[1]
Cisco MCS-7825-1133	Compaq DL320	900	1 GB[1]
Cisco MCS-7825 H-2.2-CC1	Compaq DL320-2266 G2	900	1 GB
Cisco MCS-7825 H-3.0-CC1	Compaq DL320-3.0 GHz G2	900	1 GB
Cisco MCS-7835-1000	Compaq DL380	1,000	1 GB
Cisco MCS-7835-1266	Compaq DL380 G2	1,266	1 GB
Cisco MCS-7835 H-2.4-CC1	Compaq DL380-2400 G3 (single CPU)	1,266	1 GB
Cisco MCS-7835 I-2.4-CC1	IBM 345 2400 (single CPU)	1,266	1 GB
Cisco MCS-7835 H-3.0-CC1	HP DL 380-3.0 GHz G3 (single CPU)	1,266	1 GB
Cisco MCS-7845 H-2.4-CC1 (dual CPU, using Windows 2000 Advanced Server OS)	Compaq DL380-2400 G3 (dual CPU)	2,600	4 GB
Cisco MCS-7845 H-3.0-CC1 (dual CPU, using Windows 2000 Advanced Server OS)	Compaq DL380-3.0 GHz G3 (dual CPU)	2,600	4 GB
IBM 330-1266		1,000	1 GB[1]
IBM 342-1266		1,266	1 GB
IBM 345-2400		1,266	1 GB

1. You must add 512 MB of memory to this sever to bring its total memory to 1 GB.

Point Values for IPCC Express

The following table shows each performance criterion and its corresponding point value. You do *not* have to use these point values to configure your IPCC Express server manually; rather, they are intended as a guide for planning system capacity. To determine the actual configuration and capacity of your system, use the IPCC Express Configuration and Ordering Tool because it incorporates the latest revisions and values.

Table 6-3 Point Values for Performance Criteria

Performance Variable	Point Value	Notes
IVR Variables		
IP IVR server software	0	
IP IVR port	5	For the maximum number of IVR ports that are supported on a single server, see Server Capacities and Limits, page A-1.
IVR port for IPCC Express (call duration less than 2 minutes)	5	For the maximum number of IVR ports that are supported on a single server, see Server Capacities and Limits, page A-1.
IVR port for IPCC Express (call duration greater than 2 minutes)	4	For the maximum number of IVR ports that are supported on a single server, see Server Capacities and Limits, page A-1.
IVR Software Redundancy Option (per port)	0	• The number of ports in the backup IVR server does not affect the performance of the primary IVR server. • The configuration of the backup IVR server is assumed to be the same as that of the primary IVR server.
VoiceXML parser (per port)	4	• Each VoiceXML parser requires an IVR port, so add 5 points for each port that you will use. • Because VoiceXML invokes the ASR and TTS functions, add 16 points for each VoiceXML port that will be used. • The notes that apply to TTS and ASR variables also apply to the VoiceXML parser variable.
IP Queue Manager software	0	
IP Queue Manager port	5	Equivalent to IVR port.
Text-to-Speech (TTS) Variables		
TTS client and server port, for any language except German	8	• On an MCS-7825-800, an MCS-7835-1000, or an MCS-7825H-22-CC1, use this configuration only for a single language. • If installed, TTS consumes some system resources even if it is not being used. • Each TTS port requires an IVR port, so add 5 points for each TTS port that you will use. • These point values were determined for short TTS prompts. For longer prompts (text files greater than 5 KB, which equate to audio files greater than 2.5 MB), fewer ports are supported. You should make application-specific measurements before deployment. • The Point Value for a German TTS client and server port is 10 points.

Table 6-3 Point Values for Performance Criteria (continued)

Performance Variable	Point Value	Notes
Nuance Vocalizer 1.0 TTS; each language after the first language)	200	• If installed, TTS consumes some system resources even if it is not being used. • In some scenarios, to make more efficient use of server memory, Cisco recommends that you run the TTS server on a separate MCS. • Each TTS port requires an IVR port, so add 5 points for each TTS port that you will use. • These point values were determined for short TTS prompts. For longer prompts (text files greater than 5 KB, which equate to audio files greater than 2.5 MB), fewer ports are supported. You should make application-specific measurements before deployment. • If you move the TTS server to a separate MCS, the 200 points per port for each additional language apply to the TTS server on the separate MCS.
Nuance Vocalizer 3.0 TTS client and server port, for each additional language	10	• If installed, TTS consumes some system resources even if it is not being used. • In some scenarios, to make more efficient use of server memory, Cisco recommends that you run the TTS server on a separate MCS.
TTS client-only port	8	• If installed, TTS consumes some system resources even if it is not being used. • If a system is using multiple languages, each language consumes an additional 10 MB plus 2 MB per channel or license, whichever is higher. • Each TTS port requires an IVR port, so add 5 points for each TTS port that you will use. • These point values were determined for short TTS prompts. For longer prompts (text files greater than 5 KB, which equate to audio files greater than 2.5 MB), fewer ports are supported. You should make application-specific measurements before deployment.
Automatic Speech Recognition (ASR) Variables		
ASR Client and server port	20	• An MCS-7815-1000 (IBM xSeries 200) or an MCS-7815I-2.0-CC1 (IBM xSeries 205-2000) supports a maximum of 24 ASR client and server ports. • For the maximum number of ASR ports that are supported on other servers, see Server Capacities and Limits, page A-1. • Each ASR port requires an IVR port, so add 5 points for each ASR port that you will use. • If installed, ASR consumes some system resources even if it is not being used.

Table 6-3 Point Values for Performance Criteria (continued)

Performance Variable	Point Value	Notes
ASR client port	16	• For the maximum number of ASR ports that are supported on a single server, see Server Capacities and Limits, page A-1. • Each ASR port requires an IVR port, so add 5 points for each ASR port that you will use. • If installed, ASR consumes some system resources even if it is not being used. • Point values for an ASR client increase with a complex grammar, so Cisco recommends that you test a prototype of the desired speech recognition application to determine the point values for your specific scenario.
ASR; each language after the first language	180	• An MCS-7815-1000 (IBM xSeries 200) or an MCS-7815I-2.0-CC1 (IBM xSeries 205-2000) supports a maximum of 24 ASR client and server ports. • For the maximum number of ASR ports that are supported on other servers, see Server Capacities and Limits, page A-1. • Each ASR port requires an IVR port, so add 5 points for each ASR port that you will use. • If installed, ASR consumes some system resources even if it is not being used. • Cisco recommends that you use a separate ASR server for a scenario with more than one language. • If you move the ASR server to a separate MCS, the 180 points per port for each additional language apply to the ASR server on the separate MCS.
Other Variables		
Media Termination	0	•
Agent (call duration less than 2 minutes): • IP Phone Agent • Cisco Standard Agent Desktop • Cisco Standard Supervisor Desktop • Cisco Enhanced Agent Desktop • Cisco Enhanced Supervisor Desktop	10	• For the maximum number of agents and supervisors that are supported on a single server, see Server Capacities and Limits, page A-1. • A maximum of 40 agents are supported on an MCS-7825-800 or on an MCS-7825 H-2.2-CC1. • If a supervisor will perform silent monitoring or on-demand recording, also add the points for the appropriate feature.
Agent (call duration greater than 2 minutes)	7	Note that the types of agents and the notes are the same as for the previous row.

Table 6-3 Point Values for Performance Criteria (continued)

Performance Variable	Point Value	Notes
Cisco Standard or Enhanced Historical Reporting: • One session • Database size of 100,000 records	200	Add 20 points per session for each additional record, up to 500,000 records. Do not continue to add points for databases larger that 500,000 records. If you run Historical Reporting sessions on the Cisco IPCC Express or IP IVR Server during normal contact center hours of operation: • 2 Historical Reporting sessions can run at a time on an MCS-7845H-2.4-CC1 dual CPU using Windows 200 0 Advanced Server OS. • 1 Historical Reporting session can be run at a time on other servers. If you run Historical Reporting sessions on the Cisco IPCC Express or IP IVR Server outside normal contact center hours of operation: • 1Historical Reporting session can run at a time on an MCS-7825-800. • No more than 2 Historical Reporting sessions can run simultaneously on an MCS-7825-1133, MCS-7825H-2.2.-CC1, MCS-7835-1266, MCS-7835-1000, MCS-7835H-2.4-CC1, or an MCS-7825 H-2.2-CC1. • No more than 3 Historical Reporting sessions can run simultaneously on an MCS-7845 H-2.4-CC1 dual CPU. If you run Historical Reporting sessions on an Historical Reports Database Server: • No more than 6 Historical Reporting sessions can run simultaneously when using Microsoft SQL Desktop Edition (MSDE). • No more than 13 Historical Reporting sessions can run simultaneously when using Microsoft SQL Server 2000. (Microsoft SQL Server 2000 is supported in Cisco IPCC Express and IP IVR Release 3.1(2) and later.) If you run Historical Reporting sessions on a Call Statistics, Recording, and Monitoring Server or on a Call Monitoring Server: • No more than 3 Historical Reporting sessions can run simultaneously when using Microsoft SQL Desktop Edition (MSDE). • No more than 7 Historical Reporting sessions can run simultaneously when using Microsoft SQL Server 2000. (Microsoft SQL Server 2000 is supported in Cisco IPCC Express and IP IVR Release 3.1(2) and later.)
IPCC Express Enhanced server	0	
Cisco CTI option	0	

Table 6-3 Point Values for Performance Criteria (continued)

Performance Variable	Point Value	Notes
Cisco on-demand recording, using G.711 codec (per session)	20	• For the maximum number of call recording sessions that are supported under any deployment scenario, see Server Capacities and Limits, page A-1. • Each minute of recording takes approximately 1 MB of disk space, so allow enough disk space for the amount of recording that will be done. • Agents and supervisors can initiate recording.
Cisco on-demand recording, using G.729 codec (per session)	25	• For the maximum number of call recording sessions that are supported under any deployment scenario, see Server Capacities and Limits, page A-1. • Each minute of recording takes approximately 1 MB of disk space, so allow enough disk space for the amount of recording that will be done. • Agents and supervisors can initiate recording.
Barge-in	0	
Local silent monitoring, using G.711 codec (per session)	20	• Silent monitoring is not a feature that is sold separately, so add points for this feature if it will be used. • Only supervisors can initiate silent monitoring.
Local silent monitoring, using G.729 codec (per session)	25	• Silent monitoring is not a feature that is sold separately, so add points for this feature if it will be used. • Only supervisors can initiate silent monitoring.
Remote silent monitoring, using G.711 codec (per session)	35	• Remote silent monitoring is available in Cisco IPCC Express and IP IVR Release 3.1 and higher. • Remote silent monitoring is available only with G.711 codecs; G.729 codecs are not supported. • Only properly configured supervisors can access remote silent monitoring.
Call intercept	0	

Supported Co-Resident Scenarios

In a co-resident system, Cisco IPCC Express and/or IP IVR run on the same server with Cisco CallManager. A co-resident system is typically used in light load situations. Although the scenarios listed in this section are all supported, the following factors can affect the performance of your installation:

- Cisco CallManager features that you have implemented

  If several Cisco CallManager features are running simultaneously, there might not be enough CPU resources available for Cisco IPCC Express or IP IVR features. For example, music on hold (MOH) consumes significant additional CPU resources.

- Related applications and configurations

  For example, certain voice gateways may consume more or fewer resources than other voice gateways.

Other applications, including Cisco Personal Assistant (PA), Cisco Unity, Cisco WebAttendant, and Telephony Call Dispatcher (used by the Attendant Console application), are not supported on the same server as Cisco CallManager and the IPCC Express or IP IVR system.

Cisco IP IVR Supported Scenarios

The following co-resident IP IVR deployment scenarios are supported only on an MCS-7835-1266 or equivalent server. These scenarios have not been tested on any server other than an MCS-7835-1266 or equivalent, and therefore are not supported on any other servers.

Note A system with a complicated ASR grammar might require deployment on multiple servers.

The following table describes three supported IP IVR co-resident scenarios:

Scenario	Note
15 IVR ports	—
4 IVR ports with single-language ASR/TTS	TTS must be Vocalizer 1.0
4 IVR ports with VXML	TTS must be Vocalizer 1.0

Cisco IPCC Express Supported Scenarios

This section describes co-resident IPCC Express deployment scenarios that are supported. These scenarios have only been tested on the servers shown and therefore are not supported on any other servers.

The following table describes the supported IPCC Express co-resident scenarios, lists the servers on which each scenario is supported, and lists the busy-hour call completions (BHCC) tested for Cisco CRS and Cisco CallManager.

Scenario	Server	BHCC Tested
• 150 IP Phones • 10 Agents • 1 supervisor not taking ACD calls • 10 prompt and collect ports or 5 IVR ports • 1 monitoring and 1 recording session • 1 historical reporting sessions running during off-peak hours	MCS 7815I-3.0-CC1 (As of this writing, Cisco is not shipping this server. Please contact your Cisco SE for the latest part number and availability dates.)	• Cisco CRS—300 • Cisco CallManager—900
• 500 IP Phones • 10 agents • 1 supervisor not taking ACD calls • 10 prompt and collect ports or 5 IVR ports • 1 monitoring or 1 recording session • 1 historical reporting session running during off-peak hours	MCS 7825H-3.0-CC1 (HP DL 320-3.0 GHz G2)	• Cisco CRS—300 • Cisco CallManager—3,000

Scenario	Server	BHCC Tested
• 1,000 IP Phones • 10 agents • 1 supervisor not taking ACD calls • 10 prompt and collect ports or 5 IVR ports • 1 monitoring or 1 recording session • 1 historical reporting session running during off-peak hours	• MCS 7835H-1266 (HP DL 380 G2, IBM xSeries 330-1266, IBM xSeries -1266) • MCS 7835H-2.4 CC1 (HP DL380-G3 Single 2.4G Xeon) • MCS 7835I-2.4 CC1 (IBM xSeries 345 2400 Single Xeon) • MCS 7835H-3.0 CC1 (HP DL380-3.0 GHz G3 Single Xeon) • MCS 7835I-3.0 CC1 IBM xSeries 345 3.0 Single Xeon)	Cisco CRS—300 Cisco CallManager—6,000
• 3,000 IP Phones • 10 agents • 1 supervisor not taking ACD calls • 10 prompt and collect ports or 5 IVR ports • 1 monitoring or 1 recording session • 1 historical reporting session running during off-peak hours	MCS 7845H-3.0 -CC1 (HP DL 380 -3.0 GHz G3 Dual Xeon)	• Cisco CRS—300 • CallManager—18,000

IPCC Express Silent Monitoring and Recording Considerations

The silent monitoring and recording features of the IPCC Express agent desktop are implemented with a Voice over IP (VoIP) Monitor Server in IPCC Express. The following scalability limits apply to the VoIP Monitor Server and are based on Cisco Agent Desktop Release 4.5.5 server capacity data:

- You can have a maximum of four instances of the VoIP Monitor Server per logical contact center. You can use multiple VoIP Monitor Servers in a single Cisco CallManager cluster.
- There are no hard limits in the VoIP Monitor Server on the number of calls that can be monitored by the server, nor on the number of supervisors; its capacity is limited only by the hardware.
- The VoIP Monitor Server must be on the same VLAN as the agent phones, and it requires an available Switched Port Analyzer (SPAN) port. The VoIP Monitor Server and agent phones cannot be separated by a WAN, but they can be on different Cisco Catalyst switches if those switches support SPAN. Otherwise, voice monitoring and recording will not work.

- The VoIP Monitor Server can support up to 256 simultaneous calls and up to 32 simultaneous monitoring and recording sessions. A single recording application for the desktop client can have up to 16 simultaneous recordings.
- The VoIP Monitor Server can monitor IP phones connected to a Cisco CallManager. It can also monitor: (a) the agent desktop softphone if Cisco Media Termination Service (MTS) is installed or (b) the Cisco IP SoftPhone when used for an agent's IPCC Express extension.
- The Ethernet port for the VoIP Monitor Server must be manually configured to monitor all ports connected to agent IP phones as source ports. If the voice packets going to and from an agent's IP phone are not sent to the VoIP Monitor Server's port for any reason, that conversation will not be available to the supervisor. Any attempts to run silent monitoring will return an error message, and attempts to record the call will not succeed (even if there is no error message).
- The SPAN or port monitoring requirement can create network design issues under the following conditions:
  - The VoIP Monitor Server resides on a different VLAN than the agent phones (for example, in a server farm). Silent monitoring and recording works only when the VoIP Monitor port is a member of the same VLAN as the port being monitored.
  - The voice VLAN is trunked back to a distribution switch. A monitor port cannot be a multi-VLAN port or a trunk port. If the phones reside on a remote switch (that is, not on the same switch as the VoIP Monitor) and the voice traffic runs over a trunk, it is necessary to use Remote Switched Port Analyzer (RSPAN). RSPAN allows phones to reside on a downstream switch, but a Cisco Catalyst 6000 or 4000 is required at the access layer to take advantage of this topology. RSPAN is not supported on a Cisco 2900 or 3500 switch. Furthermore, the VoIP Monitor and agent phones must still reside in the same VLAN.

Note Silent Monitoring and Recording was designed to work in a deployment where supervisors and agents are located at a single site or at remote sites. A remote supervisor can silently monitor and record conversations between callers and agents as long as the agents and the VoIP Monitor Server are on the same VLAN segment.

IPCC Express Historical Reporting Considerations

Requests for historical reports are generated on a desktop client where historical reporting is installed. However, historical reports are generated by querying the IPCC Express or IP IVR database running on the IPCC Express server. This poses some additional overhead on CPU utilization, depending upon the size of the dataset requested.

Requests for large historical datasets can affect network performance, depending on bandwidth constraints, and such large requests can affect the ability of agents to monitor their phones during peak work hours. This effect is particularly noticeable where the IPCC Express traffic is marked with a default Type of Service (ToS) setting (for example, TOS0) and where QoS is not properly provisioned in the network. See the section on Multi-Site WAN Deployment with Centralized Call Processing, page 2-5, for more details on the network impact of running historical reporting across a WAN connection.

If you run Historical Reporting sessions on the Cisco IPCC Express or IP IVR server, please consult the entry on Standard or Enhanced Historical Reporting in Table 6-3.

Sizing the Cisco CallManager Servers

Note A new Cisco CallManager sizing tool for IPCC Express is being developed, but is not yet available at the time this Design Guide must go to press. Updated information about Cisco CallManager sizing with IPCC Express will be added to the next version of this guide. When complete, the new CallManager sizing tool will also be available on line.

This chapter documents general best practices and scalability considerations for sizing the Cisco CallManager servers used with your IPCC Express deployment. Within the context of this document, scalability refers to server capacity of both the IPCC Express server and the Cisco CallManager server handling CTI messages and call processing for IPCC Express.

Before applying the guidelines presented in this chapter, you should perform the following steps:

- Determine customer call center application requirements.
- Calculate call sizing estimates.
- Determine the types of call center resources used in IPCC Express.
- Also determine the equivalent call resource in Cisco CallManager.
- Select a deployment model.
- Select a single or multi-server IPCC Express deployment scenario.

To proceed with sizing the Cisco CallManager servers, you will also need the following information:

- Number of required IPCC Express agents
- Number of required IVR or IPCC Express CTI ports
- Number of PSTN lines
- Estimated Busy Hour Call Attempts (BHCA) rate for all agents

Impact of IPCC Express on Cisco CallManager Scalability

The Configuration and Ordering Tool (available online at http://www.cisco.com/en/US/partner/products/sw/custcosw/ps1846/prod_how_to_order.html) uses a point-weighting system for sizing the IPCC Express server. At this time, the Configuration and Ordering Tool does not include server sizing information for co-located Cisco CallManager and IPCC Express configurations.

The IPCC Express point system can help determine system capacity for the IPCC Express server. However, the IPCC Express server adds a second, indirect impact on the Cisco CallManager and CTI Manager servers through the JTAPI subsystem of IPCC Express. This impact can be attributed to the following interactions:

- Call processing requests between the IPCC Express and Cisco CallManager servers, such as transferring calls to agents or to an extension, or placing a call on hold while in queue.
- Third-party monitoring of agents.
- Call processing requests between an individual desktop agent and Cisco CallManager. These requests (for example, an agent transferring a call to another extension) take the form of CTI messages directly to Cisco CallManager, originating from the desktop agent software.
- Transcoding resources if G.729 is used for RTP streaming to agents.

Each device that registers with Cisco CallManager has a weight measured in terms of device units. The weight depends on:

- Device type
- Device utilization, or Busy Hour Call Attempts (BHCA)

Therefore, in addition to using the point system for IPCC Express, we have to determine scaling capacity for this indirect impact on the Cisco CallManager that serves as the primary CTI Manager.

For proper sizing of the Cisco CallManager and CTI Manager servers, you must use the latest Cisco CallManager sizing tools, which are available through your Cisco Systems Engineer (SE) or Partner.

Impact of IPCC Express on the Cisco CallManager Performance

Cisco CallManager system performance is influenced by many criteria such as:

- Software release versions
- The type and quantity of devices registered
  - CTI ports
  - Gateway (GW) ports
  - Agent phones
  - Route points
  - CTI Manager, etc.
- The load processed by these devices (calls per second)
- Application call flow complexity
  - IVR self service
  - Call treatment
  - Routing to agents
- Special CCM configuration and services
  - MOH
  - Tracing levels, etc.
- Server platform type

- Standard
- High Performance

The following performance observations are based on scalability and load testing of Cisco CallManager servers with the IPCC Express application. Various simple and complex call flow scenarios were used. In all tests conducted, CPU measurements were taken for each individual resource that added up to the total CPU consumption required to process the call load. The CPU percentages stated in Table 7-1, are all relative to each other regardless of load. These percentages may vary from one release to the next, and are not to be used to manually configure and size Cisco CallManager servers. The intent is to give you an idea of the relative CPU weight of each resource if it were configured/registered in one Cisco CallManager node by itself. Hence, it is important to balance all resources equally as much as possible if you are using more than one primary Cisco CallManager server. This balancing of resources prevents overloading one server at the expense of another.

Table 7-1 Effect of Major Performance factors/criteria on CCM CPU resources with IPCC and IVR applications

Performance Criterion	Effect on CCM System
CTI Route Points	• CPU resource consumption: Very low. • The portion of the CPU consumed by the CTI route points is about 3% of the total CPU consumption. This result is mainly because the route point is only redirecting calls and not terminating media (In CCM 4.x, route points are capable of terminating media and the CPU consumption will vary based on use scenarios.)
CTI Manager (JTAPI monitoring)	• CPU resource consumption: Low. • The portion of the CPU consumption by the CTI Manager averaged about 10%. This included both JTAPI users monitoring agents and JTAPI users monitoring CTI ports (all resources monitored were not in this node).
IPCC Agent Phones	• CPU resource consumption: Low. • The portion of the CPU consumed by the IPCC agent phones is about 10% of the total CPU consumption. This only includes the phone portion similar to any other IP Phone.
IVR ports used with either self service or with IPCC Express for call treatment before transferring calls to agents or other devices.	• CPU resource consumption: Moderate. • The portion of the CPU consumed by the CTI ports was 20% on average of the total CPU consumption. This does not include the CTI Manager portion of the application monitoring the CTI ports (see CTI Manager)
GW ports	• CPU resource consumption: Heavy. • The portion of the CPU consumed by the gateway ports is about 56% of the total CPU consumption.
Total CPU load distribution for all the above resources	Distribution of CPU load for the above five major resources based on where they were registered: • 3%—Incoming Route Points • 10%—CTI Manager/JTAPI • 10%—Agent Phones • 20%—CTI Ports • 57%—Gateway Ports

Additional Performance Considerations

In addition to the major performance factors described in Table 7-1, there are additional performance considerations worth noting:

- Busy-hour call completion rate

  As the call rate increases, more CPU resources are consumed on the Cisco CallManager server.

- Average Call duration:

  Longer average call duration means a lower busy-hour call completion rate, which lowers CPU usage.

- Call Flow Complexity:
  - Simple Call flows are those that do not involve multiple call handling. e.g. IVR self Service, incoming calls from a gateway directly to a phone, internal calls and so forth.
  - Complex Call Flows are those that involve multiple call redirects and call handling of the original call. e.g. Incoming calls to central rout points redirected to a CTI route points and then to IPIVR for call treatment then transferred or redirected to another target such as an agent. These multiple call processing segments of the original call consume more CPU resources compared to simple call handling.
  - CPU consumption varies by type of call flow. For simple call flows the CPU consumption is moderate. However CPU consumption for complex call flows is much higher.

    Tests conducted with Complex call flow (call treatment then transfer to agents) using IPIVR with H323 gateways show 62% CPU increase compared to the same call flow using ISN (H.323 GW). This is due to the fact that ISN does not require calls to be routed to CCM first before call treatment; CCM is involved only when calls are transferred to agents (simple call handling). The trade-off is that ISN gateways have increased performance demands. Similarly, complex call flows using IPIVR with MGCP gateways show 34% CPU increase compared to the same call flow using ISN (H.323 GW).

- Trace enabled:
  - CCM CPU resource consumption varies depending on trace level enabled. Changing trace level from Default to Full on CCM 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, CCM release, applications installed; call flow complexity and so on.

Similar profiling has been conducted as well for Memory consumption and Disk I/O resources that are accounted for in the CallManager Capacity Tool.

Note For proper sizing of the Cisco CallManager servers, the latest Cisco CallManager Capacity Tool must be used. Check with your Cisco Systems Engineer (SE) for help in sizing the Cisco CallManager servers.

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 IPCC Express applications.

Estimating Bandwidth Consumption

Bandwidth plays a large role in deployments involving:

- The centralized call processing model (IPCC Express at either the central site or remote sites)
- Any call deployment model that uses call admission control or a gatekeeper

Remote Agent Traffic Profile

IPCC Express signaling represents only a very small portion of control traffic (Cisco CallManager CTI and ICD subsystems) in the network. For information on TCP ports and Differentiated Services Code Point (DSCP) marking for IPCC Express ICD and CTI traffic, see the sections on Serviceability and Security, page 8-2, and QoS and Call Admission Control, page 8-4.

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 msecs.
- Impact of delays inherent in the applications themselves. 25 seconds is the initial IPCC Express agent login setup time with no WAN delay. The overall time to log in agents and base delay adds approximately 30 secs of delay per 70 msecs 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 IPCC Express agent logins.

Use Table 8-1 to estimate the number of IPCC Express 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 IPCC Express 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 8-1.

Table 8-1 Number of Remote Agents Supported by IPCC Express Across a WAN Link

Frame Relay	128 KB	256 KB	512 KB	768 KB	T1
G.729	3	7	15	25	38
G.711	N/R	N/R	N/R	N/R	14

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 8-4. For provisioning guidelines for centralized call processing deployments, refer to the *Cisco IP Telephony Solution Reference Network Design* documentation, available online at

http://www.cisco.com/warp/public/779/largeent/it/ese/srnd.html

Serviceability and Security

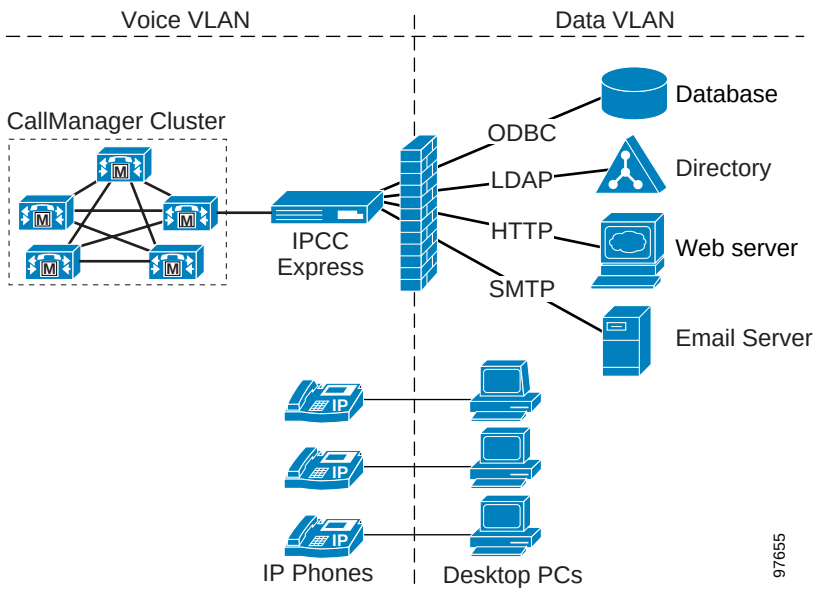
Security can be implemented on many levels. Applications security is clearly dependent upon security implemented at the infrastructure level. For more details on security at the network infrastructure level, refer to security design considerations in the *Cisco IP Telephony Solution Reference Network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

http://www.cisco.com/warp/public/779/largeent/it/ese/srnd.html

Corporate Data Access

Aside from call routing, IPCC Express or 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 8-1 shows an example of a network where voice and data components reside in separate VLANs and are separated by a firewall.

Figure 8-1 IPCC Express Accessing Data Stores



IPCC Express can communicate with these external sources through its subsystems, provided Network Address Translation (NAT) is not used. Table 8-2 lists the functionality supported on the IPCC Express and IP IVR Servers and Table 8-3 lists the various IPCC Express subsystems, the interfaces used for this communication, and the common ports associated with these interfaces.

Table 8-2 Enabled Ports on the IPCC Express Server and the IP IVR Server

Common Functionality	Port	Comments
Telnet	TCP 23	On by default
HTTP	TCP 80	Required for system maintenance
SVCHOST	TCP 135	Windows Service Loader
NETBIOS-SSN	TCP139	NETBIOS Session Service
HTTPS/SSL	TCP 443	
SMB	TCP 445	Microsoft CIFS
MS SQL	TCP 1042	SQL Server Process
RMI	TCP 1099	RMI Service
JDBC SQL	TCP 1433	

Table 8-3 Application Features Supported by IPCC Express Subsystems

Common Functionality	IPCC Express Subsystem	Interface / Protocol	Port	Comments
Call handling	JTAPI	CTIQBE	TCP 2748	
Call queuing	JTAPI	CTIQBE	TCP 2748	
	ICD	None		Queuing functions are processed internal to IPCC Express.

Table 8-3 Application Features Supported by IPCC Express Subsystems (continued)

Common Functionality	IPCC Express Subsystem	Interface / Protocol	Port	Comments
IP phone agent login	HTTP	HTTP	TCP 8080	The Cisco Agent Desktop agent login via the phone display is an IP phone service linked to an HTTP trigger on IPCC Express.
Prompt recording and playing	JTAPI	CTIQBE	TCP 2748	
	Media	Media	UDP	Media resources require CTI ports for streaming connections.
Database query for account verification or order status	Database (JDBC)	JDBC / SQL	TCP 1433	
Enterprise data to be sent to Cisco Agent Desktop	Enterprise data	CORBA	TCP port ranges: 59000, 59002-59004, 59010-59011, 59020-59021[1]	Used by Get/Set Enterprise Server Data steps.
Agent login and state control	ICD	IPCC Express CTI (subset of GED-188)	TCP 42027	Port number is configurable.
Web-enabled processing	HTTP	HTTP	TCP 8080	TCP 8080 is the default setting for the Apache Tomcat Servlet Engine, but it can be modified.
Email notification	Email	SMTP	TCP 25	
Custom Java classes		RMI	TCP 1099	TCP 1099 is the default port for Java Remote Method Invocation (RMI) calls.
IPCC Express real-time reports		RMI	TCP 1099	

1. For more information, refer to the Cisco Agent Desktop and Supervisor Desktop documentation available online at http://www.cisco.com/univercd/cc/td/doc/product/voice/sw_ap_to/apps_3_5/english/.

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, IPCC Express 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 IPCC Express 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 IPCC Express and Application-Related Traffic

Table 8-4 and the following section list TCP ports and DSCP markings for use in prioritizing IPCC Express and Cisco 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 Quality of Service design guide available online at

http://www.cisco.com/warp/public/779/largeent/it/ese/srnd.html.

Table 8-4 QoS Classifications for IPCC Express Interfaces

IPCC Express Component	Interface / Protocol	Port	DSCP Marking
CTI messaging between IPCC Express JTAPI subsystem and Cisco CallManager (both directions)	CTIQBE	TCP 2748	AF31 or EF
CTI (JTAPI) messaging from Cisco Agent Desktop to Cisco CallManager	CTIQBE	TCP 2748	None
HTTP	HTTP	TCP 8080	None
Database	JDBC / ODBC	TCP 1433	None
E-mail	SMTP	TCP 25	None
Messaging data between IPCC Express and Cisco Agent Desktop	CTI	TCP 42027	None

As Table 8-4 shows, CTI signaling is the only traffic that is automatically DSCP marked. This marking can impact the overall response of the IPCC Express application, depending upon what processing is done prior to connecting calls to an agent. For example, suppose that the IPCC Express application requires a database query to extract and pass the account information as enterprise data to the agent desktop. Because the ODBC and RMI data are not marked, they would be tagged as *best effort*.

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

Criterion	Cisco MCS-7845H-2.4 CC1 (dual CPU, using Windows 2000 Advanced Server OS) call duration < 2 minutes	Cisco MCS-7845H-2.4 CC1 (dual CPU, using Windows 2000 Advanced Server OS) call duration > 2 minutes	All Other Supported Servers
Number of agents	150	200	75
Number of supervisors (If a supervisor takes a call, the supervisor counts as an agent.)	30	32	10
Number of IVR ports	300	300	150
Number of automatic speech recognition (ASR) ports	100	100	50
Number of Vocalizer 1.0 text-to-speech (TTS) ports	200	200	50
Number of Vocalizer 3.0 TTS ports on single CRS server with 1 language and *no* ASR	100	100	25
Number of Vocalizer 3.0 TTS ports on single CRS server with 1 language and ASR	50	50	Not supported
Number of Vocalizer 3.0 TTS ports on standalone server with 1 language (ASR on a separate server)	160	160	40
Number of Vocalizer 3.0 TTS ports on standalone server with 2 languages (ASR on separate server)	80	80	20
Number of Contact Service Queues (CSQs)	75	100	25
Number of skills	100	100	50

Table A-1 Server Capacities and Limits

Criterion	Cisco MCS-7845H-2.4 CC1 (dual CPU, using Windows 2000 Advanced Server OS) call duration < 2 minutes	Cisco MCS-7845H-2.4 CC1 (dual CPU, using Windows 2000 Advanced Server OS) call duration > 2 minutes	All Other Supported Servers
Number of skills that an agent can associate with	50	50	50
Number of CSQs that an agent can associate with	25	25	25
Number of skills that a CSQ can associate with	50	50	50
Number of CSQs that a call can queue for	25	25	25
Number of simultaneous Historical Reporting sessions	2	2	1
Number of simultaneous recording sessions	32	32	16
Number of simultaneous silent monitoring sessions	32	32	16

These limits apply to an entire system. Do not exceed them even if a dedicated Call Statistics, Recording, and Monitoring Server or dedicated Call Monitoring Servers are installed.

Please use the Configuration and Ordering tool as the final authority on limits for the servers. The IPCC Express Configuration and Ordering Tool, is available online at

http://www.cisco.com/en/US/partner/products/sw/custcosw/ps1846/prod_how_to_order.html

Voice Over IP Monitoring

The VoIP monitor service for IPCC Express is targeted for the Cisco line of Catalyst switches. It relies on a Switched Port Analyzer (SPAN) session configured on the switch, which allows the IP traffic from one or more ports to be copied and sent to a single destination port. The switch issues known at this time are described in this appendex. For more information, consult the *Voice Over IP Monitoring Best Practices Deployment Guide*.

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, 2820, 2900, 2900XL, 2926GS, 2926F, 2926T, 2948G, 2950, 2980G, 3000, 3100, 3200, 3500XL, 3524-PWR XL, 3508GL XL, 3550, 5000, 5002, 5500, 5505, 5509

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

The following table gives the limits to the number of SPAN and RSPAN sessions that can exist on a switch:

Switch Model	Maximum SPAN Sessions Allowed
1200	1
1900	1
2820	1
2900	1
2900XL	1
2926GS	5
2926GL	5
2926T	5
2926F	5
2948G	5
2950	1
2980G	5
3000	1
3100	1
3200	1
3500XL	1
3524-PWR XL	1
3508GL XL	1
3550	2
4003	5
4006	5
2912G	5
5000	5
5002	5
5500	5
5505	5
5509	5
6006	30
6009	30
6506	30
6509	30
6513	30

A

B

C

D

E

F

G

H

I

J

L

M

O

P

Q

R

S

T

V